

Faculté de philosophie, arts et lettres

Commission doctorale du domaine « Langues, Lettres et Traductologie » (CDL)

Institut Langage et Communication (ILC)

Centre of English Corpus Linguistics (CECL)

Centre Valibel – Discours et Variation

The acquisition of phraseological units by
French-speaking learners of English and Dutch in
CLIL and NON-CLIL settings
Exposure effects on range and accuracy

Amélie BULON

Thèse réalisée en vue de l'obtention du grade de

Docteur en langues, lettres et traductologie

Membres du jury :

Prof. Fanny Meunier (UCLouvain), Promotrice

Prof. Kristel Van Goethem (UCLouvain), Co-promotrice

Prof. Benoit Galand (UCLouvain), Président du jury

Prof. Jeanine Treffers-Daller (University of Reading), lectrice externe

Prof. Ana Llinares (Universidad Autónoma de Madrid), lectrice externe

Louvain-la-Neuve

Octobre 2019

Abstract

The present doctoral dissertation investigates the acquisition of the phraseological lexicon in second/foreign language acquisition (S/FLA). The term phraseology refers here to the large number of multiword combinations that compose the language; as for the phraseological lexicon, it refers to the inventory of multiword combinations that users or learners of a language know receptively and/or use productively. Corpus-based analyses have demonstrated that these combinations are actually pervasive in the language, be it written or spoken. However, the acquisition of target language (TL) phraseology seems to lag behind other areas of TL acquisition (Laufer & Waldman, 2011) and even advanced learners struggle in their use of phraseological language (e.g. Meunier & Granger, 2008; Siepmann, 2008). Patterns of overuse, underuse, and erroneous phraseological productions have been reported by the majority of studies looking at TL phraseology. The fact that phraseology is not always subject to explicit instruction (Meunier, 2012) and that implicit learning has its limits in instructed TL settings may further complicate its acquisition process. Moreover, error analyses of TL learners' phraseological productions usually point to interferences of the mother tongue (L1). Psycholinguistic aspects such as frequency, prototypicality, transparency, saliency, etc. of the word combination itself will also facilitate or impede its acquisition, depending on SLA variables and contexts (Ellis, 2008b).

This thesis is framed within a multidisciplinary research project on *Content and Language Integrated Learning* (CLIL) in French-speaking Belgium (cf. Hiligsmann et al., 2017) and examines the relation between the CLIL approach and the (longitudinal) acquisition of phraseological units by learners. The sample consists of French-speaking upper secondary school learners of English in a CLIL ($n=90$) or a traditional foreign language (NON-CLIL) setting ($n=90$), or Dutch in a CLIL ($n=132$) or a NON-CLIL setting ($n=100$). The learners were followed over the course of two consecutive school years (2015-2016 and 2016-2017).

Previous research on vocabulary acquisition in CLIL settings rarely went beyond the single word level and studies looking at the acquisition of phraseology by TL learners in general usually include only one or two types of multiword combinations. In addition, English remains the most frequently analyzed TL, also in the field of phraseology. The present study broadens the scope of analysis by scrutinizing a large set of phraseological units (referential phrasemes, textual phrasemes and communicative phrasemes) in two different target languages (English and Dutch) and in two different instructed settings.

A combination of corpus-based analyses and experimental methods are used in order to assess the learners' phraseological knowledge. The corpus analyses are based on a corpus of semi-spontaneous written productions that was collected to analyze the learners' productive use of word combinations. The written corpus data were collected at two different time points (beginning of the fifth year of secondary school and end of the sixth year of secondary school), so that a longitudinal analysis of the learners' written phraseological development could be carried out. The experimental test consists in a multiple-choice test designed to evaluate the learners' receptive knowledge of verb-noun collocations. The influence of other variables such as duration of TL instruction, frequency of informal contact with the TL outside school, general productive and receptive vocabulary knowledge, non-verbal intelligence (IQ) and socio-economic status (SES) was also considered in order to isolate the effect of CLIL.

Since the CLIL approach appears to be beneficial for S/FLA considering its increased provision of TL input in more authentic situations and its interactional and communicative nature, we may expect a better and faster acquisition of phraseological units – both at the productive and receptive level – by CLIL learners of English or Dutch compared to their NON-CLIL peers. More specifically, we hypothesize that CLIL learners will produce a larger number of multiword combinations and that the phraseological productions will show a higher level of accuracy. We also expect them to score higher at the receptive collocation test. In terms of longitudinal development, we believe that the two learner groups will increase their phraseological competence and that the linguistic gain will be even more substantial in the CLIL group.

On the whole, the results of the present doctoral research corroborate our hypotheses. A higher phraseological productive and receptive knowledge was observed in the CLIL group. The longitudinal results were however more mixed since the CLIL learners either positively progressed, but to a lesser extent than the NON-CLIL learners (in the case of English), or even decreased (in the case of Dutch). It is important to note that the results often differed according to the TL, in the sense that the difference between the CLIL and the NON-CLIL learners appeared to be larger in TL Dutch than in TL English.

It has to be acknowledged that the apparent positive influence of the CLIL approach on the acquisition of phraseology observed in this research may not be the sole factor at work. Our study also showed that other variables were at play, like for instance the impact of other input-related and vocabulary-related measures, and background variables on the learners' receptive collocational knowledge. Other factors like the motivation of the learners and their attitudes

towards the TL (especially in the context of Belgium where Dutch and English tend to have a different status), the teachers' proficiency in the TL as well as their teaching practices are undoubtedly important aspects to consider.

Acknowledgments

First of all, I would sincerely like to thank my supervisor, Prof. Fanny Meunier, for her guidance, feedback, enthusiasm and support throughout the process of writing this doctoral dissertation. She managed to let me work independently, and to be there whenever I was concerned or ran into a problem. She has a talent for putting things into perspective and each meeting with her cheered me up! Beyond being a model supervisor, she is also a very nice and generous person with a great sense of humor.

Ik ben ook dankbaar voor de steun van mijn copromotor, Prof. Kristel Van Goethem. Ze was beschikbaar als ik het nodig had en heeft me altijd goed en constructief advies gegeven. Ze is ook een zeer attent persoon en ik bewaar goede herinneringen aan onze wandeling door Verona en aan de vruchtbare vergadering bij jou thuis.

I also owe thanks to the other members of my jury, namely Prof. Benoit Galand, Prof. Jeanine Treffers-Daller and Prof. Ana Llinares, for their most constructive comments.

I am also grateful to the whole ‘ARC team’: Prof. Philippe Hiligsmann, Prof. Fanny Meunier, Prof. Kristel Van Goethem, Prof. Laurence Mettwie, Prof. Benoit Galand, Prof. Arnaud Szmalec, Dr. Luk Van Mensel, Audrey De Smet, Isa Hendrikx and Morgane Simonis. Thank you Philippe for initiating this ambitious and innovative project, and thank you all for creating such a good working atmosphere. This project is yet another illustration that there is strength in unity! Now that this is coming to an end, I realize how lucky I was to be part of this incredible team. I hereby also want to acknowledge the ARC 14/19-061, UCLouvain and the Fédération Wallonie-Bruxelles (CfB) for funding this research project.

One of the perks of working in a multidisciplinary environment is that you easily find all the help you need... Bedankt Luk en Isa voor jullie hulp bij de statistische analyses en bedankt Philippe en Laurence voor jullie medewerking bij het ontwerpen en pretesten van de receptieve test voor het Nederlands.

Je tiens aussi à remercier mes ‘collocs de bureau’ et amies Audrey, Isa, Mo’ et Alice.

Audrey, Isa, Mo’: Merci de m’avoir accueillie à bras ouverts et d’avoir rendu cette collaboration tellement facile. Merci pour ces pauses de midi relaxantes, ces repas d’anniversaire conviviaux et ces soupers improvisés. La nostalgie ressentie à la fin du projet en dit long sur les bons moments partagés ensemble et j’espère sincèrement qu’il y en aura d’autres !

Alice: Merci pour ta positivité et ton enthousiasme à toute épreuve; ça a été un réel plaisir de partager le bureau avec toi et je me réjouis de nous voir travailler ensemble prochainement.

I would also like to thank all the participating schools (the teachers and school principals), the pupils and their parents for accepting to take part in the project.

I am thankful to all the university students who helped us during the data collections organized in the schools and at the university of Louvain-la-Neuve. I also want to address a word of thanks to the students who did some preliminary work on part of my data, either in the framework of their master's dissertation or during their internship: Emmanuella Ahishakiye, Laura Nolevaux, Mathilde Dumont, Léticia Raemdonck, Natacha Buntinx and Jeremy Dupal.

I also want to thank Prof. Sylvie De Cock for having been kind enough to bring this PhD position to my attention.

J'aimerais terminer sur une touche plus personnelle en remerciant ma famille pour l'intérêt qu'elle a porté à mon travail et pour m'avoir aidée autant que possible. Je pense notamment aux trajets vers Bruxelles pour l'un ou l'autre évènement, vers Namur pour l'analyse de nos questionnaires, ou encore au prêt de la 'grosse' voiture pour notre conférence à Fribourg...

Table of contents

Abstract	2
Acknowledgments	5
Table of contents	7
List of abbreviations	12
INTRODUCTION	14
PART 1: THEORETICAL BACKGROUND	20
CHAPTER 1: Phraseology in second/foreign language acquisition	21
1.1. Introduction	21
1.2. Phraseology: from peripheral to central	21
1.3. Phraseological language: terminology, characteristics and definition	23
1.4. Two different approaches to phraseology	26
1.4.1. The ‘traditional’ or ‘phraseological’ approach	26
1.4.2. The ‘distributional’ or ‘frequency-based’ approach.....	27
1.4.3. Reconciling the two approaches	28
1.5. Analyzing phraseology: methodological approaches	29
1.5.1. Learner Corpus Research (LCR)	29
1.5.1.1. Learner corpora and phraseology	30
1.5.1.2. Contrastive Interlanguage Analysis (CIA)	32
1.5.1.3. Longitudinal analyses	33
1.5.2. Experimental analyses	35
1.5.3. Error analyses	37
1.6. Written versus spoken phraseology.....	40
1.7. Acquiring phraseological language	41
1.7.1. Implicit and explicit learning.....	41
1.7.2. Transfer from the mother tongue.....	43
1.7.3. Psycho-linguistic aspects.....	45
1.8. Teaching phraseological language	45
1.9. Target language phraseological competence: previous findings	47
1.9.1. Productive knowledge of phraseological language	48
1.9.1.1. Written phraseological language	48
1.9.1.2. Spoken phraseological language	50
1.9.2. Receptive knowledge of phraseological language.....	50
1.9.3. Development of phraseological competence.....	52
1.10. Conclusion.....	54

CHAPTER 2: Second/foreign language acquisition from a usage-based approach.....	57
2.1. Introduction	57
2.2. First language acquisition.....	57
2.2.1. Theoretical assumptions of usage-based approaches	58
2.2.2. Theoretical assumptions of emergentist approaches	60
2.3. Second/foreign language acquisition.....	61
2.3.1. The role of input and frequency effects.....	62
2.3.2. The role of recency, salience, prototypicality, and redundancy	65
2.3.3. Focus on form(S) and focus on meaning.....	66
2.4. Important learner-related factors in second/foreign language acquisition	68
2.4.1. Age	68
2.4.2. Motivation	71
2.4.3. Cognition.....	73
2.5. Conclusion.....	75
CHAPTER 3: <i>Content and Language Integrated Learning (CLIL)</i> in French-speaking Belgium	77
3.1. Introduction	77
3.2. CLIL	77
3.3. CLIL in French-speaking Belgium.....	81
3.4. Research project “Assessing Content and Language Integrated Learning in French-speaking Belgium: an integrated approach” (2014-2019).....	88
3.5. Research line on the acquisition of phraseological language in CLIL	90
3.5.1. Rationale.....	90
3.5.2. Originality of the research design.....	91
3.6. Conclusion.....	92
PART 2: ACQUIRING PHRASEOLOGICAL LANGUAGE: RESEARCH HYPOTHESES, METHODOLOGY AND RESULTS	94
CHAPTER 4: Methodology.....	95
4.1. Introduction	95
4.2. Pilot study: global written proficiency in L1 French and TL English/Dutch	95
4.3. Research hypotheses.....	98
4.4. Participants	99
4.5. Tests, tasks, data.....	100
4.5.1. Written productions	101
4.5.2. Oral interaction.....	102
4.5.3. Multiple-choice test (receptive phraseological competence)	103
4.5.4. Measures of overall productive and receptive vocabulary knowledge in the target language.....	104

4.5.5.	Background variables	105
4.6.	Extraction and classification of phraseological units and phraseological errors	105
4.7.	Statistical analyses.....	112
4.8.	Research design.....	113
4.9.	Conclusion.....	114
CHAPTER 5: Productive phraseological competence		115
5.1.	Introduction	115
5.2.	Comparison between CLIL vs NON-CLIL learners' phraseological language at a single point in time	115
5.2.1.	Method.....	115
5.2.1.1.	Research questions and hypotheses	115
5.2.1.2.	Data and participants	116
5.2.2.	Results	120
5.2.2.1.	Range.....	120
5.2.2.1.1.	Global trends	120
5.2.2.1.2.	Frequency: comparison CLIL/NON-CLIL.....	122
5.2.2.2.	Variety	125
5.2.2.2.1.	Lexical collocations.....	126
5.2.2.2.2.	Grammatical collocations	128
5.2.2.2.3.	Phrasal verbs.....	130
5.2.2.3.	Accuracy.....	132
5.2.2.3.1.	Global trends	132
5.2.2.3.2.	Frequency: comparison CLIL/NON-CLIL.....	136
5.2.2.3.3.	Grammaticality versus acceptability errors	141
5.2.3.	Discussion and implications	143
5.2.4.	Conclusion.....	145
5.3.	Longitudinal development of the CLIL and NON-CLIL learners' phraseological language over time.....	146
5.3.1.	Research questions and hypotheses.....	146
5.3.2.	Methodology	147
5.3.2.1.	Participants	147
5.3.2.2.	Design and data	149
5.3.2.3.	Longitudinal analyses.....	151
5.3.3.	Results	151
5.3.3.1.	Intra-group analysis – frequency of use (RQ1)	151
5.3.3.2.	Intra-group analysis – accuracy (RQ2).....	155

5.3.3.3.	Inter-group analysis – frequency of use (RQ3)	167
5.3.3.4.	Inter-group analysis – accuracy (RQ4).....	172
5.3.4.	Concluding remarks and discussion	178
CHAPTER 6: Receptive phraseological competence		181
6.1.	Introduction	181
6.2.	Research questions and hypotheses.....	182
6.3.	Methodology	182
6.3.1.	Participants	182
6.3.2.	Receptive collocation test.....	185
6.4.	Results	188
6.4.1.	Research question 1: do CLIL and NON-CLIL learners differ in their receptive knowledge of verb-noun collocations?.....	188
6.4.2.	Research question 2: Impact of the collocation.....	193
6.4.2.1.	Analysis of the impact of collocation frequency on receptive collocation knowledge.....	197
6.4.2.2.	Qualitative analysis of the collocation in terms of degree of similarity with the L1 (congruence).....	199
6.5.	Concluding remarks and discussion	223
CHAPTER 7: Effects of target language exposure and vocabulary knowledge on the acquisition of verb-noun collocations		228
7.1.	Introduction	228
7.2.	Measures of target language exposure	229
7.2.1.	Duration of target language instruction	229
7.2.2.	Informal target language input outside school	230
7.3.	Measures of productive and receptive vocabulary knowledge.....	232
7.4.	Non-verbal intelligence and socio-economic status	234
7.5.	Research questions and hypotheses.....	235
7.6.	Methodology	236
7.7.	Results	237
7.7.1.	Correlations between the predictors and the outcome variable	237
7.7.2.	Regression analyses.....	238
7.8.	Discussion	244
7.9.	Conclusion.....	246
CONCLUSION.....		248
Co-authored publications		258
References		259
Appendices		299

1. Writing task (English/Dutch).....	299
1.1. Guidelines e-mail about the holidays (English)	299
1.2. Guidelines e-mail about a party (English)	300
1.3. Guidelines e-mail about the holidays (Dutch).....	301
1.4. Guidelines e-mail about a party (Dutch)	302
2. Oral task (English/Dutch)	303
2.1. Guidelines exercise about the holidays (English/Dutch).....	303
2.2. Guidelines exercise about a party (English/Dutch)	304
3. Multiple-choice tests verb-noun collocations	305
3.1. English.....	305
3.2. Dutch	307
4. Text fragments illustrating the categorization of phraseological units.....	309
5. Distribution of the phraseological units produced across individual learners.....	314
6. Comparative analysis of global written complexity – statistics (Chapter 5)	315
7. Correlations between predictor and outcome variables (Chapter 8).....	316

List of abbreviations

ADJ	Adjective
ADV	Adverb
ACTFL	American Council on the Teaching of Foreign Languages
ANOVA	Analysis Of Variance
ARC	<i>Action de Recherche Concertée</i> ‘Concerted Research Action’
BNC	British National Corpus
CAF	Complexity, Accuracy and Fluency
CALP	Cognitive Academic Linguistic Proficiency
CBI	Content-Based Instruction
CDST	Complex Dynamic Systems Theory
CEB	<i>Certificat d’Etudes de Base</i> ‘Certificate of Completion of Basic Studies’
CEFR	Common European Framework of Reference for Languages
CESS	<i>Certificat de l’Enseignement Secondaire Supérieur</i> ‘Certificate of Upper Secondary Education’
CfB	<i>Communauté Francophone de Belgique/Fédération Wallonie-Bruxelles</i> ‘French-speaking Community of Belgium’
CHN	<i>Corpus Hedendaags Nederlands</i> ‘Corpus of Contemporary Dutch’
CIA	Contrastive Interlanguage Analysis
CLIL	Content and Language Integrated Learning
COLLEX	Collocating Lexis
COLLMATCH	Collocate Matching
Cor.	Corrected (text)
CPH	Critical Period Hypothesis
EAP	English for Academic Purposes
EFL	English as a Foreign Language
EMI	English Medium Instruction
EMILE	<i>Enseignement de Matières par l’Intégration d’une Langue Etrangère</i> ‘Teaching of a subject through the integration of a foreign language’ (= CLIL)
ESL	English as a Second Language
freq	Frequency
INF	Infinitive
IQ	Intellectual Quotient
KWIC	Keyword In Context
L1	First Language
L2	Second Language
LCR	Learner Corpus Research
MI	Mutual Information
Morphs	Morphemes
MTLD	Measure of Textual Lexical Diversity

MWU	Multi-Word Unit
N	Noun
Nb/nb	Number
occ.	Occurrence
Orig.	Original (text)
POS	Part-Of-Speech
PPVT-III-NL	Peabody Picture Vocabulary Test-III-NL
PPVT-IV	Peabody Picture Vocabulary Test, Fourth Edition
PREP	Preposition
PV	Phrasal Verb
RQ	Research Question
SAT	Standardized Achievement Test
SES	Socio-Economic Status
S/FLA	Second/Foreign Language Acquisition
SLA	Second Language Acquisition
STIMOB	<i>Stimulerend Meertalig Onderwijs in Brussel</i> ‘Stimulating Multilingual Education in Brussels’
TL	Target Language
T-Scan	<i>Software voor Complexiteits-Analyse van het Nederlands</i> ‘Software for the Analysis of Complexity in Dutch’
TTR	Type-Token Ratio
UG	Universal Grammar
V	Verb
UCLouvain	Université catholique de Louvain
UNamur	Université de Namur

INTRODUCTION

In times of globalization, policies increasingly promote multilingualism as a strong social and economic asset. One way to foster multilingualism in education is *Content and Language Integrated Learning* (CLIL), a pedagogical approach in which some school subjects (e.g. history, geography, mathematics, etc.) are taught with and through a different language than the mainstream school language. In the French-speaking Community of Belgium, schools have been allowed to provide CLIL education in Dutch, English or German since 1998. To this day, however, we only have an incomplete and fragmented view on how CLIL differs from the traditional language learning form of education (NON-CLIL) and on how it affects second/foreign language acquisition (S/FLA). A multidisciplinary ARC project (*Action de Recherche Concertée* – Concerted Research Action) called *Assessing Content and Language Integrated Learning: Linguistic, Cognitive and Educational Perspectives* was launched by a team of researchers from the Université catholique de Louvain (UCLouvain) and the Université de Namur (UNamur) in order to gain more insight on CLIL in French-speaking Belgium (cf. Hiligsmann et al., 2017). As the project's title suggests, it involves different research lines: linguistic, cognitive, educational, and socio-affective. The present doctoral research constitutes an integral part of this project and focuses on the acquisition of target language (TL) English and Dutch phraseological units by CLIL and NON-CLIL French-speaking learners.

This doctoral dissertation primarily falls into the field of phraseology. Corpus-based analyses of authentic language brought its phraseological character to light; a large proportion of language produced by native speakers is actually 'prefabricated', which means that proficient language users rarely produce utterances word by word, but mostly resort to chunks that seem to be stored that way in their mental lexicon, as suggested by research in psycholinguistics. S/FLA studies looking at the linguistic competence of learners of a second or foreign language agree to say that phraseological competence remains a stumbling block; even advanced learners struggle to be fluent and to produce a language that sounds natural and authentic because of their lack of intuition about the way words combine together in the second/foreign language. A large amount of input seems necessary to be able to use appropriately the many and varied word combinations that exist in a language.

The present work on learners' phraseological acquisition actually fits quite well into the usage-based and emergentist theories of language acquisition and this for several reasons. First, these theories consider language input as the driving force of language acquisition and even though their principles result from research on child language acquisition, evidence has shown that

similar learning mechanisms operate for older learners learning a TL. Since CLIL settings are believed to provide learners with more substantial input than NON-CLIL settings, one may expect them to enhance the learners' acquisition of phraseological units. Second, usage-based approaches are particularly relevant for the analysis of phraseology as they acknowledge the importance of multi-word units in language learning; formulas are actually considered as the first building blocks in the language construction process. In addition, they reject the lexis-grammar dichotomy and multi-word units are located at the interface of lexicon and grammar.

As shown in the summary table below, this doctoral dissertation aims to answer three main research hypotheses that are further broken down into specific research hypotheses in each empirical chapter. All of them focus on the TL phraseological competence of CLIL and NON-CLIL learners. Some of them relate to case studies addressing the effect of other variables on the acquisition of phraseological units (the frequency of the collocation in Chapter 6, and formal/informal TL exposure, general vocabulary knowledge, socio-economic background and non-verbal intelligence in Chapter 7). The main hypotheses were formulated on the basis of previous research showing the advantages of the CLIL approach for S/FLA and the results of a pilot study that had been carried out beforehand. The pilot study investigated the learners' global written competence in their TL, but also in their L1, as the effect of CLIL on the mother tongue appears to be a key issue in CLIL research (see Chapter 3, Section 3.2.).

Pilot study's research hypotheses: • We expect CLIL learners to show a higher level of global written competence in the TL English/Dutch; • We expect CLIL and NON-CLIL learners to perform equally well in their L1 (French).
Research hypothesis (1): We hypothesize that the TL productions of CLIL learners will be more phraseological (and target-like) than that of NON-CLIL learners, i.e. they will produce more phraseological units and display higher phraseological accuracy. <i>Specific research hypotheses (Chapter 5):</i> • The phraseological language of CLIL learners will display a greater range of phraseological units; • The phraseological language of CLIL learners will be more accurate.
Research hypothesis (2):

We hypothesize that both the CLIL and the NON-CLIL learners will show improvement in their phraseological language after a period of two consecutive school years, with CLIL learners displaying greater progress.

Specific research hypotheses (Chapter 5):

- The CLIL and the NON-CLIL learners will perform better at Time 2, both in terms of frequency of use and accuracy;
- The CLIL learners' TL phraseological language will improve more than that of NON-CLIL learners, both in terms of frequency of use and accuracy.

Research hypothesis (3):

We hypothesize that the receptive phraseological knowledge of CLIL learners will be greater than that of NON-CLIL learners.

Specific research hypotheses (Chapter 6):

- The CLIL learners will perform better at the receptive test;
- The frequency of the collocation and its degree of similarity with the L1 (congruence) will have an impact on the results.

Specific research hypotheses (Chapter 7):

- CLIL is positively related to the learners' receptive phraseological knowledge;
- Formal and informal TL exposure are positively related to the learners' receptive phraseological knowledge;
- General vocabulary knowledge is positively related to the learners' receptive phraseological knowledge;
- The learners' socio-economic status and non-verbal intelligence are positively related to the learners' receptive phraseological knowledge.

These research hypotheses are addressed using a combination of corpus-based analyses and experimental methods. Both the productive and receptive phraseological knowledge of the learners are scrutinized and the collection of longitudinal data will allow for an analysis of the learners' linguistic development. The influence of a number of predictor variables is also considered. The investigation of the learners' phraseological productive knowledge is not limited to one specific type of word combination (as it is often the case in phraseological research), but takes into account a large array of multi-word units. In addition, unlike previous research on the acquisition of phraseology in a TL that mostly focused on English, the present

study also looks at Dutch, which is one of the official languages in Belgium. All these features contribute to the originality of the present doctoral dissertation.

Another innovative aspect of the present doctoral research is that it pays attention to the phraseological competence of CLIL learners, and therefore looks at their linguistic competence beyond the word level. The large number of studies on CLIL has mainly been concerned with various aspects of language proficiency such as vocabulary (at the single word level) (e.g. Zydatiś, 2007; Jexenflicker & Dalton-Puffer, 2010; Lo & Murphy, 2010; Ruiz de Zarobe, 2010), writing skills (including morphosyntactic aspects, spelling, tense, pragmatic awareness, etc.) (e.g. Jexenflicker & Dalton-Puffer, 2010; Ruiz de Zarobe, 2010), or oral skills (including fluency and phonetics) (e.g. Admiraal, Westhoff & de Bot, 2006; Zydatiś, 2007; Lasagabaster, 2008; Ruiz de Zarobe, 2008). Research into the lexical competence of CLIL learners at the phraseological level is scarce.

The present doctoral dissertation is organized in two main parts and seven chapters. The first part (“Part 1: Theoretical background”) is devoted to the theoretical contextualization of the study and includes three theoretical chapters. The second part (“Part 2: Acquiring phraseological units: research hypotheses, methodology and results”) consists of a methodological chapter and three empirical chapters.

Chapter 1 entitled “Phraseology in second/foreign language acquisition” first reviews the current issues in the field of research on phraseology – its definition, identification, variation according to the communication mode as well as its acquisition and teaching – and then discusses the methodological approaches used for the analysis of phraseological language. The chapter also reports results from previous research conducted on the phraseological competence of TL learners.

Chapter 2 (“Second/foreign language acquisition”) presents the important theoretical principles to language learning that frame the present research. It introduces the major tenets of the usage-based and emergentist approaches to first language acquisition and their applications to S/FLA. It also discusses important learner-related factors such as age, motivation, and cognition.

In Chapter 3, “Content and Language Integrated Learning (CLIL) in French-speaking Belgium”, we first introduce CLIL and draw attention to a number of differences between this educational approach and other types of immersion programs and mainstream target language learning contexts. We then focus more closely on the situation in French-Belgium and provide a short overview of studies that were previously carried out on CLIL in French-speaking

Belgium. After that, we describe the main objectives and research lines of the overarching multidisciplinary project on CLIL in which the present doctoral research is framed. Finally, we explain the contribution of the present thesis to the overall project.

Chapter 4 deals with the methodological aspects of the study. First, the results of a pilot study investigating the global written competence of CLIL and NON-CLIL learners in their target languages English and Dutch and their first language (L1) French are described. Then, the main research hypotheses are formulated on the basis of the pilot study's findings and the theoretical chapters outlined above. We also provide a detailed description of the participants who took part in the study, the different types of data that were collected, the taxonomy of phraseological units that was used for the extraction and analysis of the learners' phraseological language, and the research design.

In Chapter 5, "Productive phraseological competence", we investigate the productive phraseological competence of the CLIL and NON-CLIL learners in the two target languages on the basis of corpus data in the form of semi-spontaneous written productions. The chapter also includes an analysis of the longitudinal development of the learners' phraseological productions. The analyses include quantitative measures of frequency, variety and accuracy.

Chapter 6 entitled "Receptive phraseological competence"¹ describes the findings of a study on the receptive knowledge of English and Dutch verb-noun collocations by CLIL and NON-CLIL learners. The learners' receptive collocational knowledge was tested using a receptive test in the form of a multiple-choice exercise.

Chapter 7 ("Effects of target language exposure and vocabulary knowledge on the acquisition of phraseological units") is a follow-up study to the previous chapter and aims to disentangle the effect of CLIL from other types of TL input (number of years of target language instruction and frequency of informal contact with the target language outside school). The relation between the learners' overall productive and receptive vocabulary knowledge and receptive collocational knowledge is also assessed. Two variables related to the learners' individual background (socio-economic background and non-verbal intelligence) are also controlled for.

¹ The fact that the analysis of the learners' receptive phraseological knowledge is dealt with after the learners' productive phraseological knowledge is explained by the chronological order of the data collections (the productive data were gathered before the receptive data). In addition, while the investigation of the productive data involves a wide range of phraseological units, the receptive data focus on only one specific type of lexical collocations, i.e. verb-noun collocations.

The doctoral dissertation ends with a recapitulative overview of the results as well as a discussion of the findings with respect to other outcomes produced within the framework of the encompassing research project. A number of limitations are exposed and future avenues for research are provided.

PART 1: THEORETICAL BACKGROUND

Part 1 of this doctoral dissertation is devoted to the theoretical background. Chapter 1 focuses on phraseology in second/foreign language acquisition (S/FLA). More specifically, it tackles various issues such as the operationalization, the analysis or the teaching of phraseological units, and reports findings from previous work on the phraseological competence of (CLIL) TL learners. Chapter 2 presents the theoretical framing of our research, that is the usage-based and emergentist approaches to language acquisition, and to S/FLA in particular. A number of learner-related factors that play an important role in the acquisition of a TL are also introduced. Chapter 3 presents the *Content and Language Integrated Learning* (CLIL) educational approach and discusses its implementation in French-speaking Belgium. It also describes the Belgian research project on CLIL that encompasses this doctoral research and explains how the present work on phraseology contributes to the overarching project. These three theoretical chapters provide the basis for the research hypotheses and analyses presented in Part 2.

CHAPTER 1: Phraseology in second/foreign language acquisition

1.1. Introduction

Phraseology has long been neglected in language description (Sinclair, 2008). However, evidence provided by research in corpus linguistics, S/FLA and psycholinguistics has shown that language is largely phraseological/formulaic/idiomatic, viz. containing numerous ready-made, largely prefabricated units going beyond the single word, so that phraseology has now become a discipline in its own right, being central in a wide range of fields. Sections 1.2. to 1.4. set the scene: they discuss how phraseology has recently become established as an important field of research (Section 1.2.), tackle the issue of defining phraseological language (Section 1.3.), and present the different approaches to the identification and classification of phraseology (Section 1.4.). Then, Section 1.5. reviews the different methodological approaches to the analysis of phraseology. Section 1.6. deals with differences between spoken and written phraseology. The following two sections (Sections 1.7. and 1.8.) focus on the acquisition and the teaching of phraseological language. Section 1.9. is devoted to the results of previous research carried out on the phraseological language of TL learners. Finally, Section 1.10. provides a summary of the present chapter.

1.2. Phraseology: from peripheral to central

Structuralist linguistics, which was the dominant approach at the beginning of the twentieth century, denied the separation of lexis and linguistic context of usage. This can be illustrated by Firth's (1957) famous quote: "You shall know a word by the company it keeps". In the view of the structuralists, "language acquisition is the learning of an inventory of patterns as arrangements of words with their associated structural meanings" (Ellis, 2008b, p. 1). These structural patterns would nowadays be referred to as 'constructions' in Cognitive Linguistics, and more particularly in Construction Grammar (e.g. Goldberg, 2003; Hoffman & Trousdale, 2013), to denote "form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner's mind" (Ellis, 2008b, p. 2). Another current term would be 'phraseologisms', mostly used in Corpus Linguistics to indicate the fact that a lexical item and one or more additional linguistic elements functioning as one semantic unit co-occur more frequently than expected on the basis of chance (Howarth, 1998).

The Generative approaches and Chomsky's Universal Grammar (UG) took over in the 1960s. Structuralist patterns were abandoned in favor of grammatical phenomena, which were considered at the core of the grammar and completely separated from performance, lexis, social

usage and the rest of cognition. Patterns, structures, constructions, etc. were all seen as peripheral. Later on, the thinkers Fillmore, Goldberg and Croft brought back constructions at the center of grammar. Fillmore (1988) and Fillmore, Kay & O'Connor (1988) argued that the core phenomena could also be accounted for by the grammatical mechanisms that are necessary for the explanation of the grammatical periphery. Goldberg's Construction Grammar (1995, 2003) "argues that all grammatical phenomena can be understood as learned pairings of form (from morphemes, words, idioms, to partially lexically filled and fully general phrasal patterns) and their associated semantic or discourse functions" (Ellis, 2008b, p. 4). The central linguistic unit of the theory of Construction Grammar is the construction and is defined by Goldberg (1995, p. 4) as follows: "C is a construction iff_{def} C is a form-meaning pair <F_i, S_i> such that some aspect of F_i or some aspect of S_i is not strictly predictable from C's component parts or from other previously established constructions". Croft's Radical Construction Grammar (2001) suggested that syntactic categories and relations were not universal, but language- and construction-specific.

The Corpus Linguistic analyses of large collections of text proved the pervasiveness of recurrent patterns of words and constructions in natural language. Sinclair's Idiom Principle (1991) summarized the corpus investigations' findings in the following way: "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments" (1991, p. 100). In an attempt to quantify the phraseological nature of language, Erman and Warren (2000) assessed that about 50 per cent of language is constructed according to this principle and Leech (2000), among others, reported that spoken language seems to be even more phraseological than written language. Differences between the two communication modes are further discussed in Section 1.6. Corpus linguistics is all the more connected to phraseological research since information extracted from corpora is based on frequencies of occurrence and co-occurrence. The use of learner corpus research (LCR) as a methodological approach for the analysis of phraseology is tackled in Section 1.5.1.

The resurgence of phraseology has also been stimulated by other research fields that in some way resonate with it. This is the case of Cognitive Linguistics (e.g. Langacker, 1987; Robinson & Ellis, 2008), which recognizes symbolic units as the only kinds of element the linguistic system contains and defines a unit as a combination of a form and a meaning/function. The more often a person encounters a symbolic unit, the more fixed this unit becomes in his/her linguistic system and the more automatically it is accessed. The person does not have to "focus

his attention specifically on its individual parts for their arrangement [...] he has no need to reflect on how to put it together” (Langacker, 1987, p. 57). Furthermore, psycholinguistic research demonstrated the learner’s sensitivity to the frequencies of occurrence of different constructions in the input (Gernsbacher, 1994; Ellis, 2002a) and to the probability of their form-meaning mappings (MacWhinney, 1987). Usage-based theories of language acquisition as well as constructionist theories of child language greatly contributed to underlying the importance of phraseology in the language acquisition process of a first or second/foreign language. Because of the close links these theories of language acquisition have with phraseology, they are the subject of the next chapter (Chapter 2).

Phraseology now holds a prominent place in S/FLA research. This is reflected in the large number of terms used to refer to structural patterns in descriptive SLA studies: holophrases (Corder, 1973), prefabricated routines and patterns (Hakuta, 1974), formulaic speech (Wong-Fillmore, 1976), memorized sentences and lexicalized stems (Pawley & Syder, 1983), lexical phrases (Nattinger, 1980), formulas (Ellis, 1994; McLaughlin, 1995), chunks (Ellis, 1996), constructions (Ellis, 2003, 2006a), etc. Phraseology is also at the heart of some proficiency tests such as the American Council on the Teaching of Foreign Languages (ACTFL), and learning English for Academic or Special Purposes involves learning the phraseology of the genre (Swales, 1990). Regarding teaching, some researchers (Nattinger, 1980; Nattinger & DeCarrico, 1992; Porto, 1998) argue that instruction should be based on Sinclair’s idiom principle and focus on patterns; this is the case of the Lexical Approach (Lewis, 1993, 2000, 2002). In addition, corpora are increasingly used in language teaching (Cobb, 2007; Wichmann & Fligelstone, 2014).

Despite the fact that phraseology now seems to be at the core of all these disciplines, many questions remain unanswered, particularly in relation to the difficulty encountered by advanced learners of a TL to sound natural and idiomatic. In the next section, we give an account of the vast terminology associated with the field of phraseology and provide the definition of phraseological language used in the present research.

1.3. Phraseological language: terminology, characteristics and definition

Phraseology is usually referred to as a subfield of lexicology involved with the study of word combinations rather than single words. A large number of terms have been (and are still) used to refer to the phenomenon of phraseological language (see Wray 2002, p. 9 for a comprehensive list). Examples are *chunks*, *formulae*, *formulaic sequences*, *collocations*, *multi-*

word units, ready-made utterances, etc². Multi-word units are usually defined and identified in relation to their features. They can for example be described “in relation to several characteristics that have been found to play a role in their acquisition, processing, and use, such as frequency, familiarity, predictability, fixedness, and pragmatic function, to name a few” (Siyanova-Chanturia & Pellicer-Sánchez, 2018, p. 3).

Frequency is very often considered as an important characteristic of phraseological units (Ellis, 2012). Corpus studies showed that phraseological units occur with high frequency in authentic language. Frequency also largely affects the way we learn and process language and it is usually adopted as a criterion for the selection of phraseological units in teaching. However, very infrequent multi-word units (e.g. idioms such as *raining cats and dogs*) are nonetheless highly conventional and familiar to the majority of native speakers. *Familiarity* or *conventionality* is therefore another important defining feature of phraseological units.

Another major feature is *predictability*. This refers to the fact that a language user is able to anticipate the upcoming word(s) and possibly skip them (as the word *fork* in *knife and fork*). As a consequence, the reading of a text will happen more quickly and the speaker’s perceived fluency will increase (Wood, 2002).

Phraseological units are also defined according to their degree of *fixedness*. Idioms and proverbs are usually viewed as fixed word combinations as no or few modifications can be made at the risk of losing their meaning, while other phraseological units allow for a great deal of flexibility (e.g. *it is X that*). Another closely related feature is *transparency* (Schmitt & Carter, 2004). Just as for fixedness, phraseological units can be placed on a continuum of transparency with opaque word combinations at one end (e.g. the idiom *kick the bucket*) and transparent ones at the other end (e.g. lexical bundles such as *I don’t know if*). These two features will clearly have an influence on their acquisition by TL learners.

Pragmatic function also characterizes phraseological units in the sense that they are very often used in speech to fulfil certain functions such as maintaining social interaction (e.g. *how interesting*), expressing solidarity (e.g. *I see what you mean*) or organizing discourse (e.g. *first of all, on the other hand*). Phraseological units actually play a crucial role in the achievement of pragmatic competence in a first or a second/foreign language.

² We mostly use the term *phraseological units* in this doctoral dissertation, interchangeably with *word combinations* and *multi-word units*, as we felt these were the most inclusive.

Ellis (2012) lists two criteria (besides frequency) that are characteristic of phraseological units from a corpus-linguistic approach: *association* and *native norms*. According to him, raw frequency is not enough; highly frequent sequences such as *and of the* or *but it is* lack clearly identifiable meaning and function, and do not appear phraseological. This is why the strength of association factor should also be taken into account. Phraseological units such as *raining cats and dogs* display a high degree of cohesion and occur together much more than would be expected. Measures of association are therefore more relevant for their identification. The other criterion is that the sequence should reflect native-like selection and fluency (Pawley & Syder, 1983). In fact, phraseological units like *have a nice day* are frequent and show a high level of coherence, but will be frowned upon if they are used in formal academic writing.

Prominence to one or more of these features will be given according to the approach. A pragmatic approach will place a high emphasis on the functional feature of multi-word units. A psycholinguistic approach will be more interested in the frequency and probability of phraseological units. A lexical bundle approach will heavily rely on frequency (Biber, Conrad & Reppen, 1998; Biber, Conrad & Cortes, 2004). Researchers working within a same approach make their own choice of different criteria to distinguish phraseological units from non-phraseological ones and do not always mention them explicitly. Gries (2008) regrets incidentally a general lack of rigor and argues for the use of explicit parameters when defining phraseology from a corpus-linguistic perspective in order to ensure the comparability of phraseological studies. He suggests the following parameters: the nature of the elements, the number of elements, the frequency of occurrence, the distance between the elements, the degree of lexical and syntactic flexibility of the elements, and the role of semantic unity and semantic non-compositionality/non-predictability.

Taking this into account, defining phraseological language is not an easy task. The definition of phraseological units adopted in the present doctoral research is the one proposed by Erman and Warren (2000, p. 31), since it leaves room for the inclusion of a large variety of word combinations: “combinations of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization”.

The next section describes two major approaches to phraseology, namely the ‘traditional’ or ‘phraseological’ approach (Section 1.4.1.) and the ‘distributional’ or ‘frequency-based’ approach (Section 1.4.2.). The benefits of combining the two different approaches for S/FLA are discussed in Section 1.4.3.

1.4. Two different approaches to phraseology

The lack of a single definition that can capture all the essential features of phraseological language at once makes its identification difficult and taxonomies of phraseological units abound in the literature. They serve lexicological or lexicographic (Gläser, 1986; Cowie, 1988; Moon, 1998), pedagogical (Nattinger & DeCarrico, 1992; Lewis, 1993) or psycholinguistic purposes (Wray & Perkins, 2000; Wray, 2002). The taxonomies differ according to the criteria used for the categorization of the word combinations and to the importance given to some criteria compared to others.

There exist two major approaches to the identification and classification of multi-word units: the ‘traditional’ or ‘phraseological’ approach (Nesselhauf, 2004) and the ‘distributional’ (Sinclair, 1987) or ‘frequency-based’ (Nesselhauf, 2004) approach.

1.4.1. The ‘traditional’ or ‘phraseological’ approach

The ‘traditional’ or ‘phraseological’ approach originates from Russian scholars (Vinogradov, 1947; Amosova, 1963; Cowie, 1998) and sees phraseology as a continuum along which word combinations are situated, with the most opaque and fixed ones at one end (pure idioms) and the most transparent and variable ones at the other (free combinations) (Cowie, 1981). Linguists working along this approach have contributed to the establishment of phraseology as a discipline, the creation of a terminology and a number of linguistic criteria that differentiate one type of phraseological unit from another, and more specifically, that distinguish between the most transparent and variable word combinations from free combinations. The traditional approach considers the most idiomatic units as the most core ones: “idioms form the majority and may be regarded as the prototype of the phraseological unit” (Gläser, 1998, p. 126). This focus on fixed and non-compositional phraseological units is also reflected in reference and teaching tools, which tend to deal more with idioms, proverbs or phrasal verbs than with more variable word combinations. It adopts a top-down approach, which means that it identifies phraseological units on the basis of linguistic criteria.

Most taxonomies emerged from the traditional approach of phraseology and are rooted in lexicology and lexicography (Granger & Paquot, 2008). Among the most influential, we find Cowie’s (1988, 1994) and Mel’čuk’s (1995, 1998) taxonomies, which are quite similar. They first distinguish between *composites* or *semantic phrasemes*, which function syntactically (i.e. they function as constituents of sentences such as objects, adjuncts, etc.), and *formulae* or *pragmatic phrasemes*, which function pragmatically. Then, *composites/semantic phrasemes* are

further subdivided into three categories that form a continuum going from the most transparent and variable word combinations to the most opaque and fixed ones (*restricted collocations* or *semi-phasemes*, *figurative idioms* or *quasi-phasemes*, and *pure idioms* or *full phasemes*). Cowie subdivides the formulae category into *routine formulae* and *speech formulae*. Another influential taxonomy is Burger's (1998), which distinguishes between three different functions of phraseological units in discourse: *referential units*, *structural units* and *communicative units*. The referential units include *nominative phraseological units* – which refer to objects, phenomena or facts of life and which are made of *collocations*, *partial idioms* and *idioms* – and *propositional units* – which refer to a statement or an utterance about these objects or phenomena. The structural units are used to establish grammatical relations and the communicative units serve an interactional purpose.

1.4.2. The 'distributional' or 'frequency-based' approach

The 'distributional' or 'frequency-based' approach has its origins in Sinclair's lexicographic work and rests on the 'idiom principle' (Sinclair, 1991), which views language as essentially made up of strings of co-selected words that constitute single choices. Phraseology is central to language and phraseological units – no matter which type – prevail over single words. It uses a bottom-up corpus-driven approach: it generates a wide range of word combinations which do not all belong to predefined linguistic categories. All the word combinations that were considered as peripheral or falling outside the domain of phraseology in the traditional approach have become essential in the frequency-based approach, because very frequent in the language.

No taxonomy has resulted from the distributional approach, but main categories of phraseological units can be created on the basis of the different extraction procedures (Stubbs, 2002). On the one hand, we have the *n-gram* or *cluster analysis*, which extracts continuous sequences of two or more words called n-grams, lexical bundles, clusters, etc. using a frequency threshold. On the other hand, we have the *co-occurrence analysis*, which extracts discontinuous combinations of two words called collocation, collocate or co-occurrence using statistical measures. The word combinations retrieved using these extraction procedures will vary depending on the settings and filters adopted, as well as the statistical measure used. A number of sub-types can be further distinguished within the co-occurrence analysis according to the focus of investigation and the type of information used for identifying collocations: *surface*, *textual* and *syntactic co-occurrence* (Evert, 2008). According to *surface co-occurrence*, two words are said to co-occur when they appear within a certain distance (collocational span),

measured by the number of intervening word tokens. *Textual co-occurrence* defines co-occurrence by the fact that two words appear in the same textual unit (sentences, clauses, paragraphs, documents, etc.). *Syntactic co-occurrence* investigates co-occurrence by looking at the (direct or indirect) syntactic relationships between the words (e.g. verb + object); it is particularly suited to long-distance dependencies between collocates. Furthermore, when considering the distance separating the co-occurring words, three different approaches can be identified: the *n-gram*, the *window*, and the *collocation networks approaches* (Gablasova, Brezina & McEnery, 2017). The *n-gram approach* identifies contiguous combinations (e.g. the bigrams *of the, minor changes* or *I think*) or contiguous combinations allowing internal variation (e.g. *minor but important/significant/observable changes*). The *window-approach* identifies looser word combinations, that is co-occurrences within a specified window span determining the number of occurring words to the left and to the right. For example, the co-occurrences of *changes* within a window span of five words to the left and five words to the right will include *minor* – as in the *n-gram approach* – but also *notify* or *place* as found in the following broader patterns: *notify somebody of any changes* and *changes which took place*. The *collocation networks approach* combines “multiple associations identified using the window approach to bring together interconnections between words that exist in language and discourse” (Gablasova et al., 2017, p. 158). The distance then results from an association with another word in the network, and not from an immediate window distance.

1.4.3. Reconciling the two approaches

Both approaches can contribute to the integration of phraseology in S/FLA studies and pedagogical applications (Granger, 2005). On the one hand, the ‘traditional’ approach “provides SLA researchers with a keener awareness of the different categories of MWUs [multiword units]” (2005, p. 2). It also yields useful insights for pedagogical research since the existence of different categories and subcategories of phraseological units implies different kinds of learning (Grant & Bauer, 2004). On the other hand, the ‘frequency-based’ approach “covers the whole range of co-occurrence patterns with no a priori exclusions” (Granger, 2005, p. 3). As a result, they have highlighted the importance of some categories of multi-word units that were neglected because of their apparent predictability (e.g. free combinations and lexical bundles). Recent learner language studies have actually shown that these types of word combinations were a problem for TL learners (e.g. De Cock, 2003; Nesselhauf, 2005; Paquot, 2005).

It seems therefore relevant to combine the two approaches. The ‘frequency-based’ approach offers a good starting point for the investigation of a wide range of word combinations, which should then be broken down into fine-grained linguistic categories resulting from the ‘traditional’ approach and that can be integrated in pedagogical applications. In an attempt to design a taxonomy of phraseological units that reconciles the two approaches, Granger and Paquot (2008) applied the insights from the distributional approach to Burger’s (1998) traditional linguistic classification. This gave rise to a taxonomy consisting of three broad categories of phraseological units – referential, textual and communicative phrasemes – further subdivided into subcategories. This is the classification that was used in the present thesis (see Chapter 4, Section 4.6.).

After having discussed different issues related to the definition, identification and classification of phraseology, we now turn our attention to the methodological approaches used for the analysis of phraseological units.

1.5. Analyzing phraseology: methodological approaches

According to Laufer and Waldman (2011), three main approaches have been adopted for the analysis of learners’ phraseological language³: the analysis of large learner corpora using corpus analysis techniques referred to as Learner Corpus Research (LCR) (Section 1.5.1.), the use of experimental techniques (Section 1.5.2.), and the analysis of errors (Section 1.5.3.). The section devoted to LCR is further subdivided into three subsections, which respectively take stock of the learner corpus research into phraseology (Section 1.5.1.1.), and describe the frequently used corpus-based Contrastive Interlanguage Analysis (CIA) (Section 1.5.1.2.) and longitudinal analysis (Section 1.5.1.3.).

1.5.1. Learner Corpus Research (LCR)

The study of S/FLA has always resorted to written and spoken data produced by learners. The difference with learner-corpus data is that it provides authentic data samples involving a large number of participants. Learner corpora can be defined as “electronic collections of natural or near-natural data produced by foreign or second language (L2) learners and assembled according to explicit design criteria” (Granger, Gilquin & Meunier, 2015). Corpus data have a number of advantages. First, they provide data from a great number of learners, which makes it more representative. Second, their electronic format speeds up the analyses and broadens the

³ Most of the analyses involve the learners’ knowledge of collocations.

scope of investigations. Third, they are more reliable than native speakers' intuitions since they make comprehensive and balanced data available. Fourth, once collected, the data remain readily accessible and language development can be analyzed. The bulk of studies that resulted from these corpora is grouped under the term *Learner Corpus Research* (LCR).

According to Callies (2015, p. 42), learner corpora are particularly useful if one wants to investigate high-frequency phenomena, but are more "limited or even unsuitable for the study of infrequent, highly L2-specific or optional phenomena". Meunier and Littré (2013, p. 69) also reported the following drawbacks of working with corpora: a) corpus data allow for description, but not for explanation since instances of a linguistic feature in the learners' productions do not necessarily mean that the learner knows the feature; b) limitation related to the corpus itself, in the case of avoidance or underrepresentation of linguistic features, and c) no record of contextual features, multi-modality and retrospective user comments.

1.5.1.1. Learner corpora and phraseology

As already explained, corpora and corpus tools have played a significant role in the increased focus on phraseological research, and phraseology has established itself as a major feature of LCR, as corpus analysis is particularly suited for the investigation of phraseological units (Ebeling & Hasselgård, 2015).

According to Paquot and Granger (2012), most corpus studies of learner phraseological language have involved foreign language learners of English, usually at upper-intermediate and advanced proficiency levels. Their focus has mostly been on learner writing (of argumentative essays, especially), even though learner speech has received increased attention.

As for the types of phraseological units investigated in LCR, Ebeling and Hasselgård (2015) mentioned lexical bundles, collocations, phrasal verbs, and collostructions. Each of these types of word combinations relate to different kinds of identification procedures that vary in their level of sophistication. Lexical bundles or n-grams (i.e. uninterrupted sequences of words of varying length, usually between two and five words) are automatically extracted from learner productions using tools such as WordSmith Tools (Scott, 2012) or AntConc (Anthony, 2014) and ranked according to their frequency. The focus can be on the length, function or accuracy of the bundles. A certain amount of dispersion of the lexical bundles across the individual texts is required and a threshold of recurrence (i.e. number of times a bundle has to occur in the corpus to be considered for investigation) has to be set. Research of this type falls into the

lexical bundle branch of the ‘frequency-based’ approach as described in Section 1.4.2. Reference studies include Stubbs (2001, 2002).

Collocations are now often referred to as “usage-based lexically restricted combinations” (Granger & Paquot, 2008, p. 42) made up of a pre-selected item (‘node’) that imposes restrictions on its co-occurring words (‘collocates’) (e.g. *heavy rain*). The collocates of a node are either retrieved and listed according to frequency of occurrence or, more often, according to some statistical association measure while comparing learner and native speaker data. The collocates of a node can also be highlighted using keyword-in-context (KWIC) concordance (e.g. Laufer & Waldman, 2011; Paquot & Granger, 2012). There exist a large number of association measures and each measure has its own advantages and disadvantages (Evert, 2005; Gries, 2008, 2012)⁴. They also highlight different aspects of the phraseological relationship between two words (Gablasova et al., 2017). The most frequently used measures in LCR are the Mutual Information (MI) score and the t-score. The MI score is used to assess the degree to which the words in a phrase occur together more frequently than would be expected by chance (Oakes, 1998; Manning & Schutze, 1999). The higher the score, the more exclusive the association between the words and the rarer the combination. As reported by Gablasova et al. (2017), the MI score rather highlights rare exclusivity than strength, tightness or coherence. It is particularly useful for the identification of sequences that are constituted by low-frequency items (e.g. *longitude and latitude*) and it tends to favor names, terms, and specialized or technical combinations, that is combinations that mostly occur in specific texts or genres. MI2 is a version of the MI score that is not negatively linked to frequency, but it has not gained in popularity yet. The t-score is an adjusted value of collocation frequency that highlights frequent word combinations and that is closely related to the raw frequency of the combination in the corpus. However, “while all collocations identified by the t-score are frequent, not all frequent word combinations have a high t-score” (Gablasova et al., 2017). The main issue with this measure is that it is not standardized (i.e. it is directly dependent on the corpus size) and therefore not comparable across corpora of different sizes (Hunston, 2002). Log Dice is not a frequently used measure in language learning research, but it has the advantage of being standardized and having a fixed maximum value. It emphasizes exclusivity (as the MI-score) but is not biased by low-frequency, and is therefore the closest to the MI2 score.

⁴ The association measures described here seem to be the most frequently used measures in recent corpus-based studies of learner writing. The list is therefore not exhaustive.

Some LCR studies (e.g. Groom, 2009) combined the two identification procedures discussed above (i.e. the lexical bundle analysis and the statistical measures of co-occurrence) in order to obtain more robust results. The reason behind this is that the lexical bundle analysis may lack reliability as it only considers “invariable and continuous recurring sequences” (Ebeling & Hasselgård, 2015, p. 213). The collocation analysis will complement this, “as it is open to variation in the sequences” (2015, p. 213) and the statistical measures will increase the certainty and strength of the collocational relationship between the items.

Another type of phraseological unit that has received a great deal of attention in LCR is phrasal verbs, which function like single verbs composed of a verb and an adverbial particle (e.g. *pick up*). They resemble free combinations of verb + preposition/adverb and their components are not always contiguous. For these reasons, their extraction is either done manually (e.g. Hagglund, 2001) or necessitates a great deal of manual post-editing of the output that has been generated by searches based on Part-Of-Speech (POS) tags (e.g. Chen, 2013) or a pre-defined set of verbs and/or particles (e.g. Waibel, 2008).

As for collocations, they are “significant associations between words and constructions” (Ebeling & Hasselgård, 2015, p. 214) and refer to the combination of a ‘collexeme’ (word or ‘lexeme’ that is attracted to a particular construction) and a ‘collostruct’ (construction associated with a particular lexeme). Concretely, the analysis “starts with a particular construction and investigates which lexemes are strongly attracted or repelled by a particular slot in the construction (i.e. occur more frequently or less frequently than expected)” (Stefanowitsch & Gries, 2003). For example, the construction *N waiting to happen* would be strongly associated with the word *accident*. Work has been carried on the development of measures for collocation analysis (Evert & Krenn, 2001; Gries & Stefanowitsch, 2004).

The present doctoral dissertation falls into LCR as it mainly resorts to learner corpora for the analysis of phraseological units. As for their identification, it combines the frequency-based approach by extracting a large range of word combinations according to their frequency of occurrence, with a phraseological approach by also taking into account different linguistic categories that are pedagogically relevant. See Chapter 4 for a description of the corpora and the identification procedure.

1.5.1.2. Contrastive Interlanguage Analysis (CIA)

One of the most frequently used methods in learner corpus analysis is Contrastive Interlanguage Analysis (CIA) (Granger, 1996). The method either compares learner data with native speaker

data, or different types of learner data coming from different L1 backgrounds. The first type of comparison (learner data vs L1 data) would help uncover patterns of under- and overuse in the learners' interlanguage (quantitative differences), as well as patterns of misuse (qualitative differences). The other type of comparison (different learner data) would help establish whether the observed differences between the learner and the native data are specific to a certain group of learner or context of use (e.g. features that are L1-related) or if they can be generalized to a wider learner population or to more situations (e.g. developmental features) (Callies, 2015, p. 40). Examples of studies comparing the phraseological productions of learners with those of native speakers are Kaszubski (2000), Granger (1998b) and Altenberg and Granger (2001). Kaszubski (2000) conducted a contrastive, corpus-based approach of combinations including the verbs *be*, *have*, *make*, *take*, *do*, *get* in different TL English corpora from various L1 speakers (Polish, Spanish, and French) and in a corpus of English native speakers. Granger (1998b) used a corpus of French learners of English and a corpus of English native speakers in order to investigate and compare their use of sentence builders and collocations made of intensifying adverbs (e.g. *completely*, *totally*, *highly*). Altenberg and Granger (2001) compared learner data of French and Swedish speakers learning English as a TL with native speaker data to uncover their respective use of the high-frequency verb *make*.

No CIA is adopted in the present thesis as the analyses only involve a comparison between two types of learners that differ according to the context in which they learn their TL (CLIL versus NON-CLIL contexts). The learners come from the same L1 background and no comparisons with native speaker material are made.

1.5.1.3. Longitudinal analyses

Studies comparing the language output of learners and native speakers or of various groups of learners at a single point in their development are referred to as cross-sectional studies. However, S/FLA also took interest in investigating learners' developmental patterns. Longitudinal studies, as defined by Johnson and Johnson (1999, p. 102), "examine the language behavior of one or more subjects as that behavior develops over time". The more data collection points are made with the same individuals, the more refined the interpretation of the development. Ployhart and Vandenberg (2010) argued for a minimum of three collection points so that it is "possible to fit a developmental line and visualize potential effects on that line (linear progression or regression, U- or reversed U-shaped behavior)" (Meunier, 2015, p. 380). The use of corpora lends itself well to this type of analysis, but research into this area evidently requires corpora representing various proficiency levels. Longitudinal corpora are ideal in this

respect, but they are not very numerous. In fact, collecting such corpora costs time and money, and the chances of losing participants throughout the process are big. This is why most studies have made use of ‘quasi-longitudinal’ (Granger, 2002; Thewissen, 2013) or ‘pseudo-longitudinal’ corpora (Johnson & Johnson, 1999; Huat, 2012). These combine comparable cross-sectional corpora representing different proficiency levels or include learner variables that “allow differentiation along a proficiency-related feature” (Ebeling & Hasselgard, 2015, p. 216) (e.g. age, Common European Framework of Reference for Languages (CEFR) scale or length of stay in the target language environment). The learners are therefore not the same, but they usually share a number of characteristics such as L1 background or learning context. With longitudinal corpora, group progress, individual variation within each group and individual trajectories can be analyzed, while the analysis of individual learning trajectories is not possible with pseudo-longitudinal corpora (Meunier, 2015).

A significant number of LCR developmental studies have focused on phraseology. Some of these used a pseudo-longitudinal research design, such as Bartning and Forsberg (2006) who analyzed the stages of development in ‘prefabricated sequences’ produced by Swedish learners of French. The proxy used to measure the time variable was proficiency level correlated with numbers of years of TL instruction. Others made use of longitudinal corpora, as in Crossley and Salsbury’s study (2011) where the development of lexical bundle accuracy in the spoken productions of six English second language learners is investigated in comparison with native speaker data. Bestgen and Granger (2014), for example, performed both a longitudinal and a pseudo-longitudinal analysis of the role played by phraseological competence in the development of L2 writing proficiency and text quality assessment, and the two research designs actually revealed different results.

Longitudinal analyses of learner corpora have also been carried out to examine the role played by phraseology in S/FLA. These revealed that particular multi-word combinations that are highly frequent, functional and prototypical may drive the acquisition process forward. Studies carried out by Myles and colleagues (Myles, Hooper & Mitchell, 1998; Myles, Mitchell & Hooper, 1999; Myles, 2004) looked at the development of chunks within individual learners of French in England. The results showed a clear correlation between the use of chunks and linguistic development, with chunks providing learners with “a databank of complex structures beyond their current grammar, which they keep working on until they can make their current generative grammar compatible with them” (Myles, 2004, p. 153).

The present thesis includes both cross-sectional studies and a longitudinal analysis. The longitudinal corpus was gathered in order to examine the development of the learners' phraseological language over two consecutive years in terms of range and accuracy (see Chapter 5).

1.5.2. Experimental analyses

Besides LCR, (quasi-)experimental analyses have also been extensively used in SLA research and most of them involve elicitation. Elicitation tasks intend to make learners produce or recognize specific linguistic items that have been selected in advance by researchers for investigation (Laufer & Waldman, 2011). The learners are usually purposefully not made aware of the actual research purpose. Elicitation is typically used when researchers want to test a specific hypothesis about a linguistic phenomenon that cannot be properly investigated in learner corpora because it is too infrequent or not produced at all. More and more researchers have also started to recognize the benefits of combining corpus data and experimental data (e.g. Gilquin, 2007; Meunier & Littré, 2013). In fact, elicitation tasks can be used to supplement corpus data in order to get a more comprehensive picture of the linguistic phenomenon under investigation and increase the validity and reliability of the results and their interpretation. This methodology is often referred to as “triangulated research designs” (Callies, 2015, p. 43). Generally speaking, experimental methods have the following advantages (Meunier & Littré, 2013, p. 69): as mentioned above, they give access to less frequent linguistic features and provide supplementary evidence to corpus findings, but they also tap “into learners' explicit knowledge of the language” and “understanding of the various constraints of tense and aspect usage” (in the case of grammaticality judgment tests and metalinguistic justifications).

Elicitation techniques that are used to tap into learners' productive knowledge of phraseological units usually involve translation tasks and fill-in exercises. This is the case of Biskup's study (1992) in which Polish and German learners of English are asked to translate L1 collocations into English. Hasselgren (1994) also investigated the use of English collocations by advanced Norwegian learners using a translation exercise of texts. Bahns and Eldaw (1993) made use of both a translation task and a cloze task with German advanced learners of English as a foreign language (EFL). The cloze task consisted in filling in sentences with a verb in order to produce the target verb-noun collocations.

Elicitation tasks used to investigate the learners' receptive knowledge of word combinations commonly ask learners to distinguish between correct and incorrect linguistic items. Bonk

(2000) administered three tests of collocational knowledge to ESL (English as a second language) learners of different proficiency levels. A number of figurative uses of a verb were presented in sentences and learners had to recognize the incorrect uses. Gyllstad (2005) designed two receptive collocational tests for advanced Swedish learners of English: COLLEX (Collocating Lexis), where learners were presented with two lexical collocations – one correct collocation and one pseudo-collocation – and had to choose the correct one, and COLLMATCH (Collocate Matching), where learners had to decide which verbs combined with which nouns when presented with a grid of three verbs and six nouns. Barfield (2007) made use of a self-assessment technique asking learners to rate their knowledge of single words (verbs and nouns) and of different word combinations using these verbs and nouns. For the single words, the learners could choose between “I have never seen X”, “I have seen but do not know X”, “I know X but I have to think about it” and “I know X without thinking”. For the word combinations, they had the choice between “I don’t know this combination at all”, “I think this is not a frequent combination”, “I think this is a frequent combination” and “This is definitely a frequent combination”.

According to Howarth (1998), we could consider these elicitation procedures as direct measures of phraseological competence, since the results are represented by the number of participants who know a given word combination. However, this method makes it difficult to establish “the validity of any predefined list of target collocations, since, as many EFL teachers might agree, this component of a learner's linguistic competence is one of the least predictable, making it hard to generalize from subjects' knowledge or ignorance of a small number of individual items” (1998, p. 32).

Psycholinguistic experimental studies took an interest in researching the processing and storing of phraseological units with the aim of finding out whether multi-word units are processed and retrieved as wholes. With this aim in mind, a number of researchers resorted to the eye-tracking methodology (e.g. Underwood, Schmitt & Galpin, 2004; Siyanova-Chanturia, Conklin & van Heuven, 2011), which makes it possible to analyze whether words in phraseological units are processed faster (and therefore with less effort) than single words on the basis of the eye movements. Another frequently used technique is the Self-Paced Reading (e.g. Schmitt & Underwood, 2004; Conklin & Schmitt, 2008; Tremblay, Derwing, Libben & Westbury, 2011). It makes each word in a text appear one-by-one on a computer screen, with the participants pressing up a button to bring up each new word and the computer timing each push to measure the time required to recognize and process the word. The dictation methodology has also been

used to address the question of holistic storage (e.g. Schmitt, Grandage & Adolphs, 2004). The participants hear stretches of dictation that – if they are long enough to overload working memory – force them to reconstruct the stretches via a number of language resources (one of these resources being the inventory of phraseological units stored in memory). It is assumed that if the phraseological units are stored and available for use, they are likely to be produced in the participants' responses.

Other experimental analyses aimed to research the prosodic characteristics of figurative phraseological language (mostly idioms and proverbs). Van Lancker and Canter (1981), Van Lancker, Canter and Terbeek (1981) as well as Van Lancker-Sidtis (2003) conducted experiments with pre-recorded idiomatic and literal utterances that were judged by listeners as sounding either literal or idiomatic. The researchers' aim was to identify the specific cues that helped listeners distinguish between literal and figurative utterances. Work carried out by Hallin and Van Lancker-Sidtis (2017) looked at the prosody of low-frequency proverbs using an elicitation task that allowed for the analysis of speech rate and tonal patterns. Participants were asked to read a target sentence silently (proverbs or matched novel sequences) and then say it out loud in a most natural way.

The corpus-based data used in the present thesis were supplemented by experimental data in order to tap into the learners' receptive collocation knowledge. A number of verb-noun collocations were pre-selected to be used in an elicitation task where learners had to correctly recognize the verb (see Chapters 4 and 6). It may seem problematic to label this type of data 'experimental' as it is a controlled elicitation task and not an experiment *stricto sensu*. However, as we have seen in this section, the term 'experimental' is commonly used in SLA to refer to elicitation. Besides, in his description of the different kinds of data used in SLA research, Ellis (1994, p. 670) distinguishes "experimental elicitation" as a sub-category of "elicited language use", and describes it as a way of inducing learners to produce or comprehend data that are specific to the researcher's interests.

1.5.3. Error analyses

Studies examining errors in the field of phraseology usually start from the word combinations extracted from the learner corpora in the case of corpus-based studies, or from the word combinations elicited by the task in experimental methods. Some researchers such as Thewissen (2008) made use of an error-tagged learner corpus, which means that each text had been manually annotated for errors. Computer-aided error analyses of phraseology are however not

very numerous. Once the phraseological units of interest have been extracted, they are usually checked in reference dictionaries or corpora, judged by native speakers, or both.

When looking for the potential sources of errors in the production of collocations, a major distinction is usually made between interlingual errors and intralingual errors (Richards, 1974). The former denotes errors that are caused by the interference of the learners' L1 and that reflect the learners' inability to separate the two languages. The latter reflects the learners' competence at a particular stage and illustrates some of the characteristics of language acquisition and rule learning. Intralingual errors originate from the structure of the TL itself and may also occur in the language of children acquiring their L1. They may indicate final competence or a transitional stage.

Interlingual errors, i.e. errors that are caused by interferences from the learners' L1, appear to be one of the most frequent causes of collocation errors. After having extracted about 2,000 verb-noun collocations from a corpus of German learners of English, Nesselhauf (2005) noted that half of the errors produced by the learners displayed an influence from the mother tongue. Sometimes, learners translate word-for-word from their L1; this is called L1 transliteration or literal translation (Tarone, 1981). They may also switch to their L1 and use the L1 words; this is referred to as language switch (Tarone, 1981). The notion of L1 transfer and its negative consequences for the acquisition of phraseological units is further discussed in Section 1.7.2. below.

Intralingual errors include different types of strategies learners resort to in order to make up for their collocational deficiency. *Synonymy* or *analogy*, for instance, refers to the fact that learners confound one of the components of a word combination with another word that is synonymous or semantically related. Consequently, they erroneously combine the component with the collocate of the confused word (e.g. the production of the erroneous collocation **reach findings* instead of *reach a conclusion* would stem from the confusion between the semantically related words *finding* and *conclusion*) (Howarth, 1998; Kuo, 2009). *Approximation* or *paraphrase* refers to the fact that “learners use incorrect words or structures but which share enough semantic features to satisfy the needs of expressing the desired meaning” (Kuo, 2009, p. 149). For example, **fell* is used to get the sense of *failed* and **talk* to express the idea of *tell*. According to Hong, Rahim, Hua and Salehuddin (2011), the features shared by the correct and incorrect words can also be phonological or morphological. Learners would resort to this strategy when the TL structure has not been firmly formed or when they are not certain about the TL structure in order to ease their linguistic burden. *Overgeneralization* refers to the

creation of deviant structures on the basis of learners' previous experience of the TL structures. Learners will for instance create a deviant form of the past tense of irregular verbs by using the morpheme *-ed* (e.g. **telled the story* instead of *told*). Nesselhauf (2005) also noticed phraseological errors that result from an *overextension* of the nonrestrictive nature of the meaning of one component of the word combination (for example the verb *make* in the erroneous collocation **make an experience*).

Most studies focused on the collocational possibilities of the lexical items composing the collocations. However, in her study of verb-noun collocations produced by German learners of English, Nesselhauf (2003) also paid attention to errors involving prepositions (e.g. *raise the question *about* instead of *of*) and determiners (e.g. *get *the permission* instead of *get permission*). These error types are related to restricted collocations due to their variability that is not always predictable, but restricted in a "non-generalizable manner" (Erman & Warren, 2000, p. 52). These types of errors need to be highlighted since they imply that besides knowing which lexical items collocate with each other, learners have to know the whole combination in order to be fully accurate.

Distinguishing between correct and incorrect instances is not always straightforward, especially when it comes to analyzing lexis (Ferris, 1999). According to James (2013), phraseological units may be erroneous in two different ways; they may be formally wrong or formally correct but contextually inappropriate. James (2013) refers to the former as *grammaticality errors* and to the latter as *acceptability errors*. Formally wrong phraseological units may resemble existing units (e.g. **keep close watch* instead of *keep a close watch*) or bear no resemblance at all to any existing phraseological unit (e.g. **make a ballad* instead of *go for a walk*). In her analysis of phraseological errors made by different groups of EFL learners, Thewissen (2008) found a higher proportion of grammatical errors in the Spanish-speaking group of learners, who also had a lower level of language proficiency, while a higher proportion of acceptability errors was found in the French- and German-speaking groups of learners, who had a higher proficiency level.

Besides looking at learners' phraseological competence in terms of frequency and variety of production, the present doctoral research also looks at the accuracy of their phraseological language. The error analyses carried out in this thesis are based on the phraseological units extracted from the learner corpora (Chapter 5) or the answers given to the experimental task measuring the learners' collocation knowledge (Chapter 6).

Section 1.5. has described the methodological approaches that are commonly adopted when analyzing phraseological units. The next section stresses the importance of the communication mode as corpus-based studies pointed to differences between written and spoken phraseology.

1.6. Written versus spoken phraseology

Studies comparing written and spoken corpora seem to suggest that phraseological units are even more pervasive in spoken than in written language. Biber, Johansson, Leech, Conrad & Finegan (1999, p. 996) looked at lexical expressions in conversation and academic prose and found that “overall, conversation makes use of a larger stock of lexical bundles than academic prose, and lexical bundles make up a somewhat greater proportion of the total discourse in conversation (especially if two-word contracted sequences are included)”. Altenberg (1998) also pointed to the importance of phraseological language in spoken discourse. He examined the phraseology of spoken English in the London-Lund Corpus of Spoken English and found that “over 80 per cent of the words in the corpus form part of a recurrent word-combination in one way or another” (1998, p. 102). However, it has to be said that this proportion was based on combinations consisting of repetitions or fragments that are of little interest (for example *in a*, *and the*, *the the*, etc.). Similarly, Erman and Warren (2000, p. 37) reported a “somewhat greater density of prefabs in spoken than in written language”. According to Ellis, Simpson-Vlach and Maynard (2008, p. 376), “speech is constructed in real time and this imposes greater working memory demands compared with writing, hence the greater need to rely on formulas”.

Spoken language is not only more phraseological than written language, it also appears to comprise different types of phraseological units. As reported by Mukherjee (2009), phraseology is one of the areas where speech and writing clearly differ. One of the characteristics of spoken language is its interactional nature, which is reflected by word combinations such as *thank you*, *you know* or *come in* (Shin, 2007). After their analysis of a corpus of spontaneous speech, McCarthy and Carter (2002) listed different types of word combinations that seemed to be important in spoken interactions: 1) combinations with discourse marking functions (*you know*, *I mean*, *do you know what I mean*, *if you see what I mean*); 2) combinations with politeness functions (*do you think*, *do you want me to*, *I don't know if/whether*); 3) combinations with hedging functions (*I think*, *sort of* before nouns, verbs, adverbs, adjectives, etc., *a bit*, *I don't know*), and 4) high-frequency combinations expressing vagueness and approximation (*a couple of*, *and something like that*, *that sort of thing*). In their comparison of recurrent phrases in the speech of learner and native speakers of English, Kizil and Kilimci (2014) showed that over the half of the phrases found in the native corpus were stance expressions, i.e. combinations that

express the speaker's attitudes (e.g. *I think, I would like to*). Referential expressions were the second most frequent type of combinations – they include markers of vagueness (*sort of*), of time (*all the time*), quantifiers (*there's a lot of*), etc. These were followed by discourse organizers (e.g. *I don't know* to close a topic) and special conversational expressions (for example *yeah that's right yeah*).

These differences between the two communication modes have implications for the teaching of phraseological language, especially as a number of studies have reported a lack of awareness among language learners, even at high proficiency levels (see for example Gilquin & Paquot, 2008). An exploratory study into the learners' spoken phraseological language is currently being carried out (see concluding chapter of this thesis for a brief description).

In the next section, we discuss a number of factors that may explain the difficulties learners encounter when acquiring the phraseological lexicon of a TL. Section 1.7.1. focuses on the relation between explicit and implicit knowledge and their interface, Section 1.7.2. on L1 transfer and Section 1.7.3. on the influence of psycholinguistic effects.

1.7. Acquiring phraseological language

While phraseology is central in child language acquisition and is an integral part of the child language creation process (see Chapter 2), it causes problems for learners, even with a high level of proficiency in a TL, which suggests that their phraseological language is created differently. Contrary to natives, TL learners in school contexts would start by learning individual words and gradually build up bigger chunks, so that it is particularly hard for them to establish strong associations between pairs of words forming collocations (Wray, 2002; Schmitt & Underwood, 2004). Also, Wible (2008) wrote that since the primary source of input for learners is text (and not speech, as is the case for L1 learners), they already know about the separate units identified by the white spaces and need to discover that some sequences of these separate units are best considered as bundled wholes. Other factors that may hinder the acquisition of phraseological units by TL learners have been considered by Ellis (2008b) and are discussed in the sections below.

1.7.1. Implicit and explicit learning

According to Ellis (2008b, p. 7), many beginner-level word combinations are explicitly learned “as wholes, fast mapped as declarative memories like other aspects of vocabulary”, while the graded statistical knowledge of most of the phraseological competence requires a more implicit form of learning through usage (Ellis, 2002a). Repeated encounters with the word combinations

in context facilitate their retrieval and production (Bybee & Hopper, 2001) and the power of input frequency in language learning has been demonstrated, notably by Ellis (2002a, 2002b, 2003). A number of empirical studies looking at the effect of frequency on the incidental learning of collocations showed the importance of repetition. Durrant and Schmitt (2010), for instance, reported that learners who had been repeatedly exposed to the target collocations showed better levels of recall. Similarly, Webb, Newton, and Chang (2013) made learners read and listen to different versions of a graded reader that varied according to the number of encounters of the target collocations and found a positive relation between the number of encounters and the learning of the collocations. They concluded that five or more encounters seemed to be required to learn the form of the collocations incidentally.

Research has however pointed out that implicit learning has its limits, especially in instructed foreign language settings where input in the TL is often limited (to the classroom) and where explicit instruction has a role to play (see Muñoz, 2006 and 2008, and Meunier, 2012). Previous research has showed the efficiency of awareness-raising activities for the learning of collocations. Among the activities that seem to improve the learners' collocational knowledge, we find the use of corpus tools. Chan and Liou (2005) showed that EFL learners who performed online activities using a bilingual concordancer improved their knowledge of verb-noun collocations. Dictionary use also appears to contribute in a positive way, as suggested by Laufer's study (2010); the fact that L2 learners resorted to the dictionary in order to fill in sentences in which the verb in verb-noun collocations was missing, significantly increased their knowledge of collocations. Tasks that promote the noticing of sound repetition in phraseological units such as alliterations (repetition of initial consonants as in **d**ay **d**ream, **g**ood as **g**old, **b**lind as a **b**at) and assonance (repetition of identical or similar vowel sounds in neighboring words, e.g. my kind of **guy**) have also proved effective (Lindstromberg & Boers, 2008a, 2008b). In a comparison of focus-on-meaning and focus-on-forms tasks, Laufer and Girsai (2008) showed that the focus-on-form approach (i.e. reading a text and completing text-based vocabulary and translation tasks) led to the acquisition of collocations. Webb and Kagimoto (2009) found that both productive (i.e. cloze task) and receptive tasks (i.e. reading and memorizing) were beneficial to the learners' collocational knowledge, with the productive ones being more favorable to higher proficiency learners and the receptive ones to lower proficiency learners.

There is actually little agreement on which phraseological units should be taught explicitly. Some researchers like Boers, Deconinck, and Lindstromberg (2010) think that high frequency

word combinations will be incidentally picked up by learners. In contrast, Nesselhauf (2005) argued that explicit attention should be paid to these frequent combinations as they do not appear to be mastered at advanced proficiency levels.

1.7.2. Transfer from the mother tongue

Learning a TL involves a reshaping of the way the learner sees the world and construes experiences (Slobin, 1996; Robinson & Ellis, 2008) so that it corresponds to the one of the native speakers of the new language, and transfer from the learners' L1 affects TL phraseology at various levels. The acquisition of phraseological units by non-native speakers is impeded by the fact that they already have a repertory of multi-word units in their L1 that may overlap with the TL to some degree, or not overlap at all (Conklin & Carrol, 2018). For example, the English idiom *to throw money out the window* largely overlaps in meaning and form with the French *jeter l'argent par les fenêtres* (literally *to throw money through the windows*). However, English and French have different expressions (entirely different forms) to say that someone is feeling sad or depressed (same meaning): *to feel blue* and *avoir le cafard* (literally *to have the cockroach*). Another example is the English *to cost an arm and a leg* and the French *coûter les yeux de la tête* (literally *to cost the eyes of our head*), which share the same meaning, but have subtly different forms. When the two languages overlap, knowing the word combination in one language may facilitate its acquisition in the other language. In spite of this, Kellerman (1979) reported that learners tend to be cautious about resorting to their L1. This would result from the fact that they are aware of the differences between the languages and the language-specific character of the idiomatic language.

A number of recent studies (Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011, 2013; Titone, Columbus, Whitford, Mercier & Libben, 2015; Pritchett, Vaid & Tosun, 2016) investigated how cross-language overlap influences the processing of phraseological units. Most of them involved collocations and idioms. What emerged from these studies is that when phraseological language overlaps in form and meaning across languages (congruence), its recognition and processing occurs more accurately and more quickly. In contrast, when there is overlap in form but not in meaning, or vice versa (incongruence), processing tends to be slower or disturbed (Conklin & Carrol, 2018). This may be partly explained by the fact that L1 translation equivalents are quickly and automatically activated while processing in the TL (Zhang, Van Heuven & Conklin, 2011). In addition, the frequency of occurrence of a (congruent or incongruent) collocation in the TL, i.e. the number of times learners will encounter the collocation in the TL input, also appeared to influence its processing.

Kellerman (1978; 2000) investigated ‘transferability’ and his findings suggested that the further a multi-word combination is situated from the central core of phraseology, that is semantically opaque, and syntactically and collocationally inflexible, the more potentially transferable it may be. However, Nesselhauf (2003) looked at erroneous collocations including the verbs *take* and *make* and concluded that the degree of idiomaticity did not seem to be the reason for transfer, but rather restricted collocations (e.g. *make dreams* instead of *have dreams*). Paquot (2008) went a step further and looked at both correct and incorrect multi-word units that function as exemplification in the TL academic writing of French-speaking learners of English. She reported two types of L1 transfer. First, multi-word units that are similar units performing similar functions in the learners’ L1 are more likely to be overused. This would be reinforced by the so-called ‘teaching effect’, as these types of multi-word units are pervasive in teaching materials. In a similar vein, Wang and Shaw (2008) suggested that when the L1 is similar to the TL, learners tend to be more creative and to take more risks, which eventually leads to more errors. Conversely, when the L1 does not share similarities with the TL, learners are less inclined to create collocations. The second type of transfer according to Paquot (2008) involves less salient multi-word units that are essentially semantically and syntactically compositional (e.g. *let’s take the example of*, directly translated from *prenons l’exemple de*). The multi-word unit exists in English, hence the transfer of form can be seen as a positive transfer, but this does not systematically result in a positive transfer of function, register and frequency.

Selinker and Lakshmanan (1992) stated that intralingual errors (i.e. errors caused by L1 influence, see previous Section 1.5.3.) are more likely to fossilize, i.e. to persist regardless of the amount of input, instruction or error correction (Selinker, 1972), and according to Laufer and Waldman (2011), some of the phraseological errors made by advanced TL learners may suggest fossilization. Studies comparing the use of multi-word units by native and non-native writers of English pointed to instances of over- and underuse that can possibly be assigned to transfer from the learners’ L1 (Gass & Selinker, 1983; Odlin, 1989; Jarvis, 2000; Neff van Aerselslaer, 2008; Rica-Peromingo, 2009). Similar findings were found for EFL learners having French as their L1. Granger (1998b) investigated their use of intensifying adverbs and found that learners typically (over)used amplifiers that have a direct equivalent in French (e.g. *completely*, *totally*). Paquot (2008), as mentioned above, also found instances of overuse in learners’ academic writing, due to the similarity of the TL multi-word unit with a L1 equivalent performing similar functions (e.g. *for example*, *for instance*). De Cock (2003) observed that learners underused multi-word units that have no cognate forms in French (e.g. *sort of*), that

they misused some English sequences that have French partially deceptive cognates (e.g. *on the contrary*, *in fact*) and that they used atypical combinations that are literal translations of French multi-word units (e.g. *according to me*).

1.7.3. Psycho-linguistic aspects

The acquisition of a new language involves “the combined effects of form-meaning correspondences, construction types and tokens and their distributional properties, and prototypicality effects” (Ellis, 2008b, p. 8; see also Chapter 2, Section 2.3.2. in this dissertation). Some phraseological units will play a beneficial role in the acquisition of the TL: highly frequent and prototypically functional phrases like *put it on the table*, *how are you* or *it’s lunch time* are quickly learnable and are “the likely candidates as construction seeds” (Ellis, 2012, p. 29). By contrast, phraseological units such as *once in a blue moon* are much less frequent and have a nontransparent and idiomatic meaning, which makes them more difficult to acquire for learners. A large amount of language experience is needed to encounter these sequences and then commit them to memory (Ellis, 2008a, Ellis, Simpson-Vlach & Maynard, 2008). They are not seeds of learning but rather targets for learning and probably account for the ceiling effects observed in advanced levels of TL acquisition.

Howarth (1996) stated that psycholinguistic effects would be associated with the fact that TL learners have problems with the acquisition of “the large and complex middle ground of restricted collocations” (p. 42). Since collocations are semantically transparent and involve subtle kinds of restrictions, they do not attract the learners’ attention when encountered in the input for the first time. In fact, according to him, learners do not have many problems with the interpretation of the meaning of these collocations; they are therefore not perceived as salient in the TL input and are less easily memorized by the learners.

When discussing the acquisition of phraseological language in instructed contexts, it is important to examine the attention phraseology receives in the classroom. We tackle this aspect in the next section.

1.8. Teaching phraseological language

According to Kennedy (2008), language pedagogy has tended to neglect phraseology over the last few decades. This may be due to different reasons (2008, pp. 37-40). First, it is difficult to identify the multi-word units in a language, even though corpus-based analyses made this identification process easier. Second, when the corpus-based analyses of phraseology emerged and favored form-focused pedagogical approaches, language teaching theory and practice were

encouraging communicative teaching approaches focusing on meaning. Third, a tried and true methodology for teaching phraseology is lacking. Fourth, since phraseology is believed to be mostly learned implicitly through unconsciously meeting the multi-word units repeatedly in context (this is especially true for the infrequent combinations), it requires a good deal of effort from the teachers to maximize the opportunities for implicit learning. Finally, more research evidence is needed in order to support the inclusion of phraseology in the explicit curriculum.

Coxhead (2008) agrees to say that theoretical principles for the teaching of phraseological units are still needed and suggests following those used for the teaching of single words (see Nation, 2001, p. 45). These principles are ‘noticing’ (stimulate learners to notice and fill gaps in their learning), ‘retrieval’ (the learners strengthen the memory of the words by retrieving their form or meaning from memory), and ‘generation’ (the words that learners have already encountered are subsequently met or used in a different context). She applied these three principles on collocations in her English for Academic Purposes (EAP) classroom but states that “the question of how effective these techniques are still remains” (2008, p. 156). Jones and Haywood (2004) used various teaching methods to focus on phraseological units in reading activities (highlighting, use of concordances, analysis of word combinations in context, gap-fill exercises, etc.), and although they found an increase of the learners’ awareness of word combinations, the learners did not seem to have learned the combinations or to use them in their writing. In order to shed light on what prevents learners from using phraseological units, Coxhead (in preparation) interviewed a small sample of six university students on their vocabulary use in a writing task based on input texts. The participants reported the following reasons for not using the word combinations they had noticed in the texts: a lack of knowledge of the words or a lack of time, the risk of making mistakes in using them, an inaccurate recall, the inability to find them in a dictionary, or the fact that learning single words is already hard enough.

In a review of the application of recent findings on phraseological language in instructed SLA, Meunier (2012) confirmed that, at this point, vocabulary teaching still mainly focuses on single words. She states (2012, p. 122) that “our increasingly refined understanding and knowledge of the formulaic nature of language has clearly impacted L2 teaching”, but that “despite considerable efforts made by some publishers to include more authenticity (and hence formulaicity) in teaching materials, more work will be needed on that front to make such efforts more visible”, especially in language textbooks for general purposes. In relation to spoken phraseology, a number of studies showed that frequent features of spoken interactions are actually missing in EFL textbooks (e.g. Carter, 1998; Burns, 2001; Gilmore, 2004; Walsh,

2010). A concrete example of the poor treatment of phraseology in English for General Purposes textbooks was given by Gouverneur (2008). She compared the intermediate and advanced student's books and workbooks of three different series of textbooks and looked at their treatment of the phraseological patterns of two high-frequency verbs. The findings showed unjustified discrepancies across levels – intermediate textbooks paid explicit attention to verb-noun collocations, but this was less the case in the advanced textbooks, while these word combinations have been shown to remain problematic at that proficiency level. Moreover, a lack of consistency in the selection of the phraseological patterns was observed (large number of different collocations according to the textbook). Attention paid to phraseology in class thus very much depends on the teacher.

As far as reference tools are concerned, Pecman (2008) states that the number of monolingual phraseological dictionaries is increasing but that bilingual phraseological dictionaries are hard to find, except for those including proverbs and sayings. Siepmann (2008) wrote that the majority of monolingual learners' dictionaries do not focus enough on traditional non-compositional idioms and collocations and tend to neglect semantically fully transparent routine formulae.

The CLIL approach (see Chapter 3), which constitutes the background of the present doctoral research, would create more opportunities for the implicit learning of phraseological units to occur since it provides a larger amount of TL input through the teaching of content subjects. Since explicit forms of teaching phraseological units in the classroom still appear to be limited and/or inconsistent, it is interesting to investigate whether a 'high input' instructed environment such as CLIL helps foster the acquisition of phraseological language, and if so to what extent.

In the next section, we report the main findings on the phraseological competence of TL learners in general and CLIL learners in particular.

1.9. Target language phraseological competence: previous findings

Phraseological competence has now become central to linguistic competence, in both first and S/FLA (Siepmann, 2008). Phraseological language is believed to help learners achieve native-like fluency (Granger & Meunier, 2008; Handl, 2008) and acquire receptive and productive skills (Conklin & Schmitt, 2012; Ellis, 2012; Wray, 2012). However, research has shown that learners are struggling to achieve phraseological competence, even at advanced levels (e.g. Meunier & Granger, 2008; Siepmann, 2008; Meunier, 2012). This is reflected in the following sections, which give an account of the findings on TL phraseological competence in terms of

productive (Section 1.9.1.) and receptive knowledge (Section 1.9.2.), as well as on its development over time (Section 1.9.3.). Results from studies looking at the phraseological competence of CLIL learners are not the object of separate sections since they are not very numerous.

1.9.1. Productive knowledge of phraseological language

This section is further subdivided into two subsections that each focus on one communication mode: Section 1.9.1.1. deals with the TL learners' phraseological competence in writing and Section 1.9.1.2. with their phraseological competence in speaking.

1.9.1.1. Written phraseological language

A summary of the main findings from corpus-based analyses of the TL learners' phraseological competence was made by Siepmann (2008) on the basis of various studies (e.g. Howarth, 1996; Bahns, 1997; Granger, 1998a; Cobb, 2003; De Cock, 2003; Nesselhauf, 2005; Siepmann, 2005b; Paquot, 2008). His summary states (2008, p. 194) that learners a) often use a more limited set of multi-word units; b) have lexical "teddy bears" (p. 194) (i.e. words – or combinations of words – they tend to use more frequently than native speakers would do, because these are "widely usable and above all safe"⁵ (Hasselgren, 1994, p. 250)) for particular pragmatic functions; c) overuse one-word markers such as *however* over multi-word ones fulfilling similar pragmatic functions such as *that said*, and that d) in terms of accuracy, and especially for collocations, some of the learners' errors can be traced back to overgeneralization in the TL or to L1 influence.

Another recurring finding regarding the TL learners' phraseological competence is that high-frequency verbs such as *have*, *do*, *make*, *take* or *give*, which have long been seen as 'easy' verbs, actually seem to be a major problem. In fact, several studies pointed to the learners' overuse of these verbs (Hasselgren, 1994; Källkvist, 1999; Kaszubski, 2000; Altenberg & Granger, 2001), which would be related to the 'lexical teddy bear' phenomenon described above (Hasselgren, 1994). Other studies rather revealed a tendency towards underuse (Altenberg & Granger, 2001) or misuse (Nesselhauf, 2004), especially when the verbs are used as delexical verbs, i.e. combined with a noun to simply indicate an action and therefore carry very little meaning. In fact, while the core meanings of the verbs are usually acquired, their delexical uses – that mostly appear in multi-word units – remain particularly difficult. For example, Altenberg and Granger

⁵ An example of lexical teddy bear would be the highly frequent adverbial booster *very* instead of the less frequent adverbial boosters *exceedingly* or *gravely* (Hasselgren, 1994).

(2001) found that both the French-speaking and Swedish learners of English in their study either underused the delexical uses of the verb *make*, mostly when occurring with nouns expressing speech actions (e.g. *make a remark*, *make an argument*, *make a statement*, etc.) or misused them (e.g. **make a balance* [strike], **make harm* [do], **make a step* [take]).

As for the productive phraseological competence of CLIL learners in comparison to their NON-CLIL peers, the former group usually appears to perform better than the latter. In an investigation of the writing skills of Dutch learners of English in CLIL vs NON-CLIL programs in The Netherlands, Verspoor and Edelenbos (2011) found that CLIL learners used relatively more grammatical chunks, lexical chunks, and collocations, but made more errors. They assigned the higher number of chunks and collocations to a more authentic input provided in the CLIL setting, and the higher number of errors to a more risk-taking behavior and a more substantial output from the CLIL learners. Berendse (2014) analyzed whether Dutch CLIL learners of English had a higher command of prepositions than their NON-CLIL peers and found a more correct use in the CLIL group, even though they roughly used the same prepositions as the NON-CLIL learners. Klampfl (2010) compared the writing proficiency of Austrian CLIL and mainstream learners of English in terms of vocabulary and looked at what she calls “collocational vocabulary errors of prepositions” (p. 84). These collocational errors were only found in the texts written by NON-CLIL learners.

Phraseological knowledge being one integral component of general lexical knowledge, findings related to (single word) vocabulary use of CLIL learners are worth mentioning too. According to Dalton-Puffer (2008), vocabulary is one of the language competences that is positively affected by the CLIL approach. This has been confirmed in a study carried out by Mewald (2007) where CLIL students demonstrated a wider use of lexis in their oral productions – they used more word types (nouns, verbs, adjectives, and adverbs) than the NON-CLIL students (nouns and verbs) – and took more risks regarding mistakes. Lasagabaster (2008) also found better scores in a CLIL group of Spanish learners of English regarding written and spoken vocabulary. According to Agustín Llach (2010, p. 9), CLIL students learn “more words and more nuances (associations, collocations, syntax, etc.) of words than learners in instructional approaches”. Agustín Llach (2009) and Celaya (2008) stated that lexical transfers and direct borrowings from the L1 occurred much more frequently in the productions of NON-CLIL learners. However, CLIL learners appeared to produce more lexical inventions (Celaya, 2008; Agustín Llach, 2009), with an increasing use as proficiency develops (Celaya & Torras, 2001; Naves, Miralpeix, & Celaya, 2005).

1.9.1.2. Spoken phraseological language

Research on the learners' spoken interlanguage is fairly new (Kizil & Kilimci, 2014). Nevertheless, a number of studies have already provided interesting insights into the learners' spoken phraseological language. De Cock (2004) analyzed the preferred sequences of words used by native and non-native speakers of English and more specifically focused on markers of vagueness. Her main finding was that learners significantly underused vagueness tags compared to native speakers: "learners are lacking in routinized ways of interacting and building rapport with their interlocutors and of toning down and weaving the right amount of imprecision and vagueness (a typical feature of native speaker informal interactions)" (2004, p. 243). In fact, learners mostly used sequences that actually belong to more formal registers (*and so on* and *et cetera*) or misused informal vagueness markers (e.g. *or something like that*). Also, while native speakers very frequently used *kind of* and *sort of* in spontaneous speech, these sequences were very infrequent in the learners' productions; when they were used, it was mostly in front of nouns (instead of verbs in native speaker speech) or in their non-pragmatic sense (meaning *a type of*). Besides these cases of underuse and misuse, learners also appeared to overuse *of course*. Shirato and Stapleton (2007) looked at the vocabulary used in a Japanese EFL spoken corpus and a native speaker corpus. They found differences for each type of word combinations that are specific to spoken interactions, as categorized by McCarthy and Carter (2002). As far as discourse markers are concerned, *you know* and *I mean* were the most frequent in the native corpus. *(Do) you know what I mean* and *if you see what I mean* were also common among native speakers but never used by the learners. The category 'vagueness and approximation' was the one where the largest differences were found: sequences such as *a couple of*, *and something like that*, *that sort of thing*, *all this/that sort of thing*, etc. were omnipresent in the native corpus but never occurred in the learners' speech. The use of indirect forms in the category 'face and politeness' (for example *I don't know if/whether*, *what do you think*, etc.) was inherent in the native corpus, but much less frequent in the learner one. *Sort of* was the most frequently used word combination among native speakers for the category 'hedging'; similarly to what De Cock (2004) found, it was much less frequently used by the learners, mostly in the sense of '*what sort of X*' and before nouns only.

1.9.2. Receptive knowledge of phraseological language

Few studies have been carried out investigating learners' receptive knowledge of phraseological units. A number of them looked at the relation between receptive knowledge of TL collocations and general proficiency in the TL. Bonk's study (2000), which included a L2 receptive

recognition test of verb-noun lexical collocations, revealed a correlation between general English proficiency and collocation proficiency. Similarly, Keshavarz and Salimi (2007) reported a significant relationship between EFL learners' collocational knowledge of lexical and grammatical collocations and their performance on cloze tests measuring overall language proficiency. In contrast, Barfield (2003) found no correlation between general proficiency and recognition of decontextualized verb-noun collocations. Barfield's study also showed that knowledge of individual verbs and nouns did not necessarily involve recognition of their combination in a verb-noun combination. Nguyen and Webb (2017) compared receptive knowledge of verb-noun and adjective-noun collocations at different frequency levels and examined which factors best predicted receptive knowledge of collocations. Their study showed that the EFL learners knew less than the half of collocations at each frequency level and that their collocational knowledge decreased at each level. Positive correlations were found between knowledge of collocations and single-word items, node word frequency being the strongest predictor. In a comparison between the productive and receptive knowledge of adjective-noun collocations by the same Spanish university students, Jaén (2007) revealed a poor collocational competence from learners, with significantly lower scores obtained at the productive test in comparison to the receptive test. Granger (1998b) conducted a study in which English native speakers and French-speaking EFL learners were instructed to choose adjectives (e.g. *significant, reliable, difficult*, etc.) as acceptable collocates for a number of amplifiers (e.g. *highly, seriously, readily*, etc.), and to mark particularly strong collocates with an asterisk. The learners marked considerably fewer combinations than the native speakers. Gyllstad (2007) analyzed the receptive knowledge of adjective-noun and verb-noun collocations of advanced Swedish learners of English and found a close relationship between the learners' vocabulary size and receptive collocation knowledge. The findings also suggest a progression in receptive collocation knowledge concomitant of learning level, overall language proficiency, and vocabulary size.

What we know from research on receptive vocabulary knowledge in CLIL is that the results usually tend to be in favor of CLIL learners (Jiménez Catalán, Ruiz de Zarobe & Cenoz, 2006; Jiménez Catalán & Ruiz de Zarobe, 2009; Sylvén & Ohlander, 2015). Research investigating the receptive phraseological knowledge of CLIL learners is, to our knowledge, inexistent. One exception, however, is a very recent study carried out by Hendrikx (2019) in the framework of the project encompassing the present thesis. Her study looked at CLIL and NON-CLIL French-speaking learners' receptive knowledge of intensifying compounds (e.g. *brand new, ice cold*,

etc.) in the target languages English and Dutch. The results showed a greater receptive knowledge of compounds in the CLIL group (for English and Dutch). In addition, the compounds that were more frequent in the reference corpus were found to be more easily recognized by the learners. In addition, a number of studies investigated the receptive phraseological knowledge of TL learners in foreign and second language learning conditions and systematically found better results in the higher input conditions. Shehata (2008) looked at both the receptive and productive knowledge of adjective-noun and verb-noun collocations by EFL and ESL Arabic advanced learners and found that the learners' mother tongue and learning environment had a strong influence on the acquisition of TL collocations. Alsakran (2011) conducted a similar study and also found better results in the ESL group at the receptive test. Yamashita and Jiang (2010) compared Japanese EFL and ESL learners' scores at an acceptability-judgment task using both congruent (direct translation exists in the L1) and non-congruent (no direct equivalent in the L1) collocations; EFL learners were found to have more trouble dealing with incongruent collocations than ESL learners. Bueraheng and Laohawiriyanon's study (2014) involved two groups of Thai learners of English with different degrees of instructional exposure to the TL and examined their receptive and productive performance. The study showed that the group that benefited from more exposure in the TL scored better.

1.9.3. Development of phraseological competence

The development of learner's phraseological language has already been explored in a number of studies. Some findings seem to suggest that TL phraseological competence increases with proficiency in the TL. Crossley and Salsbury (2011) analyzed the frequency of lexical bundles produced by six learners of English over a year (using a longitudinal corpus) and found an increase in bigram accuracy and phraseological range. De Haan and van der Haagen (2013) studied the use of intensifiers in the adverb-adverb and adverb-adjective combinations by Dutch advanced learners of English over a period of seventeen months and noted a gradual increase towards a more native-like use of intensifying constructions. Siyanova-Chanturia (2015) analyzed the longitudinal production of noun-adjective collocations by over thirty Chinese beginner learners of Italian at the beginning, middle and end of an intensive course, and found a greater number of higher frequency items and strongly associated collocations at the end of the course. Li and Schmitt (2009) followed a Chinese ESL learner over the course of one year; the study revealed an increase in the learning of lexical phrases and an improvement in the degree of appropriate usage, but also an overreliance on a limited number of lexical phrases.

Thewissen (2008) looked at the phraseological accuracy of French-, German- and Spanish-speaking learners of English as a TL across two proficiency levels (B1 to C2) and found that the number of phraseological errors decreased as proficiency increased. In an attempt to assess the role of phraseological competence in the development of ESL writing proficiency, Bestgen and Granger (2014) explored the developmental use of bigrams between two data points and found a decrease in the use of collocations made up of high-frequency words that are less typical of native writers.

However, a number of studies found different results. Li and Schmitt (2010) found no major change in the production of adjective-noun collocations over the course of a year by advanced Chinese learners. Chen (2013) focused on the use of phrasal verbs by Chinese English learners over three years and identified no considerable increase at the end of the third year, even a decrease in the second year. Chen and Hin-tat (2015)'s pseudo-longitudinal study looked at the use of multi-word units involving prepositions at four different proficiency levels and revealed a growth in complex prepositions, while prepositional verbs and other types of multi-word units showed the reverse pattern. In her investigation of EFL accuracy development, Thewissen (2013) found an increase in phraseological accuracy between B1 and B2. However, this increase was followed by a stabilization, possibly explained by an "accuracy-complexity trade-off effect" (2013, p. 87); in an attempt to produce a wider range of phraseological units (and therefore complexify their language), higher-level learners take more risks, which usually results in mistakes. Laufer and Waldman (2011) investigated the written production of verb-noun collocations by Hebrew learners of English at three different proficiency levels and found that while the number of collocations significantly grew at the more advanced level, so did the number of errors.

All the studies discussed above differ in terms of methodology (longitudinal vs pseudo-longitudinal, use of association measures or not, etc.) and focus (collocations, phrasal verbs, intensifiers, etc.). It seems therefore difficult to establish a clear link between phraseological competence and language proficiency. Methodological issues have already been raised in Paquot, Naets, and Gries' longitudinal study (forthcoming), which showed that considering proficiency level and topic/prompt as possible predictors of phraseological complexity in learner writing provided an overall better degree of variance explanation in EFL learners' use of multi-word units. In addition, what particularly stands out is the fact that whenever an increase in the use of phraseological units is found, the accuracy suffers from it. This shows

that phraseological competence cannot be considered as a monolithic concept and that it should therefore be looked at from different perspectives.

One of the few relevant studies to our knowledge that compared the development of CLIL and NON-CLIL learners' phraseological language is the one conducted by Hendrikx (2019). She looked at the acquisition of intensifying constructions (e.g. *really nice*, *ice-cold*, *hypersensitive*, etc.) in Dutch and English by the same French-speaking CLIL and NON-CLIL learners as the ones used in the present thesis and her findings suggested that CLIL education was positively related to the development of a more target-like use of these specific constructions. Verspoor and Smiskova's study (2012) is also worth mentioning since they investigated the development of chunks over two years in a "low input group" and a "high input group" (p. 28) of Dutch high school EFL learners. The high input group, which consisted of English immersion students, showed a higher production of chunks at the end but few significant differences were found.

Interest in the writing development of CLIL learners and their NON-CLIL counterparts seems to be greater. Ruiz de Zarobe (2010), for example, carried out a longitudinal analysis of the written production outcomes (content, organization, vocabulary, language usage and formal aspects such as punctuation) of two CLIL groups (one group receiving one curricular subject through English as a foreign language, the other receiving two curricular subjects) versus one NON-CLIL group after two years. She found a positive relationship between grade, linguistic outcomes and the educational program. Juan-Garau, Salazar-Noguera and Gené-Gil (2013) analyzed CLIL and NON-CLIL EFL learners' longitudinal writing development at three data points by looking at the same criteria as Ruiz de Zarobe's (2010) study. They found greater gains in CLIL, both longitudinally and compared to the NON-CLIL group for several measures, including vocabulary. Gené-Gil, Juan-Garau and Salazar-Noguera (2015) investigated complexity, accuracy and fluency (CAF) measures development in the writing of CLIL and NON-CLIL learners over a period of three years and the CLIL's texts showed greater improvement in all three measures. CLIL and NON-CLIL EFL learners' receptive and productive skills were analyzed longitudinally by Roquet i Pugès (2011), who found progress in the CLIL learners' writing (for accuracy and lexico-grammatical skills).

1.10. Conclusion

This chapter has shown that researchers now recognize the importance of taking a wider view of phraseology; the traditional approach that only considered the most fixed and idiomatic word combinations as the core has now made way for a more encompassing, frequency-based

approach that also values a whole range of largely compositional multi-word units. Pleas for a more complementary approach that would combine the benefits of both views have been made and seem well-founded. In fact, the traditional view of phraseology offers interesting linguistic categories while the distributional view provides interesting new techniques and informs about new types of phraseological units.

Methodologically speaking, learner corpora are one possible entry point to investigate phraseological language at a single point in time or its development over time, as reflected by the large number of studies carried out in LCR. The methods used to extract the multiword units are varied and the types and number of word combinations extracted differ a lot according to the method used. Since corpus data have their limits, more experimental methods that tap into a more explicit type of knowledge and that allow the analysis of less frequent phraseological units are also used. Finally, the accuracy of the learners' phraseological language has also been investigated by classifying the different error types and looking for their potential sources.

Even though linguistic research has demonstrated that both written and spoken language is essentially made of patterns, proficient TL learners are still struggling to master native-like phraseology. A great deal of phraseological competence is acquired implicitly through exposure to language input and learners appear to be sensitive to frequency of exposure and repetitions in the language input, but the input provided to TL learners in instructed settings does not seem to suffice. In fact, we cannot expect learners of a TL in instructed settings to acquire the language unconsciously as L1 and second language learners do. The acquisition of a TL in an input-poor environment should be compensated by more explicit approaches to learning (e.g. awareness raising activities, a greater focus on form(s), etc.). Some explicit forms of instruction have actually shown to be beneficial for the learning of phraseological units, but even though progress has been made accordingly, there is still work to be done. According to Granger and Meunier (2008), learners and teachers should be made aware of the pervasiveness of phraseology and of its importance in language learning. The phraseological information should be more easily available. In addition, more empirical evidence is needed in order to know which types of phraseological units are relevant for teaching and what the best ways to teach them are. We have also seen that L1 transfer comes under consideration and affects the acquisition process of phraseological units in different ways.

Previous empirical studies show that vocabulary is one of the linguistic aspects that benefit from the CLIL approach. Studies looking more specifically at phraseology seem to indicate better results in CLIL learners when compared to their NON-CLIL peers. However, reviewing

the empirical studies on the phraseological competence of CLIL learners led us to draw a number of observations. First, the studies that actually investigated phraseological performance in CLIL are scarce. Second, the very few CLIL studies that went beyond the single word level involved participants from a limited number of schools (one or two). Third, they always focused on English as a TL. Fourth, their interest in phraseology is either secondary (they primarily focus on vocabulary in general or writing proficiency) or very often limited to one specific type of phraseological units (prepositions, intensifying compounds, etc.). Only the doctoral research carried out by Hendrikx (2019) in the framework of the project encompassing this thesis overcomes the first three limitations.

The close links between phraseology, input and language acquisition have been touched upon in the present chapter and are further demonstrated in the following chapter (Chapter 2) which describes language acquisition from a usage-based and emergentist approach.

CHAPTER 2: Second/foreign language acquisition from a usage-based approach

2.1. Introduction

This chapter tackles S/FLA from a usage-based perspective. Usage-based theories of language acquisition have become more and more influential in the investigation of child language acquisition and have contributed in some way to the revival of phraseology (see Chapter 1). Usage-based approaches are actually closely related to the field of phraseology since they hold that language is learned through usage/communication and that it initially develops from ‘phrases’ or ‘formulas’ (Ellis, 2008b). Section 2.2. in this chapter describes the major principles of the usage-based and emergentist approaches to first language acquisition as they adopt the same perspective, namely that input and language use are at the origin of language acquisition. Their application to S/FLA is discussed in Section 2.3., with a focus on a number of important factors such as input and frequency effects, psycholinguistic effects (salience, prototypicality, redundancy and recency) and instruction effects (focus on form(s) and focus on meaning). Finally, some learner-related variables such as age, motivation and cognition are discussed in Section 2.4., before concluding the chapter in Section 2.5.

2.2. First language acquisition

Chomsky’s Universal Grammar (1986), which was one of the key theories of language, argued that the input learners are exposed to, is insufficient to acquire the complexity of the language system. This argument is referred to as ‘the poverty of the stimulus’ (Chomsky, 1980). As a result, it had to be assumed that humans have an innate grammatical configuration that cannot be learned based on linguistic experience alone and that would be comprised of several abstract principles that are the ‘core’ of the language and that regulate the language acquisition process. As for the lexicon, irregular constructions and pragmatics – which represent the periphery of the language – these would be acquired from the input over a number of years.

However, advances in the domains of developmental psychology, computational linguistics and cognitive science changed the view on this issue and new approaches to language acquisition emerged. The so-called ‘usage-based’ and ‘emergentist’ theories state that language is a system that is not predetermined, but that can be acquired from language use itself on the basis of social skills and general cognitive abilities such as generalization (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Tomasello, 2003).

The following two subsections present the theoretical assumptions that underlie the usage-based approaches (Section 2.2.1.) and the emergentist approaches (Section 2.2.2.).

2.2.1. Theoretical assumptions of usage-based approaches

The term ‘usage-based’ comes from the belief that actual language use is a primary shaper of linguistic form and that making meaning is central to how language is configured. The central idea is that the emergence of language results from exposure to numerous *usage events* (Kemmer & Barlow, 2000), which are “situated instances of the language user understanding or producing language to convey particular meaning in a specific communicative situation” (Tyler, 2010, p. 271).

According to usage-based theory, no level of grammar is independent of language use: “all abstract grammatical rules were at some point induced from concrete and particular usage events” (Behrens, 2009, p. 385). As a result, the formalist approaches to grammar that distinguish between words and rules that operate words (e.g. Pinker, 1999) are rejected. Language knowledge is rather seen as a structured inventory of form-meaning patterns (Langacker, 1987) “in which abstract grammatical patterns and the lexical instantiations of those patterns are jointly included, and which may consist of many different levels of schematic abstraction” (Tummers, Heylen & Geeraerts, 2005, pp. 228-229).

A number of powerful cognitive mechanisms are involved in the generalization of schemas and regular aspects of a language. *Entrenchment* is one of these. It refers to “the fact that repeated encounter of a unit leaves memory traces that stabilize the more often this unit recurs” (Behrens, 2009, p. 386). It applies to both words and larger units or constructions, if they can be retrieved without attention to detail. In order to generalize and form categories, the mind also needs to compare previously stored units with new ones to filter out dissimilarities and register similarities. The new units are then categorized according to similarities with stored units. This is how *schemas* are formed, through abstraction and generalization. A schema is defined by Langacker (1987, p. 492) as “a semantic, phonological, or symbolic structure that, relative to another representation of the same entity, is characterized with lesser specificity and detail”. These processes of schema formation and categorization are responsible for “the extraction of very abstract rule-like phenomena that are not tied to specific lexical material” (Behrens, 2009, p. 386). Type frequency leads to schema formation or abstraction, while token frequency contributes to entrenchment and storage as ‘chunk’ (see Section 2.3.1. below for a description of these frequency effects).

It is believed that schemas can be integrated to form larger units by means of *composition*. Compositionality refers to the ability of children to create new structures based on previous

experience (O’Grady, 2005). However, compositionality is only partial since the new composite structure has new emerging qualities. *Emergentism* highlights the fact that it is possible for new and more complex structures to emerge from simpler, basic facts (MacWhinney, 1999). This concept of emergence is at the core of ‘emergentist’ theories of language learning and is further developed in the next section (Section 2.2.).

Usage-based models provide a “domain-general perspective on language, language learning and language change” (Behrens, 2009, p. 389). The same learning mechanisms account for the acquisition of all of language. All linguistic units are symbolic, i.e. form-meaning pairings, which implies that the units’ formal aspects contribute to their meaning.

A final major tenet of usage-based models is that patterns of usage (collocational and frequency information) are central to the acquisition of the system. Phraseological units are actually “one of the hallmarks of child second language development” (McLaughlin, 1995, p. 8). Longitudinal studies looking at the emergence of creative linguistic competence in the language of children (e.g. Dabrowska 2000; Tomasello 2003; Dabrowska and Lieven 2005) demonstrated that “the usual developmental sequence is from formula to low-scope slot-and-frame pattern, to creative construction” (Ellis, 2008b, p. 5). The idea that L1 acquirers first learn and use phraseological sequences and then analyze their subparts is shared by Bolinger (1977), Weinert (1995), Bybee (1998) and Schmitt and Carter (2004), among others. Nattinger and Decarrico (1992) described this ‘whole to parts’ acquisition process in the following way:

One common pattern in language acquisition is that learners pass through a stage in which they use a large number of unanalyzed chunks of language in certain predictable social contexts. [...] Many early researchers thought these prefabricated chunks were distinct and somewhat peripheral to the main body of language, but more recent research puts this formulaic speech at the very center of language acquisition and sees it as basic to the creative rule-forming processes which follow. For example, first language learners begin with a few basic, unvarying phrases, which they later, on analogy with similar phrases, learn to analyze as smaller, increasingly variable patterns. They then learn to break apart these smaller patterns into individual words and, in so doing, find their own way to the regular rules of syntax. (1992, p. 15)

The concepts of *limited scope formulae* (Braine, 1976) and *slot and frame patterns* (Pine & Lieven, 1997) indicate that parts of children’s utterances are likely to represent unanalyzed

chunks that act like a frame into which a limited number of words and phrases can be inserted. Children will for example move from “rote-learned chunks” such as *I dunno* to the general schematic construction *NP aux neg Verb* (Eskildsen, 2008, p. 336).

From a usage-based perspective, first language acquisition can be summarized as a process of pattern recognition and generalization based on the accumulation of language experiences provided by a substantial amount of input and using elaborate cognitive learning mechanisms.

2.2.2. Theoretical assumptions of emergentist approaches

Like usage-based approaches, ‘emergentist’ theories also fall into the theories that see language as a system that emerges through input and interactions and that does not assume innate linguistic endowment. Their major contribution was to show that learners could go beyond experience to create new structures, thanks to previous experience and generalization mechanisms. In fact, emergentist theories are based on the principle that “a combination of smaller quantitative differences can produce a new quality, for example, human language as opposed to the communication systems found in other species” (Behrens, 2009, p. 387).

Evidence for generalization over missing evidence actually comes from connectionist models. According to Elman (2003, p. 400), “indirect positive evidence” may be the key to acquisition of complex linguistic rules. Morris, Cottrell & Elman (2000) also showed that a supposedly unlearnable construction (e.g. the question formation from sentences with embedded clauses like *The boy who is smoking is crazy*) can be learned on the basis of simpler utterances containing the relevant morphosyntactic information and the learners’ ability to generalize across domains, and this even though no positive evidence is provided.

The Complex Dynamic Systems Theory (CDST) of language development (Larsen-Freeman, 1997) has recently positively contributed to emergentist theories. CDST hinges upon the concept of dynamic development as it emerges over time. Each development of the language system relies on “the previous state of that system in dynamic interaction with the state of a larger number of relevant systems” (Lowie, 2017). Due to the highly interactive nature of language development, this process cannot be predetermined. The first main characteristic of CDST is that it sees language learning as a complex process of interacting subsystems that changes over time. The relevant subsystems for language learning – phonology, syntax and lexicon – are closely interdependent and influence each other in different ways according to the moment (Lowie, 2017). For example, at early stages of language acquisition, syntax is very much dependent on the lexicon. At later stages of acquisition, however, the lexical development

may depend on the acquisition of syntactic structures. Second, from a CDST perspective, the language learning process is highly individual. Each step in time in the developing language system depends on the previous state of the system, which varies according to the individual. The language learning process may take place slowly or rapidly according to the learning context including distraction and communicative needs (Lowie, 2017). Another characteristic of CDST is that this complex interaction of subsystems leads to a high degree of variability in language acquisition and use (Lowie, 2017). Variability is actually the “motor of change” (Lowie & Verspoor, 2015, p. 74), which means that it is a necessary feature for language development to occur. It results from the learners’ adaptation to the environment and their tendency to try out new structures. According to Spoelman and Verspoor (2010), an increase in variability is associated with developmental change. In contrast, no variability means no change.

In sum, from an emergentist perspective combined with the CDST of language development, language learning is seen as “a highly individual process in which numerous subsystems continuously interact in shaping development” (Lowie, 2017, p. 4). The present thesis considers individual variation to some extent; the analyses are performed on individual scores and measures of variation are reported, but they focus on groups of learners and do not scrutinize the development of a small number of individuals.

Section 2.2. has discussed first language acquisition from a usage-based and emergentist perspective. The question whether the insights from these approaches can be applied to S/FLA is addressed in the next section.

2.3. Second/foreign language acquisition

Usage-based researchers believe that a child acquiring his/her L1 and a learner acquiring a TL undergo the same process. MacWhinney (2008) states that the same general learning mechanisms such as frequency, analogy, competition between forms, or tracking recurrent sequences are at work in L1 and TL acquisition: “both groups need to figure out the patterns that govern word combination in syntactic constructions” (2008, p. 341). According to Ellis (2002b), the established L1 acquisition sequence going from formula, to low-scope pattern, and eventually to construction, can reasonably be applied to S/FLA. Some studies investigated the validity of the usage-based linguistics trajectory of learning a second or a foreign language and found positive evidence for the usage-based path of acquisition. Eskildsen and Cadierno (2007) looked at the development of *do-negation* by a Mexican learner of English as a second language

(ESL) and found that the learner initially relied on one specification of the pattern (*I don't know*), which was gradually expanded to be used with other pronouns and verbs. Mellow (2006) investigated the acquisition of embedded clauses by a Spanish ESL learner and also found that the learner initially learned the constructions with particular lexical items and gradually expanded the constructions to apply new lexical items. Zyzik (2006) examined the use of alternating verbs in English-speaking learners of Spanish as a foreign language (e.g. transitive *romper* and intransitive *romperse*) and found a consistent pattern of overgeneralization of *se* to transitive contexts and variable performance on verbs belonging to the same class. However, in more recent work, Ellis & Cadierno (2009) and Ellis (2013) claimed that TL acquisition is more complex than L1 acquisition because of the competition between L1 and TL constructions (see also Section 1.7.2. in the previous chapter).

As for emergentist theories, their contribution may help shed some light on the development of S/FLA, as discussed by Lowie (2017). First, the dynamic interaction of subsystems provides explanation for the fact that the knowledge of different languages in multilingual speakers are interconnected in a complex way, as evidenced by code switching and code mixing. Second, the individual nature of the language learning process makes it difficult to generalize for groups of learners and has implications for language teaching. Third, the fact that variability is a necessary condition for language development to take place, implies that there will always be change in the subsystems, no matter the proficiency level. In fact, there is no final state of learning and the construct of 'native speaker' should be treated with care. Also, an insufficient amount of variability will lead to fossilization, which may require a strong perturbation (any form of input, use, or instruction) to reorganize one or more subsystems. This notion of variability has implications for the assessment of second/foreign language learners; measuring only one aspect of language use at one given moment in time may not be representative of learners' language development.

The next sections present some important factors that are at play in S/FLA, namely input and frequency effects (Section 2.3.1.), recency, salience, prototypicality, and redundancy (Section 2.3.2.) and the role of 'focus on form(S)' and 'focus on meaning' types of instruction (Section 2.3.3.).

2.3.1. The role of input and frequency effects

Input is the driving force of language acquisition within usage-based and emergentist theory. Part of the human's ability to schematize is believed to be importantly connected to their large

memory capacity and sensitivity to linguistic input and input frequency (Tyler, 2010). Research on frequency effects (e.g., Bybee & Moder, 1983; Bybee, 2003, 2006; Ellis, 2002a, 2002b, 2008d,e) has actually demonstrated that humans are sensitive to the frequency with which they have encountered words, and the constructions in which words are encountered.

Similarly, input is arguably the most important construct in S/FLA (Piske & Young-Scholten, 2008). In the same way as children learn their L1, the second/foreign language learner will induce the rules of the TL from language input, and the frequency with which the learner encounters and uses particular patterns will determine their order of acquisition and their level of entrenchment. A construction that occurs frequently and cohesively (without structural changes) in the learner's input is believed to be acquired earlier and become more entrenched (Bybee, 2010; Diessel, 2016). This implies that "frequency of use plays an important role in determining cognitive structures" (Bybee, 2008, p. 218).

Two different types of frequency effects can be distinguished: *token frequency* and *type frequency*. *Token frequency* refers to frequency of a concrete expression which "in the language learner's experience tends to entrench that expression in terms of the concrete words and morphemes involved" (Tomasello, 2003, p. 107). It counts the frequency of a particular form in the input and is characterized by three types of effects (Bybee, 2008): the *conserving effect*, the *autonomy effect* and the *reducing effect*. The *conserving effect* implies that frequent constructions tend to withstand conformance to the generalized form based on analogy (e.g. the high-frequency verbs *keep* and *sleep* have acquired the irregular past forms *kept* and *slept*, while the low-frequency verbs *weep*, *leap* and *creep* have developed the regular past forms *wept*, *leaped* and *creeped* alongside their irregular forms). Irregular forms that frequently occur in the learner's input will be more likely used in an accurate way, while infrequent irregular forms will be more likely produced incorrectly as regular forms. The *autonomy effect* refers to the fact that autonomous constructions are so frequent in the input that they can be acquired as chunks without involving a simultaneous acquisition of the constituents making up the construction (e.g. the construction *gimme* may be produced before that the elements *give* and *me* are acquired). The causal relationship between repetition and chunking would lead to the negative phenomenon of fossilization in S/FLA; the learner produces the construction so often incorrectly that it can hardly be adjusted after its automatization (Long, 2003; Ellis, 2006b). The *reducing effect* involves the phonetic reduction of frequent constructions such as pronouns, modals, negatives, auxiliaries, or frequent phrases (e.g. *going to* is often shortened to [aiməʊ] when used with the pronoun "I"). High-frequent words and phrases tend to reduce more than

low-frequent ones. *Type frequency*, on the other hand, “determines the abstractness or schematicity of the resulting construction” (Tomasello, 2003, p. 107). It counts the frequency of distinct realizations of a given construction (Wulff, 2018). The type of a construction needs to repeatedly occur with a particular frequency for the learner to identify its structure (Bybee, 2008). The English past tense morpheme *–ed*, for example, has a very high type frequency since it occurs with thousands of different verbs (in comparison to irregular past tense forms). In relation to S/FLA, a learner will first have to encounter a pattern with a number of different items in it before applying it productively to new forms.

The quantity and quality of language input is therefore crucial in the acquisition of a second/foreign language. However, input is encountered and perceived differently in L1 and TL contexts; while TL input often occurs in instructional settings and written communication, L1 input is encountered in naturalistic settings, alongside intense support from parents (Snow, 1999). In terms of input quantity, TL learners very often experience a lack of exposure to sufficiently rich and varied input so that they may fail to notice some features of the input because these are not very salient or are particularly redundant (Ellis, 2004) (see following section below). Zyzik (2009, p. 57) incidentally considers instructed SLA as a “potential poverty of the stimulus situation” in the usage-based sense of the term. As far as input quality is concerned, a number of studies showed that the instructed input might be distorted in ways that differ from actual usage, and that this would possibly impede the task of input-driven language learning. Most of these studies unsurprisingly dealt with textbooks, as these are one of the most important sources of input for TL learners. Biber and Reppen (2002), for example, found major discrepancies between ESL grammar textbooks and corpus-based frequency findings and concluded that material developers fail to reflect the patterns of actual usage in the textbooks (e.g. nouns functioning as adjectives, such as *leather seat*, are very frequent in written language but are rarely tackled in textbooks). Goodall (2008) reported similar findings when comparing the input found in Spanish textbooks with data from a corpus; the coverage of the present progressive form and the different forms of reflexive verbs in the textbooks was not in line with the frequencies of occurrence of the constructions in authentic language use. Besides pedagogical materials, the other source of input for learners in classroom environments is the teacher. Meara, Lightbown, and Halter (1997) examined the vocabulary available in the speech of various ESL teachers. They found a low number of unusual words in the teachers’ speech, which would indicate poor classroom vocabulary. According to Bybee (2008, p. 233), we should not expect instructed settings to reflect naturalistic contexts in terms of quality and

quantity of input: “natural frequency distributions do not have to be exactly reproduced in the classroom”. High-frequency items are so pervasive in the input that acquiring them should not be a problem, even in instructed settings. Infrequent ones, by contrast, will need more attention.

2.3.2. The role of recency, salience, prototypicality, and redundancy

Other factors besides token and type frequency effects play a facilitating role in the acquisition of language constructions. This is the case of *recency*, *salience*, *prototypicality*, and *redundancy effects*. *Recency* or *priming effects* refer to how recently a construction has occurred in the input: “in a situation where two variants alternate, a variant is more likely to appear if it has been preceded by the same variant shortly before.” (Ellis, 2016).

The *salience* of a feature will also affect its acquisition process. According to Terrell (1991), some grammatical form-meaning pairings have a low perceived saliency and are therefore less readily learned. This is the case of grammatical particles and inflections (e.g. the third person singular -s of English) (Goldschneider & DeKeyser, 2001; Ellis, 2006b, 2008c).

Prototypicality assumes that “information concerning a category is represented in such a way that some of its instantiations have a more important cognitive status than others” (Hu, 2002, p. 357). The members of a category that show a high level of similarity with other members of the same category (Rosch & Mervis, 1975) and that are frequently encountered (Barsalou, 1985) are considered as prototypes. In relation to language learning and use, prototypical members have been shown to be acquired earlier and more easily than peripheral members, and to be processed more quickly and accurately (de Villiers, 1980; Lakoff, 1987; Taylor, 1998). Ellis and Ferreira-Junior (2009), for example, showed that second language learners first used prototypical and generic verbs in verb-argument constructions (*go*, *put* and *give*).

Regarding *redundancy*, some grammatical form-meaning pairs are not only low in salience, but they are also redundant for the understanding of an utterance, which explains that they tend to pass unnoticed by the learners. Tense-marking morphemes are a case in point; they are not salient, nor essential, when they occur together with a temporal adverb, as in the utterance “*Hier nous sommes allés au cinéma* (‘Yesterday we went to the movies’)” (Terrell, 1991, p. 59). They will therefore be more hardly acquired than the temporal adverbs fulfilling the same function.

2.3.3. Focus on form(S) and focus on meaning

If a particular construction is neither frequent, salient, prototypical or has not been encountered recently, a form-focused type of instruction may help the learner notice the correct form of the construction in his/her TL input (Ellis, 2013).

Different instructional strategies succeeded one another over the years, going from the traditional *focus on formS* approach, to a radical shift with the *focus on meaning* approach, and finally leading to the in-between *focus on form* approach (Long, 2000). The *focus on formS* approach to teaching focuses on the TL as object. It divides the TL into different types of segments (phonemes, words, collocations, notions, etc.) and presents them to the learners one at a time following a specific order based on frequency, valency or difficulty. The learners are then expected to synthesize all the linguistic items and to use them in a native-like way. Common practices of the *focus on form* approach involve explicit grammar rules, memorization of short dialogues, repetition of models, etc.

In response to some major problems related to this approach, the focus radically shifted from the TL towards the learner and learning processes (Long, 2000). In fact, the *meaning-focused approach* considers the TL as a medium to communicate, while acquiring something else (a content); the new language is believed to be acquired incidentally and implicitly, in the exact same way as in L1 acquisition. The TL learning conditions should therefore be very similar – if not identical – to the L1 learning conditions, i.e. purely communicative content-based lessons during which the learner has to subconsciously induce the rules of the TL by him-/herself from positive evidence encountered in the input.

A third approach that lies somewhere between the two is the *focus on form approach*, which temporarily draws the learners' attention to problematic linguistic elements in context as they incidentally arise in lessons whose main focus is on meaning/communication (Long, 2000). It is the learner's developing language system that determines which forms are targeted and when; it is not scheduled in advance. The goal is to induce *noticing* (Schmidt, 1993, p. 7), that is “registering the simple occurrence of some event” in order to store it in memory (in comparison to *understanding*, which involves recognizing a general principle, rule or pattern and figuring out how it is organized in the linguistic system). The *focus on form approach* is not to be confused with *form-focused instruction*, which is an umbrella term for all the procedures and activities used to draw the learners' attention to form.

The relationship between the three approaches is very well described by Doughty and Williams (1998, p. 4):

We would like to stress that focus on formS and focus on form are not polar opposites in the way that 'form' and 'meaning' have often been considered to be. Rather, a focus on form entails a focus on formal elements of language, whereas focus on formS is limited to such a focus, and focus on meaning excludes it. Most important, it should be kept in mind that the fundamental assumption of focus-on-form instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across.

It is now acknowledged that the combination of a form-focused and meaning-focused type of instruction seems to be the most beneficial for language acquisition. As Ellis (1995, p. 136) stated, “a blend of explicit instruction and implicit learning can be superior to either just explicit instruction or implicit learning alone”. In relation to the focus of our present research, i.e. phraseology, the previous chapter has shown that implicit learning through a substantial number of input is clearly needed, but that it has its limitations (especially for low-frequency multi-word units) and that more explicit forms of teaching proved beneficial. The CLIL approach (that is fully described in Chapter 3) actually aims to integrate both language (form) and content (meaning). This being said, CLIL programs can be placed on a continuum depending on their focus and a number of studies looking at CLIL teaching performances have shown a rather unbalanced focus on meaning at the expense of form. Pérez-Vidal (2007) looked at teachers’ strategies in CLIL classrooms in Spain and showed that CLIL teachers mostly paid attention to meaning; they promoted interaction, encouraged learners’ output and mainly referred to the subject content. They devoted little time to form: “no FoF⁶ strategies are used: language is seen as the means to communicate, not an end in itself” (Pérez-Vidal, 2007, p. 49). Similarly, De Graaff, Koopman, Anikina & Westhoff (2007) analyzed CLIL teaching performances in The Netherlands and reported no evidence of explicit form-focused instruction in the CLIL lessons. Examples of more implicit forms of feedback were found (e.g. recasts, confirmation checks or clarification requests), but they were not very numerous. Besides, Lyster (2004) who studied teaching practices in Canadian immersion programs, stated that types of corrective feedback such as recasts (i.e. corrective rephrasing) tend to be problematic in immersion classrooms; because of the primary focus on meaning, learners do not notice when the teachers attempt to

⁶ Focus-on-Form, referred to by Pérez-Vidal as “Explicit out-of-content teaching of the L2. Sentences generated to focus on a grammar point, reference to phonics, presentation of lexis, negative feedback” (2007, p. 47).

focus on form. Koopman et al.'s small-scale research project (2014) carried out in CLIL classrooms in The Netherlands also reported a restricted focus on form. Several CLIL teachers explained that the teaching of formal, linguistic features was the domain of the language teachers. As a result, voices have been raised (Lyster, 2007) to complement current practices in CLIL education with a greater focus on the language and its formal aspects. The few studies that compared the phraseological competence of CLIL and NON-CLIL learners (see previous chapter) found better performances in the CLIL group overall in terms of productivity but less so in terms of accuracy, which could be the result of a lack of focus on form in CLIL classrooms.

After having presented a number of crucial factors related to language input and different types of instruction, we now focus on important internal factors that each individual learner brings with him/her to the learning situation.

2.4. Important learner-related factors in second/foreign language acquisition

The learner-related aspects discussed in this section involve age (Section 2.4.1.), motivation (Section 2.4.2.) and cognition (Section 2.4.3.).

2.4.1. Age

Penfield (1963) and Lenneberg (1967) are considered as the two great founders of research on the age factor in language learning and particularly of the *Critical Period Hypothesis* (CPH). According to this hypothesis, children are born with an innate language faculty and there exist a critical age from which this faculty starts to wither so that it becomes difficult to acquire proficiency in a TL after that age. Evidence in favor of these views mostly come from research conducted with immigrants in naturalistic language learning settings. The studies concluded that even though the older starters showed a faster learning rate at the beginning, they were eventually caught up and surpassed by younger starters (e.g. Snow & Hoefnagel-Höhle, 1978). On the whole, “in the case of immigrant children, earlier does seem better, but only in the long run, and only when children are given plenty of time and opportunity to make the most of the abundant language input they are exposed to” (Myles, 2017, p. 2). More recent studies (for example Kinsella & Singleton, 2014), however, demonstrated that it was possible for adult second language learners to achieve native-like proficiency levels too. More importantly, they showed that linguistico-cultural identity and socio-affective factors were more crucial than maturational factors. Jia and Aaronson (2003) and Moyer (2013, 2014) showed that the quantity and quality of the language learning experience had an impact on the eventual level of achievement in the new language and that it differed according to age; children seem to enjoy

the L2 experience more and seem to associate with the L2 identity to a greater extent than adolescents.

The findings of research on age and SLA in naturalistic settings were assumed to apply to foreign language acquisition. Nevertheless, studies that analyzed the outcomes of introducing a foreign language early in instructed contexts failed to verify the capacity of early instruction to deliver higher proficiency levels (e.g. Cenoz, 2002, 2003; García Lecumberri & Gallardo, 2003; García Mayo, 2003; Lasagabaster & Doiz, 2003; Muñoz, 2003; Navés, Torras & Celaya, 2003; Perales, García Mayo & Liceras, 2004; Álvarez, 2006; Miralpeix, 2006; Mora, 2006; Muñoz, 2006; Torras, Navés, Celaya & Pérez-Vidal, 2006; Kalberer, 2007; Jaekel, Schurig, Florian & Ritter, 2017). In fact, they showed that older learners were faster than younger learners (as in naturalistic contexts), and that this advantage remained in the mid or long term (in contrast to naturalistic contexts). The very few studies that found a small advantage for an early start either involved instructed settings with many hours of instruction per week (e.g. Larson-Hall, 2008) or concerned learners who had received substantial parental support (Pfenninger and Singleton, 2018) (which once again confirms the prominence of affective factors).

Muñoz (2008) gave several reasons that may explain why the results differ according to the language learning setting. First, the difference between rate of learning and ultimate attainment in the two settings should be considered. Investigating ultimate attainment in foreign language learning contexts cannot analyze the same time span as in naturalistic contexts. In fact, examining long-term effects with instructed learners is difficult; they can be followed during their school years, but once they finish school their learning trajectories may greatly vary in type and amount of exposure. Second, the amount and type of exposure to the language differs according to the setting. The limited exposure provided in foreign language settings would not make it possible to confirm the benefits of an early start and would prevent learners from attaining nativelikeness. According to Muñoz (2008), amount of input never ceases to be a determinant factor in foreign language learning settings. Recent research findings actually indicated that amount and type of input may play a more crucial role than starting age in long-term achievement. Muñoz's study (2011) showed no correlation between starting age and language outcomes, but between language outcomes and a number of input measures (length of instruction in years, number of recent curricular and extracurricular lesson hours, number of hours spent abroad, and current contact with the target language). Finally, Celaya (1991 as cited in Celaya, 2012) also suggested differences in communicative needs between the two learning

contexts. Learners in naturalistic learning contexts need to master the language for real life communication, and this plays a decisive role in language acquisition. The importance of motivational aspects is dealt with in the following section (Section 2.4.2.).

Educationally oriented research on length of exposure has emphasized the fact that starting early brings a time advantage rather than a critical-period advantage. The belief that earlier is always the better would largely be a matter of time spent learning the language rather than the actual age of starting learning (Carroll, 1969; Genesee, 1988; Blondin et al., 1998; see Johnstone, 2007 for a review). In Johnstone's review (2007) of language education findings, time – in the sense of 'time for learning' and 'time on task' – was systematically found to positively relate to successful learning.

Studies comparing different types of Canadian immersion programs according to their starting time (Swain & Lapkin, 1989; Turnbull, Lapkin, Hart & Swain, 1998) reported that even though early immersion learners appeared to reach higher TL proficiency levels overall in comparison to middle and late immersion learners, the differences were not as large as expected and not always significant. Early immersion learners were found to be better at listening and speaking skills, but late learners seemed to perform better on literacy-related tests. Swain and Lapkin (1989) explain the late immersion learners' better literacy-related skills by the fact that they have already learned how to read and write in their L1, and that they can "apply their knowledge of literacy functions and skills to their developing competence" in the TL (1989, p. 152). This echoes Cummins' research (1981) that indicated that older learners would eventually learn a new language more efficiently thanks to greater cognitive-linguistic skills and language learning strategies. Older learners have developed a higher cognitive academic linguistic proficiency (CALP) in their L1, i.e. academic skills and concepts, and they can apply these to the TL when they start learning it. This transfer of skills from one language to the other would speed up the process of TL acquisition. Swain and Lapkin (1989) also mentioned a greater cognitive maturity; older learners can better abstract, generalize and classify as well as "reflect on their own language use and [...] attend consciously to language per se" (1989, p. 152).

Based on these findings, we can work on the assumption that the CLIL approach – that is usually considered closer to the late immersion programs implemented in secondary education in Canada (Lasagabaster & Sierra, 2009a, p. 271) – contains the key ingredients for the successful learning of a TL. As we will see in the following chapter (Chapter 3), CLIL programs provide foreign language learners with more exposure time to the TL and make the learning process more meaningful since the mastery of the TL is needed to acquire a subject content. In

comparison to learners in traditional foreign language learning classes, the CLIL learners will have spent more time learning the TL in a more authentic environment.

2.4.2. Motivation

S/FLA research now recognizes language attitudes and motivation as an important factor in the acquisition of a new language. In fact, learning a second/foreign language not only involves learning rules; it is a deeply social event as the new language is socially and culturally bound, and therefore requires the consideration of a wide range of features of the TL culture (cf. Gardner, 1979; Williams, 1994).

Theory on motivation introduced key aspects such as the *integrative motive*, defined as “a motivation to learn a second language because of positive feelings toward the community that speaks that language” (Gardner, 1985, pp. 82-83). These positive feelings may, in extreme cases, lead to complete identification with the other community. In foreign language learning contexts, where learners are not in contact with a salient TL group in their environment, they will rather identify with the intellectual and cultural values associated with the TL, or with the TL itself. Gardner (1985) also distinguished between *integrative motivational orientation* and *instrumental integrative orientation*. The former refers to the aspiration to be part of the TL community, while the latter denotes a functional or practical motivation (the TL is a good means to find a job, earn money, etc.).

A large number of studies analyzed language attitudes in terms of attitudes towards the TL community, towards the learning situation, and towards learning the TL. However, few of them have researched attitudes towards the TL itself, usually considering it as an underlying part of motivation, while this is of particular relevance in some contexts such as Belgium. In fact, research carried out in Belgium found that TL proficiency was related to attitudes towards the TL itself and its community (Mettewie, 2004); learners showing positive attitudes towards the TL and/or its community appeared to be more proficient in the TL. However, in the context of Belgium where strong political tensions have been growing since 2010-2011 (Mettewie, 2015), attitudes towards the other language and/or community (Dutch for French-speaking learners and French for Dutch-speaking learners) tend to go from neutral to negative. In fact, French-speakers often consider Dutch as a difficult, unattractive language – especially compared to English.

The CLIL approach (which is described in further detail in the next chapter) would enhance motivation and positively influence language attitudes (Coyle, 2005; Dalton-Puffer, 2008;

Lasagabaster, 2009), thanks to its meaningful language learning situations, its focus on meaning, its space for interactions, and the potential transfer from positive attitudes towards the content subject taught in the TL to the TL itself (Harrop, 2012). Most of the existing research in CLIL showed higher scores of motivation and language attitudes among CLIL learners (e.g. Lasagabaster & Sierra, 2009b in Spain; Pirskanen, 2009 in Finland; Lorette & Mettwie, 2013 in French-speaking Belgium).

The research project that encompasses the present thesis involves two different target languages (viz. English and Dutch), two different learning conditions (CLIL and NON-CLIL settings) and deals with the particular context of French-speaking Belgium. Dutch and English actually have a different TL status in the French community of Belgium (*Communauté française*). Dutch is one of the national languages and could therefore be considered as a second language, but French-speakers in Belgium generally encounter Dutch infrequently in day-to-day life. Dutch is spoken in the Flemish community, and the different communities have relatively much political autonomy – the education systems are organized at communal level instead of national level. As a consequence, the learners in the French community may be exposed as little to Dutch as to English. They may even be exposed more often to English than to Dutch, since English is an international language closely linked to youth culture on radio, television, the internet, etc.⁷

Based on these elements, and knowing that the specific learning context of CLIL appears to have an impact on the learners' attitudes and motivation, one of the project's research lines is devoted to socio-affective variables (see Chapter 3, Section 3.4. for a more detailed description of the project). Studies carried out within that research line (De Smet, Mettwie, Galand, Hiligsmann & Van Mensel, 2018; De Smet, Mettwie, Galand, Hiligsmann & Van Mensel, submitted) showed that in terms of emotional engagement in the classroom, CLIL learners experienced less anxiety than NON-CLIL learners, and learners of English reported significantly more enjoyment and less anxiety than learners of Dutch. In terms of attitudes and motivation, the learners reported more positive attitudes and a higher motivation in CLIL compared to NON-CLIL, and in English compared to Dutch. All these differences mostly appeared at the secondary level.

⁷ In order to avoid any misunderstanding, the term 'target language' (TL) will be used in this doctoral dissertation to refer to English or Dutch as the new languages acquired by the learners in this study.

Some empirical chapters in the present thesis actually reveal differences between English and Dutch (see Chapters 4, 5 and 7). This also seems to hold for other studies carried out in the project and looking at various aspects of learners' TL linguistic proficiency (Hendrikx, 2019; Van Mensel, Bulon, Hendrikx, Meunier & Van Goethem, submitted). Possible interactions between these motivational aspects and our results are not investigated in the framework of the present doctoral research. However, they are taken into account in the discussions of the results.

2.4.3. Cognition

Language is not solely a social construct; it also has a biocognitive and neurocognitive basis (Ullman, 2006). The concept of cognitive aptitude is a composite term used to refer to different cognitive abilities, namely working memory and aptitude for implicit and explicit learning. Working memory is a cognitive system that stores and manipulates information under attentional control (Baddeley, 2007). Aptitude for implicit learning denotes the “ability to recognize and acquire patterns underlying the input through implicit induction” (Yi, 2018, p. 835), while aptitude for explicit learning is the aptitude that relates to explicit learning and that involves “metalinguistic reasoning, explicit induction or analysis” (Yi, 2018, p. 835). Studies have shown that cognitive aptitude was significantly related to overall language learning success in instructed settings (Carroll, 1973; Ehrman & Oxford, 1995; Harley & Hart, 1997) and that individual differences in aptitude resulted in learning differences (De Graaff, 1997; Robinson, 1997; Williams, 1999). Research has demonstrated that working memory as well as implicit and explicit learning aptitudes are related to both TL proficiency development and use (Atkins & Baddeley, 1998; Granena, 2012; Linck, Osthus, Koeth, & Bunting, 2014; Suzuki & DeKeyser, 2017).

According to Kormos (2013), the basic cognitive capacities that are at work when children are learning their L1 are the same for learners acquiring a TL. These cognitive capacities involve memory for verbal material (which helps learners remember words) and phonological short-term memory (which assists in decoding sequences of sounds and associating them with words and their meanings). They also include verbal reasoning skills (which aid in discovering the rules and regularities of the language) and working memory resources (which help regulating attention to relevant linguistic features, maintaining chunks in memory for further processing and inhibiting irrelevant stimuli and automatic response patterns). From a cognitive point of view, learning another language also includes the acquisition of declarative knowledge (e.g. the meaning of words) and skills, i.e. procedural knowledge and its automatic application (e.g. construct a sentence fluently, making use of syntactic building procedures) (Kormos, 2013).

Research has shown that bilingualism appears to be beneficial in terms of cognitive functioning. More specifically, bilinguals would be able to handle a heavier cognitive load at one time. They also seem to have an increased metalinguistic awareness, defined as “the knowledge we have about the structural properties of language, including the sounds, words and grammar of language” (Cloud, Genesee & Hamayan, 2000, p. 3). This heightened metalinguistic awareness allows bilinguals to compare their languages and eventually leads to a more precise use of the language. It can also serve as a tool in language learning, for example “to decode words in a text by drawing on knowledge from both of their languages” (Mehisto & Marsh, 2011, p. 35). Bilinguals also appear to have enhanced problem-solving skills, which could be related to their greater executive control (Bialystok et al., 2005; Bialystok & Barac 2012). This means that they are better at determining which information is worthy of attention (i.e. relevant to solve the problem) and which is not. This is in turn closely linked to inhibitory control; they can ignore the distractions of irrelevant stimuli and therefore think more effectively (McLeay, 2003; Colzato et al., 2008). Finally, bilinguals would have superior episodic and semantic memory (Kormi-Nuori et al., 2008), which are two functions of long-term memory. The former “includes information about such elements as time, place, feelings, and activities”, while the latter “includes general knowledge about, for example, ideas, facts and problem-solving” (Mehisto & Marsh, 2011, p. 36). More recent studies have however failed to find coherent evidence for these advantages (e.g. Paap & Greenberg, 2013; Gathercole et al., 2014).

Learning a TL in an immersive context such as CLIL is believed to be quite demanding at the cognitive level. Immersive settings would lead to competition between the consolidated linguistic representations of the L1 and the newly acquired representations of the TL (Baus, Costa & Carreiras, 2013). Immersion primarily relies on implicit or procedural memory, which has the advantage of being similar to naturalistic language development in children (Morgan-Short, Steinhauer, Sanz & Ullman, 2012). This type of memory is also more resistant to interference and less prone to forgetting. The better results found in terms of TL linguistic outcomes in CLIL may be a result of these memory advantages. Moreover, the immersive contexts make learners divide their attention between the subject content and the form of the linguistic input, both during production and comprehension.

Research findings reported by Simonis (2019), who investigated the cognitive abilities of CLIL and NON-CLIL primary and secondary school learners in the framework of the project encompassing the present thesis, showed, however, no substantial cognitive advantages (or

disadvantages) of learning the TL in the immersive context of CLIL in terms of executive control and attention.

2.5. Conclusion

The present chapter has provided the main theoretical principles that underlie usage-based and emergentist approaches of language acquisition. These approaches do not assume an innate language faculty, but rather see language as a product of accumulated linguistic experiences that is learned by means of powerful cognitive mechanisms (e.g. generalization, categorization, abstraction). The principles underlying these approaches that are important in the framework of this thesis is that prefabricated chunks are at the center of the language learning process since they serve as a basis for the creation of rules and that exposure to language input is crucial, both in L1 and TL acquisition. Various frequency and psycholinguistic effects (recency, saliency, prototypicality, etc.) also determine the way constructions are acquired. Constructions that are frequent, salient, prototypical, and recent in the input are believed to be more easily acquired. This explains why the quantity and quality of the input learners are being exposed to, is decisive. When constructions are not likely to be noticed by learners in the language input – because the input is too limited or because the characteristics of the construction do not make it noteworthy – instructional strategies that are meant to draw the learners' attention to form may be needed. In fact, an ideal S/FLA combines both a meaning- and form-focused type of instruction. This coincides with the CLIL approach, which theoretically focuses the instruction on both the language (form) and the content subject (meaning). Emergentist theories also emphasize the role of variability in language development and we intend to take this into account by assessing learners at different points in time and by looking at different aspects of language use.

We have also seen that learner-related variables can have an important impact on the acquisition of a second/foreign language depending on the learning context. While age of onset was initially considered as crucial, studies looking at language acquisition in foreign instructed settings now agree to say that time spent learning and using the TL language seems to be more essential than an early start, and that socio-affective factors such as motivation are also particularly decisive. CLIL programs, which provide learners with more time of exposure to the TL, appear to enhance learners' motivation and improve their attitudes towards the TL. The fact that motivation appears to be closely related to the TL itself, especially in the context of French-speaking Belgium, is of importance in the present research as it involves two different target languages (viz. English and Dutch). As for cognition, it receives pride of place in the usage-based approach and clearly has an impact in both L1 and TL acquisition. Although some studies

suggested that learners' cognitive abilities could benefit from bilingual contexts, the CLIL learners investigated in the present study did not show any cognitive advantages.

The next chapter describes the characteristics of the CLIL approach, its implementation in French-speaking Belgium, as well as the research project on CLIL that encompasses the present thesis.

CHAPTER 3: *Content and Language Integrated Learning (CLIL) in French-speaking Belgium*⁸

3.1. Introduction

Chapter 3 focuses on *Content and Language Integrated Learning (CLIL)*, the specific language learning context that is at the heart of the present research. Section 3.2. describes the approach in comparison to other types of bilingual education programs as well as to mainstream language learning contexts. Section 3.3. focuses on its implementation in French-speaking Belgium and gives a summary of previous CLIL research in the context of French-speaking Belgium. Section 3.4. presents the ARC project, of which this doctoral thesis forms one line of research. Finally, Section 3.5. explains how the present research on the acquisition of phraseological language contributes to the project.

3.2. CLIL

CLIL stands for *Content and Language Integrated Learning*⁹. It is a specific teaching method where the TL is both used as a tool to teach curricular subjects such as history, science, geography or mathematics, and taught as a subject included in the assessment. In other words, “the non-language subject is not taught in a foreign language but with and through a foreign language” (Eurydice, 2006, p. 8). CLIL is often referred to as a “dual-focused educational approach” (Dalton-Puffer, 2011, p. 194) which aims to foster the learners’ acquisition of both the TL and the subject content.

According to Brown and Bradford (2017), CLIL differs from CBI (*Content-Based Instruction*) and EMI (*English Medium Instruction*) in the sense that these two approaches respectively aim at language learning or content mastery only. This is not entirely true since these approaches can take several forms. In their description of the CBI approach, Met (1999) and Tedick and Cammarata (2012) actually demonstrated the existence of a continuum where different CBI programs are placed between the two extremes according to their focus on content or language and their degree of intensity (for a lengthier discussion, see the two above-mentioned references). As a result, it is possible that some CBI programs will share the same goals as the CLIL approach.

⁸ Some parts of sections 3.2., 3.3. and 3.4. in this chapter are literally taken from a joint publication of the team members of the ARC project (Hilgsmann et al., 2017).

⁹ The acronym used in French is EMILE: *Enseignement de Matières par l'Intégration d'une Langue Etrangère*.

In his survey of language promotion by European supranational institutions, Baetens Beardsmore (2009, p. 208) points out that:

[t]here are at least thirty-three different designations for some type of “bilingual education”, which may be the term used in some countries but which tends to be avoided at the European level, given that in certain countries it has a negative connotation (sic). The term “immersion”, though regularly used in some countries, is not favored either, given that this tends to be associated with the Canadian efforts, whereas many of the European initiatives either predated the immersion phenomena or were developed independently and with very different goals and methodologies.

In sum, language immersion first came to prominence in Canada in the 1960’s (Cummins & Hornberger, 2008), while CLIL has become a popular and widespread practice (and term) in Europe, advocated by the European Union (Coyle, Hood & Marsh, 2010). Lasagabaster and Sierra (2009a, pp. 369-373) highlighted the similarities and differences between language immersion and CLIL (see also Baker, 2011 and García, 2009, pp. 111-136). Although they share the same objective of fostering proficiency in a new TL, CLIL programs are concerned with *foreign* languages, that is languages that are only used in a formal context by the learners (for example at school) and that are not commonly spoken by the community as a whole. In contrast, language immersion is involved with *second* languages, i.e. languages that are typically used within the local context of the learners (be it at home, in the society at large, or both). Another difference lies in the starting age of the learning of the TL:

the vast majority of immersion programs are of the early (starting age) immersion type, whereas the CLIL approach shows certain similarities with the late immersion programs implemented in secondary education in Canada. This means that the amount of exposure to the local and the foreign language is far from comparable. In fact, CLIL programs endeavor to develop the language skills of students who have had traditional foreign language teaching throughout their primary education (Lasagabaster & Sierra, 2009a, p. 371).

Moreover, language immersion is seen as “a model under which the second language is added to the person’s repertoire and the two languages are maintained” (García, 2009, p. 52) (see also Aronin & Singleton, 2012, pp. 109-114). CLIL, by contrast, is considered as a more dynamic form of bilingualism or plurilingualism assuming that “language use in the twenty-first century

requires differentiated abilities and uses of multiple languages as citizens cross borders either physically or virtually” (García, 2009, p. 54).

CLIL programs not only vary from one country/region to another (see for example Baetens Beardsmore, 1993; Eurydice, 2006; Marsh & Wolff, 2007; Tedick, Christian & Fortune, 2011), but also within the same region (e.g. for French-speaking Belgium, see below). Comprehensive overviews of all the forms CLIL can take are found in Baetens Beardsmore (1993), Marsh and Wolff (2007), Cummins and Hornberger (2008), García (2009), Coyle et al. (2010), Dalton Puffer, Nikula, and Smit (2010), Baker (2011), Ruiz de Zarobe, Sierra, and del Puerto (2011), Aronin and Singleton (2012). The basic criteria that need to be taken into account – no matter the type of bilingual education – are the following:

- the choice of the target language;
- the class time spent in the target language;
- the age of the learners (a distinction is made between early [starting in the third year of nursery school or first year of primary school], middle [starting in the third year of primary school] or late immersion [starting in the first year of secondary school]);
- the content subjects that are taught in the target language;
- the target language itself as an additional and separate subject;
- the main goals (linguistic, educational, socioeconomic and sociocultural).

The main educational aim of CLIL is the fostering of bi-/multilingualism, i.e. the development of learners’ communicative competence or language proficiency in the new TL. Next to this educational goal, the learning of a new TL is also encouraged for various societal reasons (see García, 2009; Baker, 2011). First, to preserve a language that is slowly disappearing (for example the teaching of Basque to Spanish-speaking Basque children in the Basque Autonomous Region). Second, to assimilate minority language groups into mainstream society (for example the teaching of Dutch as a second language in the Netherlands). Third, to obtain increased harmony between language groups (for example French-speaking learners of English and English-speaking learners of French in Canada). Fourth, to enhance economic and employment opportunities. These reasons include extra-curricular goals (e.g. reading books in another language outside the classroom) and learner-centered goals (the learner’s mental development as individual) (Cook, 2002). Furthermore, recent studies have also shown that CLIL “is an effective means of facilitating primary school students’ second language acquisition without undermining their competence in their first language” (Cheng, 2012, p. 379;

see also Braun & Vergallo, 2010; Van de Craen, Mondt, Ceuleers & Migom, 2010; Mattheoudakis, Alexiou & Laskaridou, 2014; Merisuo-Storm & Soininen, 2014; Nieto-Moreno-de-Diezmas, 2018).

In comparison to mainstream language learning contexts, CLIL programs provide learners with a larger amount of TL input since some content subjects are given through the TL instead of the L1 (the proportion of the curriculum devoted to CLIL in French-speaking Belgium is discussed in the next section). Moreover, learners in CLIL programs are in contact with more authentic and directly relevant input (Coyle, Hood & Marsh, 2010) since the TL is used as a means to an end – a tool used to learn a content subject. In traditional TL classes, however, the teaching of the TL is at the heart of the lessons. This makes the learning environment seem more artificial and the input less immediately relevant (Coyle et al., 2010). Furthermore, even though the CLIL approach tends to pay attention to form, it is in essence communicative (with a focus on meaning) (Lasagabaster & Sierra, 2009a) and therefore gives learners more opportunities for active involvement (Nikula, 2010). These specific CLIL features are in agreement with various theories of language acquisition, such as Krashen's (1985) input hypothesis, Long's (1996) focus on interaction, and Swain's (1995) output hypothesis, as the input/interaction/output nexus is constantly activated.

It is therefore not entirely surprising to see recently published surveys' findings showing CLIL learners outperforming NON-CLIL learners as far as TL test scores are concerned (Lorenzo, Casal, and Moore, 2005; Admiraal, Westhoff & de Bot, 2006; Zydati, 2007; Lasagabaster, 2008; Ruiz de Zarobe, 2008, 2010; Llinares & Whittaker, 2010; Dalton Puffer, 2011; Gené-Gil, Juan-Garau & Salazar-Noguera, 2015; Martínez Adrián & Gutiérrez Mangado, 2015; Sylvén, 2019). As for the effect of CLIL on L1 development, it appears to be a source of concern for parents (e.g. Pladevall-Ballester's 2015 study on CLIL in Catalonia). Many of them believe that following between 50 up to 75% of the education in a TL might be detrimental to the acquisition and mastery of their children's L1 (Van de Craen, Surmont, Ceuleers & Allain, 2013, p. 252). Nevertheless, studies from the 1970's onwards have shown that CLIL does not seem to have a negative impact on the L1 (e.g. Genesee, 1989 as cited in Genesee, 1991; Marsh, 2002; Wesche, 2002; Knell et al., 2007; Braun & Vergallo, 2010; Merino & Lasagabaster, 2015; Pérez Cañado, 2018). Some researchers even reported a positive influence of CLIL on the L1 (Harley et al., 1986 as cited in Wesche, 2002; Vesterbacka, 1991; Merisuo-Storm, 2002; Merisuo-Storm & Soininen, 2014).

Despite the very abundant literature on CLIL (see for example Baetens Beardsmore, 1993; Marsh & Wolff, 2007; Cummins & Hornberger, 2008; García, 2009; Coyle et al., 2010; Dalton Puffer et al., 2010; Baker, 2011; Ruiz de Zarobe et al., 2011; Aronin & Singleton, 2012), it nonetheless remains largely unclear how much, in what respect and thanks to which (internal and external) processes/factors CLIL students are better than traditional learners (cf. Dalton Puffer, 2011). As stated by Dalton Puffer (2008, p. 139):

CLIL is still far from being a consolidated and fully articulated educational model [...] A great deal more needs to be done, for instance, in order to consolidate the theoretical underpinnings of CLIL and create a conceptual framework that is both coherent and applicable to different local conditions” (see also Van de Craen, Mondt, Allain, and Gao, 2007, 2010; Ruiz de Zarobe et al., 2011).

In recent years, voices are also beginning to rise downplaying the conclusions drawn from CLIL research (see Seikkula-Leino, 2007; Bruton, 2011; Ruiz de Zarobe, 2011; Pérez Cañado & Ráez Padilla, 2015). To this day, we only have an incomplete and fragmented view on whether and/or how CLIL differs from traditional TL education. In fact, limited empirical research has thus far been conducted to evaluate CLIL effectiveness in relation to learners’ linguistic achievement, their cognitive development and the teaching and learning processes with regard to teacher education (see Coyle, 2007; Coyle et al., 2010; Cheng, 2012). Indeed, as De Graaff, Koopman, Anikina, and Westhoff (2007, p. 605) states, “little attention is paid to the pedagogic repertoire of CLIL teachers and how this contributes to the pupils’ target language proficiency”.

The next section focuses on the specificities of CLIL in French-speaking Belgium and provides an overview of previous research on CLIL in the context of French-speaking Belgium.

3.3. CLIL in French-speaking Belgium

Although Belgium can be considered a multilingual country on a societal and governmental level, comprising four linguistic regions (the Dutch-speaking region, the French-speaking region, the bilingual region of Brussels-Capital and the German-speaking region, see *The Belgian Constitution*. Article 4), education is still largely set up as monolingual by each of the three “language community” governments, with languages traditionally being taught in foreign language classes. In recent decades, however, the growing perception of the importance of language skills on the job market (Mettewie & Van Mensel, 2009) alongside the discontent with ‘traditional’ types of language education (see e.g. Ginsburgh & Weber, 2007) have led to an increased number of educational stakeholders to call for bilingual education, resulting in the

organization of CLIL-type bilingual education programs. As can be drawn from the overview in Table 3.1., Dutch-medium education in Flanders only recently adopted a policy allowing for CLIL education, and only in secondary education¹⁰. The reasons for the reluctance in implementing such programs are mainly political and ideological in nature (Bollen & Baten, 2010). An exception to this reluctance can be found in Dutch-medium education in Brussels, where a limited number of schools have been participating in the STIMOB-project (Stimulerend meertalig onderwijs in Brussel), which was set up in 2001 to offer CLIL-type immersion (Dutch/French) to the linguistically highly diverse pupil population (see e.g. Van de Craen, Lochtman, Ceuleers, Mondt, and Allain, 2007).

Table 3.1. Implementation of CLIL in different parts of Belgium

Part of Belgium		Official CLIL since	# CLIL-programs: Primary/Secondary		
French-speaking	Wallonia + Brussels	1998	191	100	(1)
Dutch-speaking	Brussels (STIMOB)	2001	10	2	(2)
	Flanders	2014	/	60	(3)

(1) Statistics for school year 2015-2016 (see *enseignement.be*)

(2) Statistics for school year 2013-2014 (provided by *Scholengroep Brussel*)

(3) Statistics for school year 2016-2017 (see *onderwijs.vlaanderen.be*)

The French-speaking community of Belgium, responsible for French-medium education in Brussels and Wallonia, has allowed bilingual education programs since 1998¹¹. The compulsory legal framework for this was scaffolded in the *Décret relatif à l'enseignement en immersion linguistique*¹² in 2007 and has been regularly amended (2003, 2008, 2009, 2010 and 2011). The Decree stipulates that schools can offer a CLIL program (often labelled 'immersion') from kindergarten onwards, under certain organizational and pedagogical conditions¹³. These include (see *enseignement.be*):

¹⁰ A maximum of 20% of the curriculum (not including the foreign language class) may be taught in the target language (French, English or German); CLIL subjects must be taught in Dutch as well (see *onderwijs.vlaanderen.be*).

¹¹ See http://www.galilex.cfwb.be/document/pdf/22229_018.pdf (accessed 22 May 2019).

¹² See http://www.galilex.cfwb.be/document/pdf/32365_003.pdf (accessed 22 May 2019).

¹³ Note that CLIL can be organized for all types of education (general, technical, vocational, artistic).

- a) Schools (direction or organizing power) have to submit a proposal to the Ministry of Education of the CfB.
- b) Upon acceptance, the CLIL project should explicitly be mentioned as being part of the School Project ('projet d'établissement').
- c) No official selection procedure is allowed; pupils should be enrolled according to the chronological order of enrollment¹⁴.
- d) Normally, the authorization from the CfB is valid for three years, and can then be extended upon request.
- e) Subjects in the target language are preferentially taught by a native speaker or by someone with a nativelike level in the target language. In order to be allowed to teach in a CLIL track, teachers have to fulfill three requirements: they need to be in possession of a pedagogical degree, they need to have an advanced knowledge of the target language, and they need to have a 'functional' knowledge of French.

However, provided these conditions are met, schools are in fact relatively free in deciding which CLIL program they wish to provide, and this has led to a wide range of CLIL-formulas throughout the CfB. These may differ with respect to (see *enseignement.be*):

- a) The target language: schools can offer a CLIL program in Dutch, English or German. An exception to this rule applies to the schools situated in the officially Dutch-French bilingual Region of Brussels, as well as those in a few communities along the language border (Comines-Warneton, Mouscron, Flobecq, and Enghien), where only Dutch is allowed as the target language¹⁵. Furthermore, schools are allowed to offer up to two CLIL language tracks, but a pupil can only attend one CLIL program;
- b) The age of onset: *Early CLIL*: 3rd year kindergarten, 1st year primary, or 3rd year primary; *Late CLIL*: 1st year secondary, or 3rd year secondary;
- c) The proportion of the curriculum provided in the target language: the number of weekly hours that the school may provide in the target language varies from 25% to 75% of the curriculum. Primary: between 8 and 21 hours per week (including the foreign language class). Secondary: between 8 and 13 hours per week (including the foreign language class);

¹⁴ This is different from the situation in other countries, e.g. Germany (Rumlich, 2016) or The Netherlands (Maljers, 2007).

¹⁵ From Year 3 in secondary education onwards, the third language (English) may also be offered in the Brussels Capital Region.

- d) The subjects that are taught in the target language: with the exception of religion or ethics and the French language class, all subjects may be taught in the target language;
- e) The choice of the language for initial literacy instruction;
- f) Other curricular differences that are subject to the authority of the organizing body of the schools.

There are obvious advantages to the flexibility afforded by the official educational bodies. Schools can for instance easily adapt to local needs and demands from parents and other stakeholders, such as the implementation of a German CLIL program near the border with Germany or a Dutch CLIL program near Flanders or the Netherlands. They can equally adjust their program according to the availability of teachers that are adequately trained to teach both language and subject. On the downside, however, support from the Ministry of Education to address the numerous challenges these schools have to cope with is limited. Specific teacher training, for instance, was until very recently non-existent¹⁶. In the absence of any official guidelines for immersion teaching in the CfB, schools and teachers have worked with academics to compile their own pedagogical toolkits ('CLIL/ EMILE Digest', Bya & Chohey-Pacquet, 2005). Furthermore, teachers regret the paucity of adequate didactic materials available (Vandeputte, 2011). Finally, many researchers attest to the organizational and logistical challenges that schools face when introducing a CLIL track in terms of classrooms and timetables, as well as regards finding (or substituting) adequate teachers (Blondin & Straeten, 2002; Chohey-Pacquet, 2007, 2008; Bouillon & Descamps, 2011; Buyl & Housen, 2014).

The number of schools offering a CLIL program and pupils attending these programs have been steadily increasing since 1998 (see Figure 3.1.). In 2015-2016, a total of 191 primary schools and 100 secondary schools offered a CLIL-program (see *enseignement.be*) to more than 32,000 pupils (figure for school year 2013-2014, see Chohey-Pacquet, 2015). The proportion of CLIL pupils within the total pupil population of the CfB is around 6% for primary schools and 3.9% for secondary schools.

¹⁶ Since 2014, the Université de Louvain has offered the possibility to obtain an 'immersion teaching certificate' (*Certificat didactique de l'enseignement en immersion*). See <https://uclouvain.be/fr/etudier/iufc/didactique-de-l-enseignement-en-immersion.html> (accessed 21 May 2019). The Université de Mons offers a similar program (*Certificat d'Université de spécialisation en méthodologie de l'immersion linguistique*). See <https://web.umons.ac.be/langage/training/specialisation-en-methodologie-de-limmersion-linguistique/> (accessed 21 May 2019).

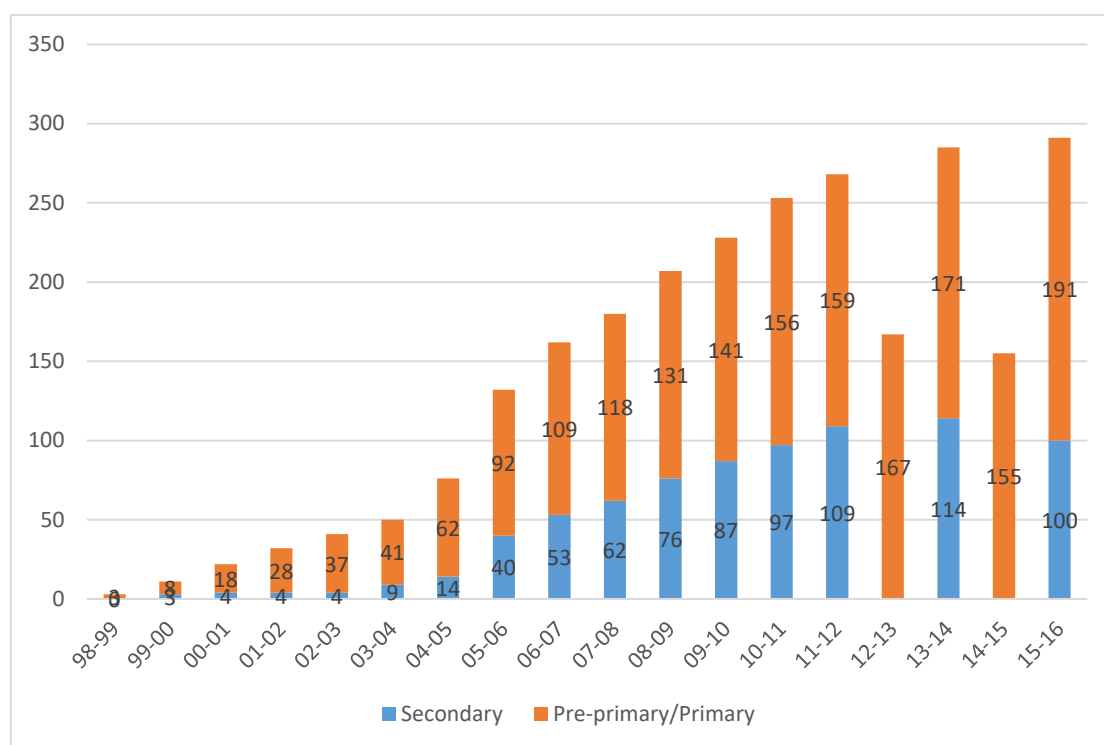


Figure 3.1. Evolution of number of schools providing a CLIL program in the CfB (our calculation, based on various documents from *enseignement.be*, no figures available for secondary 2012-2013/2014-2015)

As shown in Figure 3.2., the majority of pupils enrolled in CLIL schools follow a Dutch language track. CLIL in German is almost exclusively provided in the province of Liège, near the German border. CLIL in the Brussels Capital Region is, as is to be expected (cf. the above-mentioned legal framework), almost exclusively provided in Dutch.

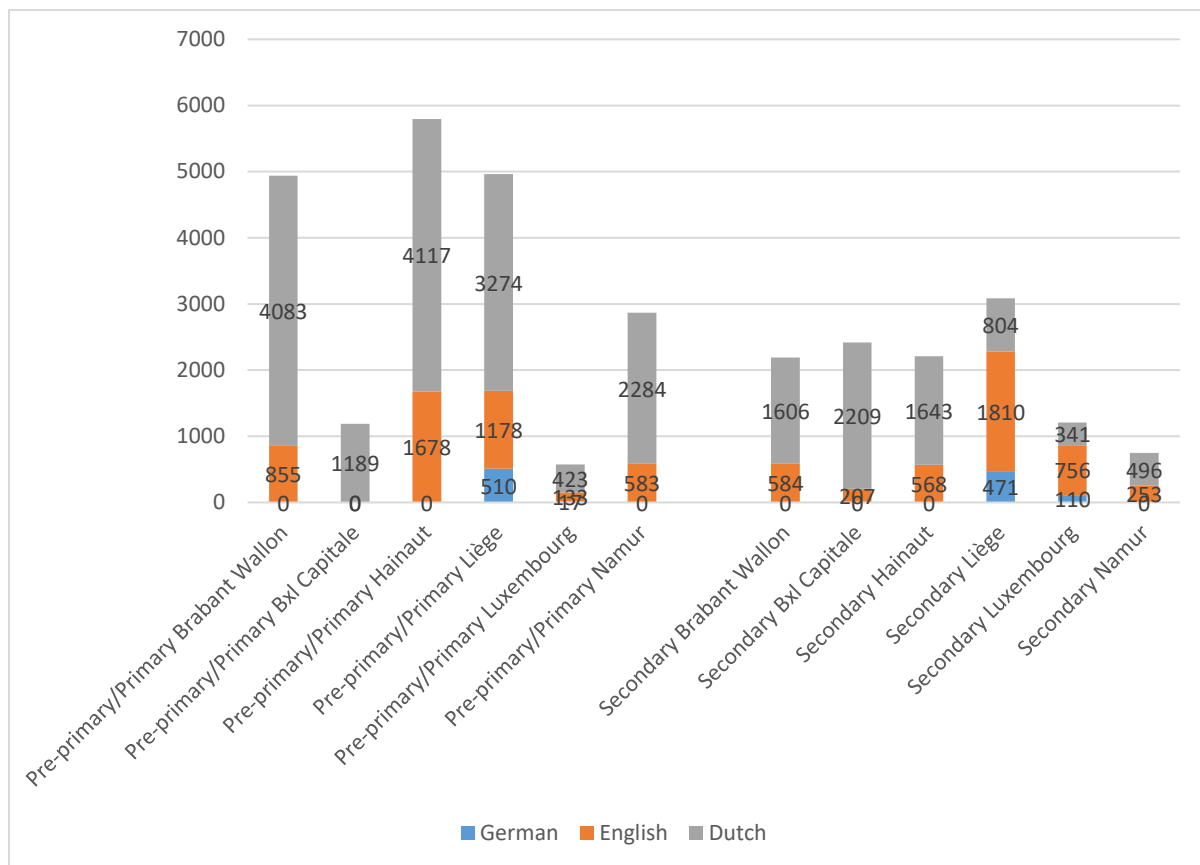


Figure 3.2. Number of pupils enrolled in CLIL programs in the CfB, per province and target language (our calculation, based on various documents from *enseignement.be*, figures for school year 2013-2014)

Although the first CLIL programs were set up almost two decades ago, research on CLIL in schools in the French-speaking Community in Belgium is relatively scarce, and not often published in peer-review journals. All in all, the few studies that were carried out project a rather positive image, even if some results are inconsistent. The findings can be summarized as follows: first, CLIL learners seem to display similar (or even higher) levels of competence in the L1 when compared to their counterparts in the French-medium track (Comblain & Rondal, 2001; Lecocq et al., 2007a, 2007b; Lecocq et al., 2009; Braun & Vergallo, 2010; Bouillon & Descamps, 2011). Second, CLIL appears to have no negative effects on the acquisition of content subjects¹⁷ (Blondin, 2003; Babault & Markey, 2016). Third, CLIL pupils' competences in the TL seem higher than those of pupils in the French-medium track, although without attaining nativelike levels (particularly in terms of productive skills, see Braun, De Man-De Vriendt & De Vriendt, 2006; Rasier, Bui, Jouniaux & Hiligsmann, 2014; Castel, Hiligsmann &

¹⁷ Note that the pupils' final assessment in Year 6 of content matter taught in the TL still happens solely in French (through Standard Achievement Tests: CEB for primary and CESS for secondary education in the CfB).

Leloup, 2015). Finally, CLIL appears to be positively associated with the pupils' openness toward the other language and its community (Blondin & Straeten, 2002; Mettwie & Lorette, 2014). However, we can also discern a number of shortcomings among these studies, mainly with regard to a lack of breadth and scope. All studies discussed involve a (very) small group of participants, mostly from one school (and/or one cohort). As mentioned above, schools are given much flexibility regarding the practical implementation of CLIL in the CfB. In order to make general statements about the outcome of CLIL, a larger sample comprising pupils from different schools is required, enabling the researcher to take into account the various contextual and pedagogical factors surrounding CLIL education. Furthermore, many of the studies discussed above solely focus on one specific – often linguistic – aspect of CLIL education, and an integrated approach (as strongly championed by CLIL scholars) is still lacking. Moreover, available studies usually did not include many covariates (e.g. SES, IQ, academic background, etc.) and have very low control on the equivalent of the samples in CLIL and NON-CLIL. Consequently, one cannot rule out the hypothesis that observed differences are due to unmeasured initial differences between students attending CLIL and NON-CLIL classrooms. Nevertheless, perhaps the most important limitation of many of the studies conducted so far is the use of a cross-sectional study design to compare CLIL and NON-CLIL pupils. Whereas such an approach obviously provides us with valuable information on potential differences between both groups at the time of testing, the cross-sectional design precludes any strong inferences on the impact of the CLIL experience as such. This means that we cannot verify whether these differences can be ascribed to the CLIL experience (as indirectly or directly suggested by some of the authors mentioned above) or perhaps to other factors. In order to state any claims on this issue, a longitudinal design would be more appropriate as only such designs allow for the analysis of “group progress, individual variation within groups and individual trajectories” (Meunier, 2015, p. 382). Finally, and more generally, the studies conducted until now (with the exception of Chohey-Pacquet, 2015) mainly focus on the outcomes of the CLIL programs. What happens within the classroom, for instance, remains largely unknown. This observation aligns with the misgivings voiced by, among others, De Graaff et al. (2007) and Bruton (2011), with respect to research on CLIL in general (as mentioned earlier).

In order to tackle these shortcomings of research on CLIL in French-speaking Belgium, researchers from the Université catholique de Louvain and the Université de Namur have set up a joint integrated research project (2014-2019). The design and the goals of this multidisciplinary and longitudinal study are briefly presented in the next section.

3.4. Research project “Assessing Content and Language Integrated Learning in French-speaking Belgium: an integrated approach” (2014-2019)¹⁸

On the basis of a large-scale longitudinal study, the research project aims to gain insight into the linguistic, cognitive, socio-affective and educational aspects of CLIL and to understand how the interplay between these perspectives may underlie TL acquisition processes. The data collected for the project involves French-speaking CLIL and NON-CLIL learners (control group) with Dutch or English as a TL, at different times during their final two years of primary and secondary school education. The dominant language studied in CLIL research is English (cf. Dalton-Puffer, 2011, p. 183). Interestingly, this project also includes Dutch in order to shed light on the actual TL acquisition process in terms of developmental patterns and errors typologies, and to assess the potential impact of attitudinal, motivational and cognitive factors on the learning of the two languages. At the methodological level, a longitudinal, contrastive and multidisciplinary perspective is adopted, including the following types of data collection:

- a detailed sociolinguistic and attitudinal questionnaire that has been administered to all the pupils, in order to shed light on socio-affective variables, learning trajectories as well as background and contextual information;
- a set of controlled tasks that tests both the linguistic and cognitive dimensions;
- (videotaped) classroom observations that provide materials for the linguistic and educational studies;
- focus-groups that support the socio-affective and educational facets of the project;
- parent and teacher questionnaires that provide background information on the home/classroom context.

The purpose is to build a complementary multi-layered database to be explored for the linguistic, cognitive and educative research lines in order to obtain a more comprehensive view on CLIL. A total of five groups of variables are measured: control, socio-affective, cognitive, instructional, and linguistic variables. Table 3.2. provides an overview of the research procedure.

Table 3.2. Overview procedure CLIL project (UCLouvain-UNamur)

YEAR 1 (2015-2016)	YEAR 2 (2016-2017)
5 th primary/secondary school	6 th primary/secondary school

¹⁸ The coordinated research project (ARC 14/19-061) is funded by the *Fédération Wallonie-Bruxelles (CfB)*.

T1	T2	T3	T4	T5
Student questionnaires			Student questionnaires	
Corpus data			Corpus data	
Computerized tasks			Computerized tasks	
Parent questionnaires	Teacher questionnaires	Focus groups		
		Classroom observations		Standardized achievement tests (SATs)

More specifically, the research project involves four major research lines (work packages). Two of these work packages deal with linguistic aspects: (a) the acquisition of nativelike lexical and morpho-syntactic use in the domain of intensification (Hendriks's doctoral thesis, 2019), and (b) the acquisition of nativelike use in the domain of phraseology (the present doctoral thesis). As previously mentioned in Chapter 2 (Sections 2.4.2. and 2.4.3.), one work package studies the various socio-affective conditions related to language learning in both the CLIL and NON-CLIL environments (De Smet's doctoral thesis, in preparation) and another investigates and compares the cognitive abilities of the CLIL and NON-CLIL pupils (Simonis' doctoral thesis, 2019). Investigating the interactions between the different research lines is also one of the goals of the research project.

For the first data collection, which took place between September and November 2015, data were collected among more than 900 pupils from thirteen primary and nine secondary schools (see Table 3.3.). The participating schools have contrasted profiles, notably in terms of location (all provinces are covered), socio-economic level, CLIL type (early vs. late), and education authority (official education and publicly subsidized schools).

Table 3.3. Initial sample CLIL project (UCLouvain-UNamur): number of pupils in year 5 of primary/secondary education ($n = 928$)

	CLIL Dutch (1)	CLIL English (2)	NON-CLIL English (3)	NON-CLIL Dutch (4)
Primary	175	103	106	71
Secondary	141	104	114	114

- (1) CLIL with Dutch as the target language
- (2) CLIL with English as the target language
- (3) NON-CLIL with specific foreign languages curricular option in English (4 hours/week)
- (4) NON-CLIL with specific foreign languages curricular option in Dutch (4 hours/week)

While this section has presented the overarching ARC research project on CLIL in French-speaking Belgium, the next section zooms in on the research line that is at the heart of the present thesis, that is the acquisition of phraseological language.

3.5. Research line on the acquisition of phraseological language in CLIL

This section demonstrates the relevance (Section 3.5.1.) and originality (Section 3.5.2.) of the present doctoral research within the framework of the ARC project described above.

3.5.1. Rationale

Research has shown that progress made in corpus-, psycho- and cognitive linguistics changed our understanding of language and impacted the field of S/FLA (Meunier & Granger 2008, Stengers, Boers, Housen, and Eyckmans, 2011). It is now a fact that language is to a large extent phraseological, viz. made of ‘chunks’, that lexis and grammar interface in producing form-meaning pairings and that spoken and written language develop through input and use. Since the early nineties, studies assessing language proficiency no longer focus exclusively on single words or isolated grammatical items; much literature has been devoted to the prefabricated nature of authentic language use. Phraseology has now developed as a very complex research field, characterized by an ever-expanding list of terms and a plethora of methodologies, and pleas for a more complementary approach to phraseology have been made.

Usage-based and emergentist approaches have demonstrated the importance of the quantity and quality of the input, as well as the key role played by word combinations in language acquisition. According to these approaches, language develops implicitly through accumulated linguistic experiences, while a number of frequency and psycholinguistic effects make the acquisition of word combinations more or less easy. The importance of explicit instruction is also acknowledged, especially when the input is too limited.

Previous empirical research has demonstrated that phraseological language is difficult to acquire by TL learners, and this for various reasons. Despite the growing body of research demonstrating the importance of phraseology for language acquisition, phraseological units do not seem to be the object of much explicit teaching and the limited quantity and quality of the

TL input provided in instructed settings does not seem to suffice for implicit learning to occur. In sum, the acquisition of phraseological units in instructed settings requires enough opportunities for both implicit and explicit learning. Investigating the acquisition of phraseological units in CLIL contexts (in comparison to NON-CLIL contexts) may help us better understand the effect of a larger quantity of TL input (more hours of exposure) and of a more authentic type of TL input (focus on meaning and form through the integration of content and language learning).

3.5.2. Originality of the research design

The originality of the research design lies in different aspects. First, the study examines the phraseological competence of CLIL and NON-CLIL learners, an area that is clearly underresearched. In fact, studies assessing TL proficiency in CLIL contexts have paid much attention to lexicon (Ruiz de Zarobe, 2011) – probably because the content-based approach typical of CLIL provides more opportunities to learn target vocabulary in meaningful situations – but too rarely went beyond the word level (see also Llinares, Morton, and Whittaker, 2012), therefore neglecting the formulaic/phraseological nature of language.

Second, as the practical implementation of CLIL may vary from one school to another, the present study consists of a large sample of learners coming from nine secondary schools with contrasted profiles (see Section 3.4.). The very few existing studies that investigated the phraseological competence of CLIL learners involved participants from only one or two schools.

Third, the present research investigates two different target languages: English (which is the most studied language in CLIL research) and Dutch (which is much less represented in CLIL research but is of particular relevance to the specific Belgian context).

Fourth, the data collected are extensive and varied, which allows for different types of analyses: longitudinal and cross-sectional, corpus-based and experimental, quantitative and qualitative. This enables us to take into account the multidimensional aspect of phraseological competence by investigating the learners' productive use in terms of frequency, variety and accuracy, as well as their receptive knowledge. In addition, the development of their language competence over time can be examined thanks to the longitudinal data, which appears to be missing in CLIL research in general (see Section 3.3.).

Fifth, the analyses are not limited to one or two types of word combinations, as has usually been the case, but involve a large array of phraseological units. Even though choices had to be made in terms of terminology and methodology, the approach adopted in this study is meant to be as encompassing as possible by including the benefits of both the traditional and the frequency-based approach.

As we have seen, the analysis of the phraseological competence of learners of English or Dutch in CLIL and NON-CLIL settings is interesting in its own right. Nevertheless, another innovative aspect of this study is that it is framed within a research project that aims to investigate the interplay between linguistic, educational, cognitive, and socio-affective aspects (see Section 3.4.). The links between educational factors and the acquisition of phraseological units would arguably be strong. As discussed earlier, input frequency is not the only key determinant to the successful acquisition of word combinations. Explicit forms of instruction also seem to be necessary, and a number of teaching practices such as focus on form activities (awareness raising, focus on salience, etc.), for example, would undoubtedly reinforce the benefits of rich input in CLIL classes. As for the connection between phraseological competence and cognitive factors, research has shown that learners are sensitive to frequency effects; a particular form of a word combination is more likely to be represented in the learners' mental lexicon if it has been frequently encountered. This means that the additional hours of TL exposure provided in CLIL may reinforce the cognitive entrenchment of phraseological units in the learners' mind. Finding direct links between socio-affective factors and the acquisition of phraseological units may be more difficult. Nevertheless, some pupils may not be motivated in learning multiword units because they think learning single words is already hard enough (as reported by Coxhead, in preparation) or because they may give preponderance to communication over accuracy (although it could be argued that effective communication and accuracy often go hand in hand) (Granger & Meunier, 2008). The more authentic learning environment of CLIL may increase motivation and make the learning of word combinations appear more meaningful, and therefore easier.

3.6. Conclusion

This chapter has shown that the popularity of CLIL programs has only grown since its first implementation in French-speaking Belgium. As for the studies carried out on CLIL in French-speaking Belgium, they are relatively scarce and mostly focus on the linguistic outcomes of the approach. Moreover, similarly to international research on CLIL, they usually fail to assess the real impact that CLIL has on TL acquisition. The ARC research project aims to fill this gap by

taking an integrated and multidisciplinary approach. More generally, the project intends to contribute both theoretically and empirically to the public debate in French-speaking Belgium and the ongoing international scientific discussions on multilingualism in general and CLIL in particular.

The present research, which forms an integral part of the project, focuses on the CLIL and NON-CLIL secondary school learners' productive and receptive knowledge of phraseological units in the target languages English and Dutch. It will contribute to the body of research on CLIL, which has so far focused on the lexical competence of learners without taking into account its phraseological aspect. In addition, the present research may give more insights into the impact of TL input and an implicit form of learning.

PART 2: ACQUIRING PHRASEOLOGICAL LANGUAGE: RESEARCH HYPOTHESES, METHODOLOGY AND RESULTS

Part 1 was concerned with the theoretical background of the present doctoral dissertation. Part 2 is devoted to the investigation of our research, that is the acquisition of English and Dutch phraseological language by CLIL and NON-CLIL learners in French-speaking Belgium. Chapter 4 is a methodological chapter, which presents a pilot study into the learners' global written competence in their L1 and TL English/Dutch, the research hypotheses, the participants, and the different tests, tasks, and data that were used for the analyses. A number of studies have been carried out in order to test the research hypotheses and answer our research questions. These studies are thoroughly described in Chapters 5, 6, and 7. Chapter 5 deals with the learners' productive phraseological competence in the target languages English or Dutch, as well as its development over time using a longitudinal corpus. Chapter 6 is devoted to the learners' receptive knowledge of a specific type of phraseological unit, namely verb-noun collocations. Finally, Chapter 7 is a follow-up study to Chapter 6 in which the data from the receptive collocational test are integrated in a regression model, together with overall vocabulary knowledge measures and TL input-related measures, in order to isolate the effect of CLIL on the learners' receptive knowledge of verb-noun collocations.

CHAPTER 4: Methodology

4.1. Introduction

This methodological part first presents the results of a pilot study into the learners' global written proficiency (Section 4.2.), which subsequently serve to refine the thesis' research hypotheses (Section 4.3.) based on the theoretical insights introduced in Part 1. It then describes the participants (Section 4.4.), as well as the data collected (Section 4.5.). It also provides an extensive description of the procedure followed for the extraction and classification of the learners' phraseological productions (Section 4.6.) and briefly describes the statistical analyses carried out (Section 4.7.). Finally, it outlines the thesis' research design (Section 4.8.), before wrapping up in Section 4.9.

4.2. Pilot study: global written proficiency in L1 French and TL English/Dutch

This pilot study was a first exploratory analysis of the learners' written productions¹⁹. One of the objectives was to get insight into the learners' global written proficiency in their L1 (French). Based on previous findings showing no negative effect of CLIL on L1 performance (see Chapter 3, Section 3.2.), we expected CLIL learners to perform equally well in French as their NON-CLIL peers. The other objective was to examine the influence of CLIL on the learners' global written proficiency in their target languages (English or Dutch). We believe that starting with the examination of CLIL and NON-CLIL learners' TL global proficiency level before zooming in on their TL phraseological competence will help us refine this thesis' main research hypotheses (see Section 4.3. below). As TL written proficiency and phraseological complexity appear to be related (e.g. Paquot, 2019), if differences are found between CLIL and NON-CLIL learners at the overall proficiency level, we may hypothesize similar differences at the phraseological level. We actually expected to find a higher level of written proficiency in CLIL as it provides greater exposure to the TL and appears to be beneficial in terms of overall general written competence (see Chapter 1, Section 1.9.3.).

The written competence of CLIL and NON-CLIL learners was compared using corpora of semi-spontaneous written productions and a number of selected measures as proxies for general written proficiency. For the English texts, we used Coh-Metrix, a computational tool “which

¹⁹ The results were the object of an online publication: 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, *Travaux du Cercle Belge de Linguistique*, 11(1), 1-25.

analyzes texts on over 200 measures of cohesion, language and readability” (Graesser, McNamara, Louwerse & Cai, 2004, p. 93). For the Dutch texts, we used a similar tool, T-Scan, which stands for Software voor Complexiteits-Analyse van het Nederlands (Pander Maat et al., 2014). For the French texts, as software tools are scarce, we used Wordsmith Tools (Scott, 2012) for basic counts and then used more complex measures included in a software tool for French developed by François (2011). Among the measures offered by each tool, we selected both lexical and syntactic measures that were similar for English, Dutch and French and that were reported as good indicators of overall language proficiency (see Bulon et al., 2017 for a detailed discussion of the measures). The selected measures were the following: total number of sentences and total number of words per text, sentence length (in number of words), word length (in letters and syllables), TTR (Type-Token Ratio) and MTLD²⁰ (Measure of Textual Lexical Diversity). For English and Dutch, TTRs were computed for content words and for all words. For French, TTRs were computed for inflected forms and lemmas.

The participants in this pilot study are the same as the ones described in Section 4.4. below. All the learners were in the fifth year of their secondary school education at the time of the data collection. The corpus of semi-spontaneous written productions was collected using a writing task that took place between October 19th and November 9th 2015 in computer rooms in Louvain-la-Neuve. The task, which is described in detail in Section 4.5.1., consisted in writing an e-mail on two possible topics (either their last holidays or a party they attended). They wrote a first e-mail in the TL (Dutch or English) in the morning and a second one in their mother tongue (French) in the afternoon. As Table 4.1. shows, we were able to collect a total of 843 productions.

Table 4.1. Number of texts collected per condition (language – type of education)

French (L1)	Dutch CLIL	Dutch NON-CLIL	English CLIL	English NON-CLIL
431	132	100	90	90

The results are presented in Table 4.2., which shows the median scores per complexity measure for the texts written by the CLIL and the NON-CLIL pupils (see Appendix 6 for the statistics of the comparison between the CLIL and the NON-CLIL scores).

²⁰ Koizumi (2012) found that MTLD was least affected by text length compared to TTR and other recent indices (Guiraud index and D), when used with texts of at least 100 tokens. That is why we chose to include it along with TTR for English and Dutch (we were unable to compute this score for the French texts).

Table 4.2. Median scores per measure for each language group (green = CLIL significantly higher score, orange = CLIL significantly lower score)

		DUTCH		ENGLISH		FRENCH	
		NON-CLIL	CLIL	NON-CLIL	CLIL	NON-CLIL	CLIL
Sentences per text	Median	21,0	25,5	22,5	22,0	18,0	19,0
	Variance	50,72	41,86	64,23	54,32	59,03	53,65
Words per text	Median	194,5	274,0	267,5	328,0	288,0	296
	Variance	3571,47	3339,65	6930,62	4071,49	6836,66	5248,65
Words per sentence	Median	9,0	10,7	11,0	14,5	14,9	14,8
	Variance	6,14	9,77	29,71	35,20	51,66	20,64
Syllables/ morphs per word	Median	1,23	1,27	1,30	1,28	-	-
	Variance	0,004	0,002	0,002	0,002	-	-
Letters per word	Median	4,27	4,35	3,86	3,84	4,03	4,06
	Variance	0,04	0,03	0,29	0,02	0,04	0,04
TTR content words	Median	0,77	0,77	0,70	0,71	0,68	0,68
	Variance	0,006	0,004	0,12	0,004	0,001	0,002
TTR all words	Median	0,54	0,51	0,49	0,49	0,76	0,77
	Variance	0,005	0,003	0,01	0,003	0,001	0,001
MTLD	Median	58,5	73,0	60,2	70,4	-	-
	Variance	323,82	258,74	205,67	269,42	-	-

The comparison of the French texts showed no significant differences between the learners in CLIL and in NON-CLIL. This finding seems to indicate that, as expected, CLIL education does

not have any (positive or negative) influence on the complexity of the learners' L1. The comparison of the English and Dutch texts showed different results according to the TL. We can see that most significant differences are found in the texts written in Dutch: the texts written by CLIL learners were significantly more complex than those written by NON-CLIL learners, except for TTR content words and TTR all words. In contrast, only half of the scores diverge significantly in English: the texts written by CLIL learners were more complex regarding text length, sentence length and MTLTD, and the NON-CLIL learners used significantly more syllables per word.

Our results show more differences between the groups in Dutch than in English. One possible explanation for this may be that NON-CLIL learners of English are more proficient in the TL than NON-CLIL learners of Dutch are. The greater availability of English outside school and/or its greater attractiveness as a TL may also play a role (Berns et al., 2007; De Le Vingne, 2015; De Smet et al., submitted), so that the benefits of CLIL in English are less evident compared to CLIL in Dutch. Another potential explanation is that NON-CLIL learners of Dutch would have more difficulty acquiring Dutch than NON-CLIL learners of English acquiring English, which would be due to the fact that French shares more (syntactic, morphological and lexical) properties with English than with Dutch (Laufer & Paribakht, 1998; Fox, 2002; Forlot & Beaucamp, 2008; Pasquarella, Chen, Lam, Luo & Ramirez, 2011; Pierce, 2012).

This exploratory analysis has shown differences in the global written proficiency of CLIL and NON-CLIL learners, with a higher level of competence in the CLIL groups. These findings provide valuable insights for the formulation of our research hypotheses. Also, the fact that the differences appeared to be more substantial in the TL Dutch will be considered while investigating the learners' linguistic output at the more specific level of phraseology.

4.3. Research hypotheses

The aim of this research is to complement the understanding of TL acquisition by using both corpus-based and experimental methods and tools to assess whether the output of CLIL learners is more target-like than that of NON-CLIL learners. More specifically, it investigates whether input-rich CLIL settings speed up the acquisition of overall phraseology.

In light of previous findings pointing out the advantages of CLIL for SLA, given that CLIL settings offer more target like input and input-rich environments than NON-CLIL settings and based on the findings of the pilot study, we formulate the following main research hypotheses:

1. The written TL output of CLIL learners will be more phraseological (and target-like) than that of NON-CLIL learners, i.e. they will produce more phraseological units and display higher phraseological accuracy;
2. Both the CLIL and NON-CLIL learners will show improvement in their phraseological language after a period of two consecutive school years, with CLIL learners displaying greater progress;
3. The receptive phraseological knowledge of CLIL learners will be greater than that of NON-CLIL learners.

These hypotheses apply to both English and Dutch as target languages and will be addressed using various tests and tasks that are presented in Section 4.4. As explained in the introduction, these main hypotheses will be broken down into specific hypotheses in the empirical chapters. The influence of other variables (frequency of occurrence, general vocabulary knowledge, etc.) will also be taken into account to measure their effect on the acquisition of a particular type of phraseological units.

It seemed important to cast a multifocal light on the learners' phraseological competence. In fact, in this thesis, we not only look at the learners' competence in terms of frequency of production (i.e. the number of phraseological units produced), but also in terms of variety (i.e. the number of different phraseological units produced) and accuracy (i.e. the quality of the phraseological units produced). It would be slightly misleading to speak of phraseological competence as a monolithic concept as learners may produce a large number of phraseological units, but always use the same. Moreover, more does not necessarily better in the sense that accuracy does not always correlate with an increase in production, as reported in a number of studies mentioned in Chapter 1, Section 1.9. According to Huang (2015), quality may even be more important than quantity when investigating the phraseological competence of TL learners. Besides, looking at these different aspects also allows for a more comprehensive analysis of the linguistic features that differentiate between CLIL and NON-CLIL learners.

4.4. Participants

The data collected for the present thesis involved 412²¹ French-speaking CLIL and NON-CLIL learners (control group) with Dutch or English as a TL, at different times during their final two years of secondary school education. They came from nine different schools in Wallonia

²¹ A total of 473 secondary school pupils enrolled in the study, but some data were missing due to technical problems.

(French-speaking Belgium) with different profiles (as already mentioned in Chapter 3, Section 3.4.) and all of them followed the general education track. The learners of English had chosen Dutch as their ‘second’ TL learned at school and the learners of Dutch had chosen English as their ‘second’ TL learned at school. The schools provided CLIL in Dutch and/or English, alongside traditional French-medium instruction with foreign language classes. The content subjects that were taught in the TL to CLIL learners were history and geography. The proportion of CLIL lessons provided to the learners in CLIL programs went from 15% to 30% according to the school.

Table 4.3. gives the number of participants per target-language learning condition.

Table 4.3. Number of participants per target-language learning condition

	English	Dutch	Total	Percentage
CLIL	90	132	222	53.88
NON-CLIL	90	100	190	46.12

At the time of the first data collection (fall 2015), the participants’ ages ranged from 15 to 18 and their mean age was 16.5. 223 (54.12%) of the participants were female, while 189 (45.88%) were male.

4.5. Tests, tasks, data

The data were collected between October 2015 and May 2017. Three different waves of data collections were planned as follows:

- October 19th - November 9th 2015: collection of written productions and background questionnaires at the UCLouvain;
- February/March 2017: collection of spoken productions in the schools;
- April 10th - May 15th 2017: collection of written productions and administration of a multiple-choice test at the UCLouvain.

Most of the data collected were corpus-based (written and spoken productions). The use of corpora in the study of second/foreign language learners’ interlanguage has “increased its importance in most fields of linguistics and [...] is currently the most frequently used method employed in the study of phraseology” (Gries, 2008, p. 15), as highlighted in Chapter 1, Section 1.5.1. In fact, learner corpora allow for a better understanding of how second/foreign languages

are acquired. As already mentioned in Chapter 1, Section 1.5.1.2., no comparison with native speaker data is made.

The tasks and tests that were designed in order to gather the data are described in the following subsections.

4.5.1. Written productions

A writing task was developed in order to compile a corpus of the learners' written productions. The task consisted in writing an informal e-mail to a friend about everyday topics (holidays or a party) in their TL (Dutch or English)²². As text type has been reported to have a great influence on lexical diversity (see for instance Yu, 2009), we controlled for this by making all the pupils write on the two same topics (if a pupil had written about a party in the morning, s/he was given the other topic – holidays – in the afternoon, and vice versa). The text topics were randomly assigned and the same oral and written instructions were provided to the learners. The written instructions were purposefully quite short and did not include many phrases relating to the topics suggested; the probability of pupils reusing a large number of word combinations from the rubric of the task is therefore minimal (see Appendix 1).

The goal was to make the learners produce a TL output that was as authentic and spontaneous as possible. According to Granger (2008), free compositions allow to gather learner production data that have a high degree of naturalness. The learners were to a large extent free to choose their own wording, but the task was not fully open-ended since pictures were provided for inspiration. The task can therefore be considered as semi-spontaneous.

The task was carried out in computer rooms at the UCLouvain. It was timed – learners had 25 minutes to write their e-mail. There was no word limit but the learners had to produce at least 15 lines. The spelling checkers had been turned off and the learners had no access to any reference tools. The task was presented in the online software Lime Survey²³ and the learners wrote their e-mail directly in a text window of the software.

The written learner corpus is longitudinal; the same writing task was performed at two different points in time (T1 and T2) by the same learners. As is always the case in longitudinal studies, some of the initial participants did not take part in the second round of tests in 2017, so that the

²² The learners were also asked to perform the same task in their L1 (French) (see pilot study presented above), however the analysis of the learners' phraseological productions in their mother tongue is beyond the scope of this doctoral dissertation.

²³ Lime Survey <http://www.limesurvey.org/about-limesurvey/team-contributors> (team: Schmitz, Gac, Härstedt, Flür, Kondziella, Chirițoiu).

sample used for the longitudinal study is somewhat smaller than the one described in the previous section.

Table 4.4. below gives the number of texts, number of words and the average text length for each sub-corpus.

Table 4.4. Description of the written sub-corpora

	TL English		TL English		TL Dutch		TL Dutch	
	CLIL		NON-CLIL		CLIL		NON-CLIL	
	T1	T2	T1	T2	T1	T2	T1	T2
Nb of texts	78	78	66	66	123	123	72	72
Nb of words	25,537	24,447	17,341	17,284	34,786	33,495	14,031	15,535
Mean text length	327	313	263	262	283	272	194	216
SD	65.39	71.43	82.46	54.80	58.25	50.28	62.96	47.53

See Section 4.6. in this chapter for more information on the procedure of extraction and classification of the phraseological units and phraseological errors.

4.5.2. Oral interaction

An oral task was designed in order to compile a corpus of the learners' spoken productions. The learners had to perform a dialogue exercise in pair which consisted in solving a 'problem': they each received an invitation to a party or won a free trip to an exotic country and had to decide on which party to attend or which destination to choose. This technique is usually used to elicit spoken data (Wichmann, 2008). Moreover, conversations on familiar topics have been shown to bring out a much larger proportion of phraseological language than story retelling tasks (Cordier, 2013).

The task was carried out in three different secondary schools selected from the initial sample described in Section 4.4., so that the number of participants was much more limited. Again, we tried to select three schools with contrasting profiles. The task was performed during school hours, so we had agreed with the school principals and the teachers that 5-6 groups of two pupils each would leave the class for about 15 minutes. The learners had no preparation time so that their productions were as spontaneous as possible. The interaction was timed (5 minutes) and each pupil's production was recorded using a headset. The instructions were given in the

learners' mother tongue (French) and some pictures were included for inspiration (see Appendix 2).

The topics were the same as for the written productions (the holidays or a party) so that a comparative analysis of their phraseological productions in the two different communication modes could be made. Dialogues were chosen over monologues for the same reason; the learners had an interlocutor in both the written and the oral task.

The number of oral productions, the number of words and the average length of the productions are shown in Table 4.5.

Table 4.5. Description of the spoken sub-corpora

	TL English CLIL	TL English NON-CLIL	TL Dutch CLIL	TL Dutch NON-CLIL
Nb of productions	17	11	13	20
Nb of words	4,173	2,153	3,854	3,644
Mean length of production	245	196	296	182
SD	108.02	79.08	65.78	82.84

The same procedure of extraction and classification was used as for the written productions (see Section 4.6. below). Quantitative and qualitative analyses of the spoken data are still being carried out, which explains why they are not included in the present thesis. A brief description of the study and some preliminary results are however provided in the concluding chapter as they may prompt future research on the topic.

4.5.3. Multiple-choice test (receptive phraseological competence)

A receptive language test in the form of a multiple-choice exercise was created in order to measure the learners' receptive phraseological knowledge. The exercise only tested one type of phraseological units, namely verb-noun collocations. The participants who performed this test are the ones from the initial sample (see Section 4.4.) who did not drop out (as the data were collected at T2), plus a very small number of learners who joined in at a later stage.

The multiple-choice exercise differed according to the TL. This is explained by the fact that already existing receptive tests have recently been developed in order to test the TL English, while no such material was found for the TL Dutch. The difference mainly lies in the presentation of the tested items (the collocations were contextualized in sentences for Dutch, while no context was provided for English) and in the degree of validity of the tests (the one

used for English was the object of a research thesis and proved to yield reliable results, while the one designed for Dutch was pre-tested once).

The test consisted in finding the correct verb collocate among three options available. One of these options formed a conventionalized TL collocation, while the other two were ‘pseudo-collocations’ or ‘distractors’. An example for each TL is given below (see Appendix 3 for the complete tests):

- a) a. draw a conclusion
b. pull a conclusion
c. make a conclusion

- b) Wat ik voor jou voel, is moeilijk onder woorden te
a. brengen
b. zeggen
c. maken

The exercise was presented to the learners using the software Lime Survey. More information on the exercise and the selected verb-noun collocations is provided in Chapter 6, Section 6.3.2.

4.5.4. Measures of overall productive and receptive vocabulary knowledge in the target language

The learners’ overall TL productive vocabulary knowledge was calculated on the basis of their written productions (see Section 4.5.1. in this chapter) using the Measure of Textual Lexical Diversity (MTLD). This measure was first computed in the framework of the pilot study looking at the learners’ global written competence (see Section 4.2. in this chapter) using Coh-Metrix (Graesser, McNamara, Louwerse & Cai, 2004) for English and T-Scan (Pander Maat et al., 2014) for Dutch. The measure was then used as a predictor variable of receptive collocational knowledge in Chapter 7.

The learners’ overall TL receptive vocabulary knowledge was measured using a standardized computerized test: the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-IV) (Dunn & Dunn, 2007) for English, and the Peabody Picture Vocabulary test-III-NL (PPVT-III-NL) (Dunn, Dunn & Schlichting, 2005) for Dutch. These tests consisted in identifying the picture corresponding to a pronounced word among four different options. The score that the learners

achieved – which corresponds to the number of TL words correctly identified – was also used as a predictor variable of the learners’ receptive collocational knowledge in Chapter 7.

4.5.5. Background variables

A large number of learner-related and TL-specific variables were collected and some of them were taken into account in our research. This is the case of the learners’ socio-economic status (SES). This was measured using the mother’s education level²⁴, an information that was collected via questionnaires distributed to the parents. Input-related variables such as the number of years the learner has spent learning the TL at school or the learner’s current informal contact with the TL outside school were also collected via background questionnaires administered to the learners. Another background variable considered in this study is the learners’ non-verbal intelligence (IQ), which was measured using a standardized computerized task (the Raven test). More information about these variables can be found in the following empirical chapters.

4.6. Extraction and classification of phraseological units and phraseological errors

The extraction method used in this dissertation is based on frequency, but is not a purely frequency-based approach in the sense that lexical bundles such as “going to be a”, “percent of the”, “de vader van” (‘the father of’) or “de dag dat” (‘the day that’) would not be taken into consideration. Although these combinations may be frequent, they might be less interesting from a pedagogical perspective and are therefore less central to our study. In fact, we were interested in covering a wide range of combinations with different degrees of syntactic fixedness and semantic compositionality (i.e. from idioms to free combinations), while focusing on those that carry a certain function or meaning.

Concretely, we used Wordsmith Tools (Scott, 2012) to extract automatically all word clusters (word combinations) from the texts. The cluster size was set between 2 and 4 words, and the minimum frequency of occurrence for each cluster was 2. We did not make use of association measures; these are well suited for the extraction of bigrams such as collocations (Heid, 2008), but less so for clusters of three words or more.

Then, we manually checked and pruned the list of clusters and organized them per category of phraseological unit. Examples of English word combinations that were selected for further

²⁴ This may seem discriminating but the mother’s level of education is very frequently used as a proxy for SES because it has a low rate of missing responses and is highly correlated with father’s education (Entwistle & Astone, 1994).

analysis are *nice people, pretty good, take a shower, keep in touch, birthday cake, good at, dance with, clean up, in terms of, as soon as, by the way, see you, it's a shame*, etc. Examples for Dutch include *slecht weer* ('bad weather'), *heel moe* ('very tired'), *plezier hebben* ('have fun'), *uit de bol gaan* ('go crazy'), *armband* ('bracelet'), *verliefd op* ('in love with'), *genieten van* ('enjoy'), *opruimen* ('clean up'), *in het midden van* ('in the middle of'), *zelfs als* ('even if'), *ten eerste* ('firstly'), *tot binnenkort* ('see you soon'), *jammer genoeg* ('unfortunately'), etc.

The classification was made on the basis of Granger and Paquot's (2008) taxonomy (see Table 4.6. below), which seemed well suited to our combined approach to phraseology.

Table 4.6. Taxonomy used for the classification of phraseological units

REFERENTIAL PHRASEMES		
Lexical collocations	Adjective + noun	Ex: <i>heavy rain; mooi weer</i> ('good weather')
	Adverb + adjective	Ex: <i>pretty sure; helemaal zat</i> ('completely drunk')
	Verb + noun	Ex: <i>take place; het vliegtuig nemen</i> ('take the plane')
	Adverb + verb	Ex: <i>really appreciate; lekker eten</i> ('eat well, feast')
Idioms	Ex: <i>keep in touch; uit de bol gaan</i> ('go crazy')	
Irreversible bi- and trinomials	Ex: <i>more or less; alles en nog wat</i> ('this and that')	
Similes	Ex: <i>cry like a baby; slapen als een otter</i> ('sleep like a log')	
Compounds	Ex: <i>birthday cake, seafood, make-up; verjaardagsfeest</i> ('birthday party'), <i>luchthaven</i> ('airport')	
Grammatical collocations	Adjective + preposition	Ex: <i>afraid of; gek van</i> ('keen on')
	Preposition + noun	Ex: <i>by car; op school</i> ('at school')
	Adjective + to + infinitive	Ex: <i>happy to; verplicht te</i> ('obliged to')

	Noun + preposition	Ex: <i>an invitation to</i> ; <i>soort van</i> ('sort of')
	Verb + to + infinitive	Ex: <i>try to</i> ; <i>durven te</i> ('to dare to')
	Verb + preposition	Ex: <i>depend on</i> ; <i>wachten op</i> ('to wait for')
Phrasal verbs	Ex: <i>wake up</i> ; <i>aanbieden</i> ('to offer')	
TEXTUAL PHRASEMES		
Complex prepositions	Ex: <i>in addition to</i> , <i>in terms of</i> ; <i>ten opzichte van</i> ('in relation to'), <i>in plaats van</i> ('instead of')	
Complex conjunctions	Ex: <i>even though</i> , <i>as soon as</i> ; <i>zelfs als</i> ('even if'), <i>ondanks het feit dat</i> ('despite the fact that')	
Linking adverbials	Ex: <i>in other words</i> , <i>by the way</i> ; <i>in het algemeen</i> ('in general'), <i>ten eerste</i> ('firstly')	
Textual sentence stems	Ex: <i>the problem is that</i> ; <i>dat is de reden waarom</i> ('that is the reason why')	
COMMUNICATIVE PHRASEMES		
Speech act formulae	Ex: <i>happy birthday!</i> , <i>how do you do?</i> ; <i>graag gedaan</i> ('my pleasure'), <i>tot ziens</i> ('see you later')	
Attitudinal formulae	Ex: <i>in fact</i> , <i>I think that</i> ; <i>zonder twijfel</i> ('undoubtedly'), <i>jammer genoeg</i> ('unfortunately')	
Commonplaces	Ex: <i>we only live once</i> ; <i>wie niet sterk is, moet slim zijn</i> ('necessity is the mother of invention' ²⁵)	
Slogans	Ex: <i>a bird in the hand is worth two in the bush</i> ; <i>hoge bomen vangen veel wind</i> ('the bigger they are, the harder they fall')	
Proverbs	Ex: <i>make love, not war</i> ; <i>wie zijn billen verbrandt, moet op de blaren zitten</i> ('who burns his butt, must sit on the blisters')	

The phraseological units were first classified in three major categories: referential phrasemes, textual phrasemes, and communicative phrasemes. The first category includes phrasemes that

²⁵ Traduction provided by <http://www.mijnwoordenboek.nl>

are “used to convey a content message: they refer to objects, phenomena or real-life facts” (Granger & Paquot, 2008, p. 42). The second category contains phrasemes that are “typically used to structure and organize the content (i.e. referential information) of a text or any type of discourse” (ibid.). The third category consists of phrasemes that are “used to express feelings or beliefs towards a propositional content or to explicitly address interlocutors, either to focus their attention, include them as discourse participants or influence them” (ibid.). The phraseological units were then assigned to different sub-categories (see Granger & Paquot, 2008, pp. 43-44 for a detailed description of each sub-category). The taxonomy has been extended regarding *lexical collocations* and *grammatical collocations* in order to detail the various possible combinations within these sub-categories. The examples provided in Table 4.6. come from the learners’ texts, except for the sub-categories *similes*, *commonplaces* and *slogans* for Dutch and *commonplaces*, *slogans* and *proverbs* for English (no occurrence was found in the learners’ productions for these categories of phraseological units).

The categorization of the phraseological units was challenging for some cases. Distinguishing between lexical collocations and grammatical collocations, for example, was quite straightforward. In fact, lexical collocations combine two lexical words, while grammatical collocations are combinations of a lexical and a grammatical word (typically verb/noun/adjective + preposition). However, making the distinction between English prepositional verbs (verb + preposition) and phrasal verbs, for example, was more difficult. At first glance, both types of word combinations pretty much look alike, even though in the case of prepositional verbs, the second component is a preposition and in the case of phrasal verbs, it is a prepositional adverb. The word combination *go out* is a prime example, especially as it can be used both as a prepositional verb or a phrasal verb according to the context. Two examples are given below (1-2).

(1) *We have met a few friends and all evening, we **went out**. We came back at hotel at average 4 o'clock.* (NON-CLIL, T1)

(2) *Everyday were with minimum 20 degrees celsius, we can **go out** at 5am in t-shirt by the way, and always sun.* (NON-CLIL, T2)

In example (1), the combination *went out* refers to the action of leaving a place to do something enjoyable (e.g. go to a party, to the cinema, etc.) and is used as a phrasal verb. In example (2), the combination *go out* refers to the action of departing from a room, a house, etc. and is used as a prepositional verb. While the phrasal verb in example (1) does not require an object, the

prepositional verb in example (2) implicitly refers to an object (a hotel or a house). This is one way of distinguishing between the two types of phraseological units. However, some phrasal verbs may require an object – this is the case of *pick up* (see example (3) below).

(3) *My parents went by car and they **picked us up** at the airport in Portugal.* (CLIL, T1)

We categorized it as a phrasal verb because the object can be placed before or after the prepositional adverb. This is not the case with prepositional verbs; the object has to come right after the preposition (see example (4) below).

(4) *I don't have much money right now and I can't afford to **pay for** the funerary services.*
(CLIL, T2)

Reference tools such as the *Oxford Collocations Dictionary for students of English* (2009), the *Combinatiewoordenboek van Nederlandse substantieven (zelfstandige naamwoorden) met hun vaste verba (werkwoorden)* (2003), the *British National Corpus* (BNC) and the *Corpus Hedendaags Nederlands* (CHN) were used to assess the degree of acceptability of the phraseological combinations used by the learners, i.e. to check whether the combinations produced were frequent, and somehow conventionalized English/Dutch combinations. A word combination that occurred at least five times in different texts in the reference corpora was considered as an appropriate combination. As for the classification into the various categories, we referred to some already existing lists for English such as the one compiled by Klégr (1997) for complex prepositions and the one made by Liu (2008) for linking adverbials. Appendix 4 provides a number of texts produced by the learners as illustration. The phraseological units identified are highlighted in yellow and the category assigned to the phraseological unit is found within square brackets.

To sum up, the selected combinations were those that appeared at least twice in the corpus, that corresponded to one of the (sub)-categories included in the taxonomy, and that (frequently) occurred in the reference tools consulted. No specific measure in relation to distribution and dispersion was taken during the extraction process. The participants' productions were quite short and we did not expect each individual learner to produce instances of every category of phrasemes – we think that even low frequency phenomena, i.e. word combinations that occur sporadically in a small number of texts are worth investigating. We examined the distribution of the phraseological units produced across individual learners by counting how many learners used each type of word combination for English and Dutch (see Appendix 5) and we can say that the most frequent categories of phrasemes are rather well spread out across the pupils (in

CLIL and NON-CLIL). Similarly, we did not look into the distribution of phrasemes according to their length (2-word, 3-word, 4-word,... units) because we were dealing with a large variety of word combinations. This would have probably been more relevant if the study focused on one type of word combination (e.g. connectives).

A last step involved the classification of the phraseological errors on the basis of Thewissen's (2008) and Hong, Rahim, Hua, and Salehuddin's (2001) work (see Table 4.7. below).

Table 4.7. Taxonomy used for the classification of phraseological errors²⁶

Category of error	Example
Wrong choice of words (the word exists)	*do pictures => take zonnig *tijd => weer ('sunny weather')
Wrong choice of word (the word does not exist)	*repass an exam => retake heel *plaisant => plezant ('very pleasant')
Creation of a deviant verb	*gived a present => gave mensen *ontgemoet => ontmoet ('meet people')
Misuse of item from other lexical categories	started to *cooking => cook ik hoop dat het *wel gaat => goed ('I hope that you are well')
Wrong word order	put *order in sth => put sth in order ik hoop dat het *gaat goed met je ('I hope you're doing okay')
Wrong choice of preposition	fall in love *for => with gaan *in => naar ('to go to')
Preposition missing	next *[to] kennismaken *[met] ('to get acquainted with')
Superfluous preposition	attend *to a party

²⁶ The English translations provided correspond to the correct Dutch phraseological units.

	<i>*naar</i> muziek beluisteren ('listen to music')
Article missing	have <i>*[a]</i> good time tot <i>*[de]</i> volgende keer ('see you next time')
Superfluous article	all <i>*the</i> night naar <i>*de</i> zee gaan ('to go to the sea')
Noun used in the plural instead of singular	go on <i>*holidays</i> <i>*ruzies</i> maken ('to quarrel')
Noun used in the singular instead of plural	many <i>*time[s]</i> vier <i>*ster[ren]</i> hotel ('four-star hotel')
Missing separable particle	/ stelde <i>*[voor]</i> ('to propose', 'to present')
Superfluous separable particle	/ uit het vliegtuig <i>*uit</i> stappen ('get off the plane')
Missing items	as <i>*[you]</i> know jammer <i>*[genoeg]</i> (('unfortunately'))
Superfluous items	made <i>*me</i> a friend zoals je <i>*het</i> weet ('as you know')
One word instead of two separate words (compounds)	<i>*swimmingpool</i> => swimming pool <i>*lolgemaakt</i> => lol gemaakt ('had fun')
Two separate words instead of one word (compounds)	<i>*grand father</i> => grandfather <i>*zomer vakantie</i> => zomervakantie ('summer holidays')
Superfluous dash (compounds)	<i>*living-room</i> => living room

	*groot-ouders => grootouders (‘grandparents’)
Missing dash (compounds)	*break[-]up (noun) *Middellandse[-]zeeklimaat (‘Mediterranean sea climate’)
Wrong word order (compounds)	* pool hotel => hotel pool * sport water => watersport (‘watersports’)
Superfluous items (compounds)	*guest’s room => guest room fontein * met chocolade => chocoladefontein (‘chocolate fountain’)
Missing items (compounds)	/ *verjaardag[s]/feestje (‘birthday party’)

This taxonomy has been created as the errors were spotted and aims to encompass as many phraseological errors as possible. Note that the last seven categories specifically tackle errors related to the sub-category *compounds*. Some phraseological units contain more than one error and may therefore be found in different categories. The categories *Missing separable particle* and *Superfluous separable particle* refer to the incorrect use of separable particles in phrasal verbs, which are composed of a core verb and a separable particle. These types of errors were only found in the productions of the learners of Dutch. This also holds for the category of error *Missing items (compounds)*. No inter-rater reliability was used for the classification of the errors because it was not feasible in terms of time and money. This means that there was only one coder, but as this doctoral research was carried out in the framework of a project made up of a team of various researchers, advice was sought from colleagues.

4.7. Statistical analyses²⁷

The analyses performed in this doctoral research are based on individual scores; each learner was assigned a score of written complexity, of the total number of phraseological units

²⁷ SPSS version 25 was used for the analyses.

produced, of the total number of phraseological errors made, and of the total number of correct answers given to the receptive test.

Most of the comparative analyses were carried out using an independent samples t-test or a Mann-Whitney U test²⁸. The effect sizes are systematically reported.

Two different statistical tests were conducted for the longitudinal analysis (Chapter 5): a paired-samples t-test for the comparison between the learners' productions at T1 and at T2 and a repeated measures ANOVA for comparing the progress made by the CLIL group and the NON-CLIL group. A lengthier discussion of these tests can be found in Chapter 5, Section 5.3.2.3.

Chi-square tests of independence were carried out in Chapter 6 in order to analyze the relationship between two categorical variables: the collocation used in the test and the type of answer given to the test (correct or incorrect).

In the last empirical study (Chapter 7), linear regression models (hierarchical regression analyses – forced entry) were conducted to examine whether CLIL still remained a predictor of the learners' receptive collocational knowledge after considering a number of other important variables.

4.8. Research design

The present research looking at TL phraseological competence in English and Dutch involves the following comparisons:

- the comparison of the written productive use of phraseological units in CLIL and NON-CLIL settings;
- the comparison of the learners' receptive collocational knowledge in CLIL and NON-CLIL settings;
- the comparison of the learners' phraseological development over time in CLIL and NON-CLIL settings.

To sum up, this thesis' research design involves two different target languages, two different learning conditions, a longitudinal dimension and analyses that encompass various aspects of phraseological knowledge (productive versus receptive; frequency, variety and accuracy of use).

²⁸ Non-parametric equivalent when the normality assumption is not respected.

4.9. Conclusion

This chapter has presented the research hypotheses, the participants that took part in the study, the collected data, the procedure of identification and classification of the phraseological units and phraseological errors, the statistical analyses carried out, and the research design.

Based on the theoretical background provided in the previous chapters and on the pilot study, our three main hypotheses concern the effect of TL exposure on the acquisition of phraseological units, and broadly predict a better and faster acquisition of phraseological units by CLIL learners compared to their NON-CLIL peers. Our sample of participants is relatively large and our data are varied and allow for the examination of TL phraseological competence from a multidimensional perspective.

The next chapters give a comprehensive analysis of the different data types. In Chapter 5, the same corpus of written productions as the one used in the pilot study serves as a basis for the investigation of the learners' productive phraseological knowledge. In Chapter 6, we resort to the experimental data collected from the multiple-choice exercise to analyze the learners' receptive phraseological knowledge. In Chapter 7, we use the data from the multiple-choice exercise in a regression analysis together with the overall measures of vocabulary knowledge and the background variables described above (Sections 4.5.4. and 4.5.5.) in order to examine the actual effect of CLIL on the learners' receptive collocation knowledge when controlling for these learner- and input-related variables.

CHAPTER 5: Productive phraseological competence

5.1. Introduction

Chapter 5 presents the results regarding the learners' productive written phraseological competence in English and Dutch as target languages. It first describes the findings of a cross-sectional study comparing CLIL and NON-CLIL learners' phraseological productions at a single point in time (after the first data collection) (Section 5.2.). After that, results from a longitudinal study looking at the phraseological development of the same CLIL and NON-CLIL learners are discussed (Section 5.3.). Some of this chapter's findings were accepted or submitted for publication²⁹.

5.2. Comparison between CLIL vs NON-CLIL learners' phraseological language at a single point in time

The present study compares the phraseological productions of CLIL and NON-CLIL learners of English and Dutch. The analysis is based on the semi-spontaneous written productions that were gathered during the first data collection. Section 5.2.1. describes the research questions and hypotheses, as well as the data and the participants per TL. Section 5.2.2. presents the results for English and for Dutch. The results section is followed by some concluding remarks and a discussion (Sections 5.2.3. and 5.2.4.).

5.2.1. Method

5.2.1.1. Research questions and hypotheses

This analysis aims to compare the phraseological language of French-speaking secondary school learners of English or Dutch as a TL in CLIL vs NON-CLIL settings. Our research questions are the following:

1. Is there a difference between CLIL and NON-CLIL learners in terms of phraseological range? If so, in which way and to what extent? (RQ1);
2. Is there a difference between CLIL and NON-CLIL learners in terms of phraseological accuracy? If so, in which way and to what extent? (RQ2).

²⁹ Bulon, A. (accepted). Comparing the 'phrasicon' of teenagers in immersive and non-immersive settings: does input quantity impact range and accuracy? *Journal of Immersion and Content-Based Language Education*.

Bulon, A., & Meunier, F. (submitted). Comparing CLIL and NON-CLIL pupils' learning curves in Dutch L2 phraseological competence: the winner does not take it all. *International Journal of Bilingual Education and Bilingualism*.

We expect the phraseological language of CLIL learners to display a greater range of phraseological units and to be more accurate, and this for two main reasons. First, as we have seen in Chapter 3, CLIL programs in French-speaking Belgium provide greater exposure to target input than traditional foreign language learning programs; Belgian secondary school pupils in CLIL settings receive at least twice as many hours (and up to three times) in the TL than those in NON-CLIL settings. Muñoz (2012) refers to the former as intensive exposure programs and to the latter as limited-exposure contexts. Second, CLIL programs aim at both content and language competence development; learners are expected to attain a certain level of proficiency in the TL, with a specific focus on productive and interactive skills (Brown & Bradford, 2017).

5.2.1.2. Data and participants

English

The sample consisted of 180³⁰ fifth year secondary school pupils. They all learned English as a TL through CLIL or NON-CLIL. The participants' ages ranged from 15 to 18 and their mean age was 16. The distribution of the participants across the two subsamples (CLIL/NON-CLIL) and their background characteristics are displayed in Table 5.1. Their (extra)curricular exposure to English is described in Table 5.2. The sample sizes (*n*) in the tables do not always correspond to 180 because some data were missing.

Table 5.1. Distribution across subsamples and background characteristics for the learners of English

	<i>n</i> =180	Gender (%) <i>n</i> =180		Home language (%) <i>n</i> =178		
		Male	Female	French	French + other	Mostly other
CLIL	90	44.4	55.6	64.4	23.3	12.2
NON-CLIL	90	37.8	62.2	68.2	26.1	5.7
ALL	180	41	59	66	25	9

106 of the participants were female (59%) and 74 were male (41%). Regarding the languages spoken at home, most of the participants spoke French all the time (66%), 25% spoke sometimes French and sometimes another language, and a minority spoke another language most of the time (9%). Among those who answered speaking another language than French at home, only five wrote English as one of their home languages – most of them did not answer

³⁰ The sample originally consisted of 218 participants but a number of texts were lost due to technical problems.

the question or mentioned other languages that are of little importance within the scope of this study. We chose not to exclude them.

Table 5.2. Learners' (extra)curricular exposure to English

	<i>n</i>	Nb of hours/week at		Extra	Stay	Informal
		school		classes	abroad	contact
		(%)		(%)	(%)	outside
		(n=178)		(n=178)	(n=178)	school
						(Means)
						(n=176)
		4h/week	+ 4h/week			
CLIL	90	72.2	27.8	3.3	8.9	3.46
NON-CLIL	90	76.1	23.9	0	3.4	3.02
ALL	180	74.2	25.8	1.7	6.2	3.24

The two subsamples are quite homogeneous in terms of linguistic background. We asked the participants how many hours of English as a TL they were currently given at school. The majority of the participants answered 4h/week (72.2% in CLIL and 76.1% in NON-CLIL); some of them received more than 4h/week (28% in CLIL and 24% in NON-CLIL). Only a few CLIL participants took extra classes of English outside school (3.3%). We also asked the participants whether they had stayed in an English-speaking country or region for more than three months (the purpose of the stay was not specified; it could have been holidays or studies abroad) and most of them did not (91.1% in CLIL and 96.6% in NON-CLIL). However, it is worth mentioning that CLIL participants (Mean=3.46, SD=.815) seemed to have more current informal contact with English outside school³¹ than the NON-CLIL participants (Mean = 3.02, SD = .853), $t(174) = 3.505$, $p=.001$, $d=.60$ (medium to large effect size).

The corpus was made of 180 texts that were gathered between October 19th and November 9th 2015 through a writing task which consisted in writing an informal e-mail to a friend about the holidays or a party (see Chapter 4, Section 4.5.1.). A description of the corpus in terms of number of texts, number of words, and mean text length is given in Table 5.3.

³¹ Composite measure based on different variables, on a scale from 1 to 5 (never – rarely – sometimes – often – very often): 1) frequency of use of English on the internet; 2) frequency of use of English outside school (how often do you speak); 3) frequency of input in English outside school (how often do you hear) and 4) frequency of contact with native speakers.

Table 5.3. English corpus description

	Number of texts	Number of words	Mean text length	SD
CLIL	90	29,387	326.52	65.38
NON-CLIL	90	23,747	263.86	82.46

Dutch

The sample consisted of 232³² fifth year secondary school pupils. They learned Dutch as a TL through CLIL or NON-CLIL programs. The participants' ages ranged from 15 to 18 and their mean age was 16. Table 5.4. shows the distribution of the learners across the two subsamples (CLIL/NON-CLIL) and their background characteristics. Table 5.5. provides information about their (extra)curricular exposure to Dutch. The sample sizes (*n*) in the tables do not always correspond to 232 because some data were missing.

Table 5.4. Distribution across subsamples and background characteristics for the learners of Dutch

	<i>n</i> =232	Gender (%) <i>n</i> =232		Home language (%) <i>n</i> =230		
		Male	Female	French	French + other	Mostly other
CLIL	132	54.5	45.5	83.2	15.3	1.5
NON-CLIL	100	43	57	71	26	3
ALL	232	49.6	50.4	78	20	2

117 pupils were female (50.4%), while 115 pupils were male (49.6%). Most of the participants spoke only French at home (78%), while some of them spoke sometimes French and sometimes another language (20%). Only 2% mostly spoke another language than French. Among the pupils who said speaking another language than French at home, only 8 mentioned Dutch – many of them did not provide the information or named a language that is not relevant for our study – and these were not excluded from the analyses.

Table 5.5. Learners' (extra)curricular exposure to Dutch

<i>n</i>	Nb of hours/week at school (%) (<i>n</i> =230)	Extra classes (%) (<i>n</i> =229)	Stay abroad (%) (<i>n</i> =228)	Informal contact outside school

³² The sample originally consisted of 255 pupils, but a few texts were lost due to technical problems.

						(Means) (n=222)
		4h/week	+ 4h/week			
CLIL	132	86	12	1.5	10	2.55
NON-CLIL	100	100	0	7	1	2.07
ALL	232	92	7	4	6	2.34

We can say that the two subsamples are quite comparable regarding the learners' linguistic background. The majority of the participants received 4 hours of TL courses per week (only 12% of CLIL learners had more than 4h/week)³³. Only a minority of learners (1.5% in CLIL and 7% in NON-CLIL) took extra classes of Dutch outside school. To the question whether they had spent more than three months in a Dutch-speaking country or region, most of them answered 'no' (90% in CLIL; 99% in NON-CLIL). However, it seems that the two subsamples differ in their amount of current informal contact with Dutch outside school; an independent samples t-test revealed a higher amount of contact in the CLIL group (Mean = 2.55, SD = .779) compared to the NON-CLIL group (Mean = 2.07, SD = .700), $t(220) = 4.735$, $p=.000$, $d=.67$ (medium to large effect size).

The corpus was made of 232 written texts that were collected between October 19th and November 9th 2015 through the same writing task as for English (informal e-mail to a friend). Table 5.6. gives a description of the corpus.

Table 5.6. Dutch corpus description

	Number of texts	Number of words	Mean text length	SD
CLIL	132	37,092	281.00	58.25
NON-CLIL	100	19,399	193.99	62.96

The methodology used for the extraction and classification of the phraseological units and phraseological errors produced by the learners of English or Dutch was thoroughly described in Chapter 4, Section 4.6.

³³ One CLIL pupil answered having only 2 hours of foreign language classes per week, which seems unlikely since all the other classmates received 4h/week.

5.2.2. Results

We first present the results related to the range of the learners' phraseological vocabulary (RQ1). Global trends for both groups of learners are described before turning to comparative findings related to both frequency and variety. We then proceed with the results regarding the accuracy of the learners' phraseological vocabulary (RQ2). In that section, we first look at the global trends for all learners and then at the comparative findings in terms of frequency. Finally, we apply the distinction made by James (2013) between grammaticality and acceptability errors (see Chapter 1, Section 1.5.3.).

Note that the scale of the figures presented in Section 5.2.2.1.1. below varies from one TL to another.

5.2.2.1. Range

5.2.2.1.1. Global trends

English

The entire group of learners – no distinction being made between CLIL and NON-CLIL pupils here – produced 4,790 phraseological units in total. As shown in Table 5.7., the learners mostly produced *referential phrasemes*, which is not really surprising given the type of words found in that category – it mainly includes combinations of verbs, nouns, adjectives, and adverbs. There might also be a slight task effect since learners had to write an informal e-mail and not an essay, for example, which would possibly have resulted in more *textual phrasemes*. As for *communicative phrasemes*, they are typically less frequent, whatever the text type. We would expect them to be relatively numerous in an e-mail, though.

Table 5.7. Total number of phrasemes produced by all learners of English (n=180) per category (raw frequencies, on a total of 53,134 words)

Referential phrasemes	Textual phrasemes	Communicative phrasemes
4,281	134	375

A closer look at the learners' productions per sub-category (see Figure 5.1. below) shows that they used numerous *lexical collocations*, *compounds*, and *grammatical collocations*. These three categories account for 98% of the *referential phrasemes* produced. *Complex prepositions* and *linking adverbials* are the phrasemes that were the most frequently used by learners as text organizers. When addressing their interlocutor, learners produced *speech act formulae* and *attitudinal formulae*. We found no occurrence of *proverbs*, *slogans* or *commonplaces*.

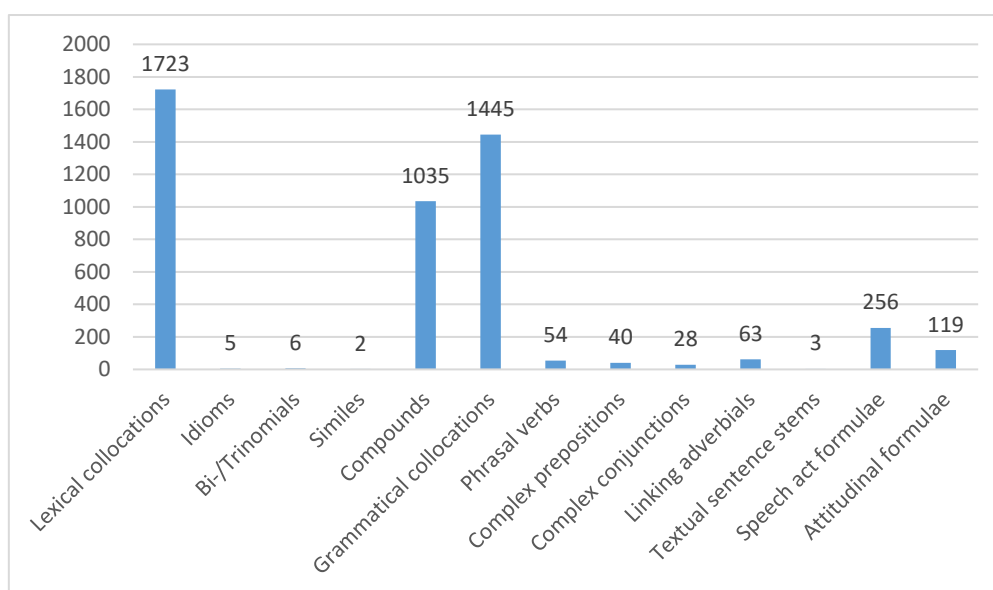


Figure 5.1. Total number of phrasemes produced by all learners of English per sub-category (raw frequencies)

Dutch

The whole group of learners (CLIL and NON-CLIL included) produced 6,367 phraseological units in total. As shown in Table 5.8. and similarly to the learners of English, the learners of Dutch mostly produced *referential phrasemes*, followed by *communicative phrasemes* and *textual phrasemes*.

Table 5.8. Total number of phrasemes produced by all learners of Dutch (n=232) per category (raw frequencies, on a total of 56,491 words)

Referential phrasemes	Textual phrasemes	Communicative phrasemes
5,321	495	551

As shown in Figure 5.2. below, the learners produced numerous phraseological units in the sub-categories *lexical collocations*, *compounds*, and *grammatical collocations*. These sub-categories account for 94% of the *referential phrasemes* produced. The learners also produced a substantial number of *phrasal verbs*. *Complex conjunctions* are clearly the phraseological units that were the most frequently used by the learners to organize their texts, followed by *linking adverbials*. As for the phraseological units that are used to perform a communicative function, learners produced *speech act formulae* and *attitudinal formulae*. No occurrence of *similes*, *proverbs*, *slogans* or *commonplaces* was found in the learners' productions.

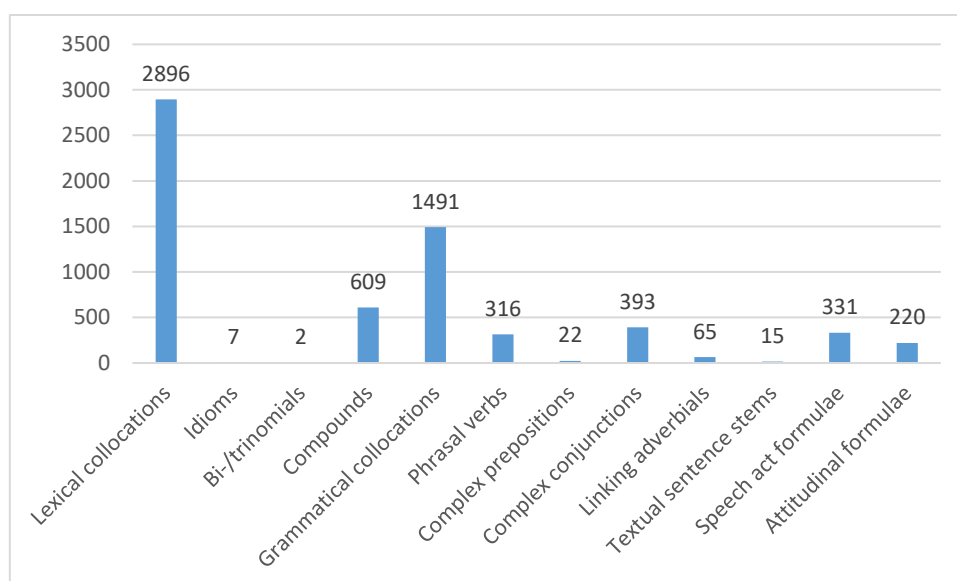


Figure 5.2. Total number of phrasemes produced by all learners of Dutch per sub-category (raw frequencies)

In the following section, we compare the CLIL and the NON-CLIL group in terms of frequency of production of phrasemes overall, but also within the three major categories (*referential phrasemes*, *textual phrasemes*, and *communicative phrasemes*) and the different sub-categories. We have decided to disregard the sub-categories *idioms*, *bi-/trinomials*, *similes* and *textual sentence stems* because of their relatively low frequencies of occurrence.

5.2.2.1.2. Frequency: comparison CLIL/NON-CLIL

English

As can be seen from Table 5.9., CLIL learners produced more phraseological units than NON-CLIL learners. We conducted an independent-samples t-test to see whether that difference was statistically significant. The test indicated a significant difference in the scores for the CLIL group (Mean = 30.90, SD = 9.367) and the NON-CLIL group (Mean = 22.32, SD = 8.909); $t(178) = 6.295$, $p = .000$, $d = .94$ (large effect size).

Table 5.9. Total number of phrasemes produced by CLIL and NON-CLIL learners of English (raw frequencies)

CLIL ($n=90$, total nb of words= 29,387)	NON-CLIL ($n=90$, total nb of words=23,747)
2,781	2,009

We looked at differences between the CLIL and NON-CLIL group within the three broad categories and used a Mann-Whitney test since the data were not normally distributed.

Table 5.10. Mann Whitney U Test results per broad category of phrasemes for each English learner group

	CLIL	NON-CLIL		
	Mean	Mean	U	r
Referential phrasemes	27.54***	20.02	2227.5	-.39
Textual phrasemes	1***	0.49	2347.5	-.37
Communicative phrasemes	2.36	1.81	3525.5	-.11

Notes. *** $p < .001$

An examination of the findings in Table 5.10. shows that CLIL learners produced more phraseological units in all three categories, with significant differences found in *referential phrasemes* and *textual phrasemes*.

We went a step further in the comparison and investigated differences within the various sub-categories.

Table 5.11. Mann Whitney U Test results per sub-category of phrasemes for each English learner group

	CLIL	NON-CLIL		
	Mean	Mean	U	r
Lexical collocations	10.22	8.92	3506	-.12
Compounds	6.82***	4.68	2649	-.30
Grammatical collocations	9.77***	6.29	2242	-.39
Phrasal verbs	0.52***	0.08	3033	-.32
Complex prepositions	0.34***	0.10	3429	-.20
Complex conjunctions	0.24***	0.07	3422	-.23
Linking adverbials	0.38	0.32	4046.5	-.001
Speech act formulae	1.49	1.36	3814	-.05
Attitudinal formulae	0.87	0.46	3529	-.13

Notes. *** $p < .001$

As shown by the results in Table 5.11., the CLIL learners scored better than the NON-CLIL learners in all sub-categories, with significant differences found in five of them – they significantly produced more *compounds*, *grammatical collocations*, *phrasal verbs*, *complex prepositions*, and *complex conjunctions*. Both groups of learners seemed to have a somewhat

equal use of *lexical collocations*, *linking adverbials*, *speech act formulae*, and *attitudinal formulae*.

Dutch

As can be seen from Table 5.12., CLIL learners produced more phraseological units than NON-CLIL learners. We conducted a Mann-Whitney test to see whether that difference was statistically significant since the data were not normally distributed. The test revealed a significant difference in the scores between the CLIL group (Mean = 31.58) and the NON-CLIL group (Mean = 21.98), $U = 2784.5$, $p = .000$, $r = -.50$ (large effect size).

Table 5.12. Total number of phrasemes produced by CLIL and NON-CLIL learners of Dutch (raw frequencies)

CLIL ($n=132$, total nb of words=37,092)	NON-CLIL ($n=100$, total nb of words=19,399)
4,169	2,198

We examined the differences between the CLIL and NON-CLIL group within the three broad categories and used a Mann-Whitney test.

Table 5.13. Mean scores per broad category of phrasemes for each Dutch learner group

	CLIL	NON-CLIL	U	r
Referential phrasemes	26.12***	18.73	3326.5	-.42
Textual phrasemes	2.83***	1.22	3458	-.42
Communicative phrasemes	2.64**	2.03	4977	-.22

Notes. ** $p < .01$. *** $p < .001$

A look at the findings in Table 5.13. shows that CLIL learners significantly produced more phrasemes in all three categories.

Furthermore, we compared and investigated the differences between the groups within the various sub-categories.

Table 5.14. Mean scores per sub-category of phrasemes for each Dutch learner group

	CLIL	NON-CLIL	U	r
Lexical collocations	14.07***	10.40	4175	-.32
Compounds	3.16***	1.92	4369	-.29
Grammatical collocations	7.03**	5.62	5114	-.19
Phrasal verbs	1.83***	0.75	3594	-.40

Complex prepositions	0.14*	0.04	6012	-.15
Complex conjunctions	2.24***	0.97	3956	-.35
Linking adverbials	0.35*	0.19	5873	-.14
Speech act formulae	1.45	1.39	6231	-.05
Attitudinal formulae	1.18***	0.64	4495	-.29

Notes. * $p < .05$. ** $p < .01$. *** $p < .001$

The results presented in Table 5.14. show that the CLIL learners scored better than the NON-CLIL learners in all sub-categories, with significant differences found in 8 of them – they significantly produced more *lexical collocations*, *compounds*, *grammatical collocations*, *phrasal verbs*, *complex prepositions*, *complex conjunctions*, *linking adverbials* and *attitudinal formulae*. Both groups of learners seemed to have a somewhat equal use of *speech act formulae*.

The findings mentioned above show that CLIL learners outperform their NON-CLIL peers in terms of frequency of use, both in English and Dutch as target languages; they significantly produced more phraseological units in nearly all categories and sub-categories. In the following section, we analyze the variety of the learners' phraseological lexicon. In fact, lexical richness is another important dimension of lexical competence and it is possible that CLIL pupils produced numerous phraseological units but always used the same.

5.2.2.2. Variety

We look at variety within the three following sub-categories: *lexical collocations*, *grammatical collocations*, and *phrasal verbs*. The first two sub-categories are the most frequently produced types of phraseological units amongst both groups of learners. They are also further subdivided into various combinations that are worth investigating in detail. As for the sub-category *phrasal verbs*, frequency results showed a big difference between the CLIL and the NON-CLIL group, and for that reason we thought it deserved further attention.

We opted for the lemma/token ratio to measure lexical variation. As Granger and Wynne (1999) stated, “although the type/token ratio is by far the most commonly used measure of lexical richness, it is arguably much less useful in assessing foreign learners' vocabulary than another measure, the lemma/token ratio” (p. 3). Hence different forms of the verb *wait for*, for example – *wait for*, *waited for*, *waiting for* – were counted as one lemma instead of three different types. This was also the case for erroneous forms (e.g. *happy to* and *happpy to* were counted as one lemma, and not as two different types).

5.2.2.2.1. Lexical collocations

English

Table 5.15. shows the number of occurrences (frequency), the number of lemmas (range), and the lemma/token ratio for both groups of learners per combination (ADJECTIVE + NOUN, ADVERB + VERB, VERB + NOUN, and ADVERB + VERB). The rightmost column shows the contrasted lemma ratios between the CLIL and NON-CLIL groups. Concretely speaking, if the CLIL group has 59 different ADJ+NOUN lemma combinations and the NON-CLIL group also has 59, then the CLIL/NON-CLIL lemma ratio will be 1/1 (identical range). In contrast, if the CLIL group has 13 ADV+VERB lemma combinations vs only 2 for the NON-CLIL group, then the CLIL/NON-CLIL lemma ratio will be 6.5/1, showing that for that specific combination the CLIL learners' range is 6.5 higher.

Table 5.15. Lexical collocations – comparison of variety between CLIL and NON-CLIL learners of English

Lexical collocations	CLIL (<i>n</i> =90, total nb of words=29,387)			NON-CLIL (<i>n</i> =90, total nb of words=23,747)			Lemma ratio CLIL/NON-CLIL
	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	
ADJ + NOUN	301	59	0.19	299	59	0.19	1/1
ADV + ADJ	290	58	0.20	312	55	0.17	1.05/1
VERB + NOUN	367	79	0.21	204	43	0.21	1.8/1
ADV + VERB	47	13	0.27	4	2	0.50	6.5/1
TOTAL	1,005	209	0.20	819	159	0.19	1.3/1

If we look at the lemma ratios of CLIL to NON-CLIL learners, we see a little bit more variety in the CLIL group for the total number of lexical collocations produced (1.3/1). Regarding the various combinations, we notice that the CLIL learners produced a wider range of lexical collocations in the combinations ADVERB + VERB, compared to the NON-CLIL group (6.5 times more in this case). Examples include *really like*, *quickly realize*, *always wanted*, *probably*

know, etc. CLIL learners also produced nearly twice as many different VERB + NOUN combinations (e.g. *have fun*, *make friends*, *tell jokes*, etc.) as NON-CLIL learners. The difference in variety is less sharp for the combinations ADVERB + ADJECTIVE (learners mostly used adverbs such as *very*, *really*, *so*, *pretty*, *too*) and inexistent for the combinations ADJECTIVE + NOUN (e.g. *hot weather*, *pretty girl*, *nice people*, etc.).

Dutch

Table 5.16. shows the number of occurrences (frequency), the number of lemmas (range), and the lemma/token ratio for both groups of learners per combination (ADJECTIVE + NOUN, ADVERB + VERB, VERB + NOUN, and ADVERB + VERB).

Table 5.16. Lexical collocations – comparison of variety between CLIL and NON-CLIL learners of Dutch

	CLIL (<i>n</i> =132, total nb of words=37,092)			NON-CLIL (<i>n</i> =100, total nb of words=19,399)			
Lexical collocations	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Lemma ratio CLIL/NON-CLIL
ADJ + NOUN	459	87	0.19	292	62	0.21	1.4/1
ADV + ADJ	626	94	0.15	343	44	0.13	2.1/1
VERB + NOUN	707	147	0.21	372	80	0.21	1.8/1
ADV + VERB	88	14	0.16	38	7	0.18	2/1
TOTAL	1,880	342	0.18	1045	193	0.18	1.8/1

If we look at the lemma ratio of CLIL to NON-CLIL learners for the total number of lexical collocations produced, we see that the CLIL learners produced nearly twice as many different types of combinations (1.8/1) as the NON-CLIL learners. Regarding the various combinations, we see a bit more variety in the combinations ADJECTIVE + NOUN (1.4/1) (e.g. *slecht weer* ('bad weather'), *vorig jaar* ('last year'), *groot probleem* ('big problem'), etc.). CLIL learners also produced twice as many different ADVERB + ADJECTIVE and ADVERB + VERB combinations as NON-CLIL learners. They mostly used adverbs such as *heel* ('very'), *zo* ('so'),

zeer ('very'), *echt* ('really'), *te* ('too'), *enorm* ('enormously'), *erg* ('terribly'), *redelijk* ('rather'), and *wel* ('quite') in combination with adjectives. With verbs, they used for example the adverbs *goed*, *lekker* or *nooit* (ex: *goed gaan* ('go well'), *lekker eten* ('to feast', 'to eat well'), *nooit vergeten* ('never forget')). Compared to the NON-CLIL group, CLIL learners produced a wider range of lexical collocations in the combinations VERB + NOUN (e.g. *tocht maken* ('make a trip'), *schade aanrichten* ('do damage'), *zin hebben* ('to feel like'), etc.) (1.8 times more in this case).

5.2.2.2.2. Grammatical collocations

English

Table 5.17. shows the number of occurrences (frequency), the number of lemmas (range), and the lemma/token ratio for CLIL and NON-CLIL learners per combination (ADJECTIVE + PREPOSITION, ADJECTIVE + TO + INFINITIVE, PREPOSITION + NOUN, NOUN + PREPOSITION, VERB + TO + INFINITIVE, and VERB + PREPOSITION). The rightmost column (lemma ratio CLIL/NON-CLIL) gives the proportion of lemmas produced by the CLIL group in comparison to the NON-CLIL group.

Table 5.17. Grammatical collocations – comparison of variety between CLIL and NON-CLIL learners of English

	CLIL (<i>n</i> =90, total nb of words=29,387)			NON-CLIL (<i>n</i> =90, total nb of words=23,747)			
Grammatical collocations	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Lemma ratio CLIL/NON-CLIL
ADJ + PREP	48	11	0.22	21	4	0.19	2.7/1
ADJ + TO + INF	25	3	0.12	9	1	0.11	3/1
PREP + NOUN	241	47	0.19	128	24	0.18	1.9/1
NOUN + PREP	73	16	0.21	32	10	0.31	1.5/1
VERB + TO + INF	204	12	0.05	88	7	0.07	1.7/1

VERB + PREP	400	43	0.10	371	35	0.09	1.2/1
TOTAL	991	132	0.13	649	81	0.12	1.6/1

The lemma ratio of CLIL to NON-CLIL learners for the total number of grammatical collocations indicates a wider range in the CLIL group compared to the NON-CLIL group (1.6/1). More specifically, CLIL learners' range of ADJ + TO + INFINITIVE combinations is three times higher than that of NON-CLIL learners (e.g. *able to V*, *afraid to V*). The number of different grammatical collocations of the type ADJECTIVE + PREPOSITION produced by CLIL learners is more than double that of NON-CLIL learners. Examples include *good at*, *proud of*, *different from*, etc. CLIL learners also used nearly twice as many different PREPOSITION + NOUN (e.g. *by car*, *at lunch*, *in love*, etc.), and VERB + TO + INFINITIVE combinations (e.g. *start to V*, *try to V*, *decide to V*, etc.) than NON-CLIL learners. We also see more variety in the CLIL group regarding the combinations NOUN + PREPOSITION (e.g. *trip to*, *problem with*, etc.) and to a lesser extent in the combinations VERB + PREPOSITION (e.g. *forget about*, *come from*, *listen to*, etc.).

Dutch

Table 5.18. shows the number of occurrences (frequency), the number of lemmas (range), and the lemma/token ratio for CLIL and NON-CLIL learners per combination (ADJECTIVE + PREPOSITION, NOUN + TO + INFINITIVE, PREPOSITION + NOUN, NOUN + PREPOSITION, VERB + TO + INFINITIVE, and VERB + PREPOSITION).

Table 5.18. Grammatical collocations – comparison of variety between CLIL and NON-CLIL learners of Dutch

Grammatical collocations	CLIL (<i>n</i> =132, total nb of words=37,092)			NON-CLIL (<i>n</i> =100, total nb of words=19,399)			Lemma ratio CLIL/NON-CLIL
	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	Nb of occ. (freq)	Nb of lemmas (range)	Lemma /token ratio	
ADJ + PREP	39	10	0.26	27	6	0.22	1.7/1
NOUN + TO + INF	8	2	0.25	2	1	0.50	2/1

PREP + NOUN	373	49	0.13	224	29	0.13	1.7/1
NOUN + PREP	47	12	0.25	19	6	0.31	2/1
VERB + TO + INF	28	5	0.18	7	3	0.43	1.7/1
VERB + PREP	447	37	0.08	272	23	0.08	1.6/1
TOTAL	942	115	0.12	551	68	0.12	1.7/1

The lemma ratio of CLIL to NON-CLIL learners for the total number of grammatical collocations indicates a wider range in the CLIL group compared to the NON-CLIL group (1.7/1). More specifically, CLIL learners' range of NOUN + TO + INFINITIVE (e.g. *zin (om) te V* ('desire to V'), *tijd (om) te V* ('time to V')) and NOUN + PREP combinations (e.g. *zin in* ('desire for'), *problemen met* ('problems with'), *respect voor* ('respect for'), *uitzicht op* ('view across/over'), etc.) is twice higher than that of NON-CLIL learners. The number of different grammatical collocations of the type VERB + PREPOSITION (e.g. *genieten van* ('to enjoy'), *gaan naar* ('to go to'), *houden van* ('to love'), etc.) produced by CLIL learners is 1.6 higher than that of NON-CLIL learners. CLIL learners also used nearly twice as many different ADJECTIVE + PREPOSITION (e.g. *verliefd op* ('in love with'), *bang voor* ('afraid of'), *blij met* ('happy about/with'), etc.), PREPOSITION + NOUN (e.g. *op vakantie* ('on holiday'), *ter wereld* ('in the world'), *met de trein* ('by train'), etc.) and VERB + TO + INFINITIVE combinations (e.g. *beginnen te V* ('start to V'), *proberen te V* ('try to V'), *durven te V* ('dare to V'), etc.) than NON-CLIL learners.

5.2.2.2.3. Phrasal verbs

English

Table 5.19. shows the number of occurrences (frequency), the number of lemmas (range), the lemma/token ratio, the number of learners using phrasal verbs (PVs), the average number of phrasal verbs per user, and the lemma ratio of CLIL to NON-CLIL learners.

Table 5.19. Phrasal verbs (PVs) – comparison of variety between CLIL and NON-CLIL learners of English

	CLIL (n=90)	NON-CLIL (n=90)

Number of occurrences (freq)	51	7
Number of lemmas (range)	14	2
Lemma/token ratio	0.27	0.28
Number of pupils using PVs	28	6
Average number per user	1.7	1.1
Lemma ratio CLIL/NON-CLIL	7/1	

From Table 5.19. it can be seen that the NON-CLIL learners used a very narrow set of phrasal verbs (2) in contrast to the CLIL learners who used a much wider range (14). In fact, in the NON-CLIL texts we only found occurrences of the verbs *wake up* and *go out (with someone)*. These were also spotted in the CLIL texts, alongside *end up*, *break up*, *find out*, *sum up*, *clean up*, *pick up*, *turn out*, *throw up*, *go on*, *come up*, *run away*, and *go by*. The range of phrasal verbs in the CLIL group is seven times higher than that of the NON-CLIL group. Moreover, when we look at the number of learners (out of 90 in each group) who used phrasal verbs, we notice that these are more numerous in the CLIL group (28 out of 90 (31%), versus only 6 out of 90 (7%) in the NON-CLIL group). They also used more phrasal verbs on average.

Dutch

Table 5.20. provides the number of occurrences (frequency), the number of lemmas (range), the lemma/token ratio, the number of learners using PVs, the average number of PVs per user, and the lemma ratio of CLIL to NON-CLIL learners.

Table 5.20. Phrasal verbs (PVs) – comparison of variety between CLIL and NON-CLIL learners of Dutch

	CLIL (n=132)	NON-CLIL (n=100)
Number of occurrences (freq)	241	75
Number of lemmas (range)	33	13
Lemma/token ratio	0.14	0.17
Number of pupils using PVs	106	45
Average number per user	1.8	0.75
Lemma ratio CLIL/NON-CLIL	2.5/1	

As indicated in Table 5.20., the range of phrasal verbs produced by the learners is less extensive in the NON-CLIL group (13) than in the CLIL group (33). In fact, the CLIL learners' range of phrasal verbs is more than twice higher than that of the NON-CLIL group. The NON-CLIL

learners used the following PVs: *terugkomen* ('to come back'), *uitnodigen* ('to invite'), *uitgaan* ('to go out'), *uitleggen* ('to explain'), *teruggaan* ('to go back'), *aanbieden* ('to offer'), *aankomen* ('to arrive'), *opbellen* ('to phone'), *overgeven* ('to throw up'), *(zich) afvragen* ('to ask oneself'), *voorstellen* ('to propose', 'to present'), *klaarmaken* ('to prepare'), and *kennismaken* ('to get acquainted with'). The CLIL learners also produced these PVs (excluding *(zich) afvragen*), together with *plaatsvinden* ('to take place'), *binnenkomen* ('to come inside'), *meebrengen* ('to bring along'), *eruitzien* ('to look like'), *doorbrengen* ('to spend', 'to pass'), *paardrijden* ('to ride a horse'), etc. Also, when looking at the number of pupils (out of 132 for the CLIL group and out of 100 for the NON-CLIL group) who made use of PVs in their texts, we see that they are 106 (80%) in the CLIL group, compared to 45 (45%) in the NON-CLIL group. The CLIL learners also produced more phrasal verbs on average.

5.2.2.3. Accuracy

We will now turn to the results related to the learners' phraseological errors³⁴. It is important to note that errors here are not to be viewed as highly problematic features that would impede communication – on the whole, both the CLIL and NON-CLIL written productions were easily comprehensible and would not have led to communication breakdowns. The phraseological errors were classified according to the taxonomy presented in Chapter 4, Section 4.6.

Note that the scale of the figures presented in Section 5.2.2.3.1. below varies from one TL to another.

5.2.2.3.1. Global trends

English

The total number of phraseological errors found in the productions of all the learners of English (both CLIL and NON-CLIL) amounts to 478.

³⁴ The term 'phraseological errors' refers to errors occurring within a phraseological unit; they may be inherently phraseological, but not always.

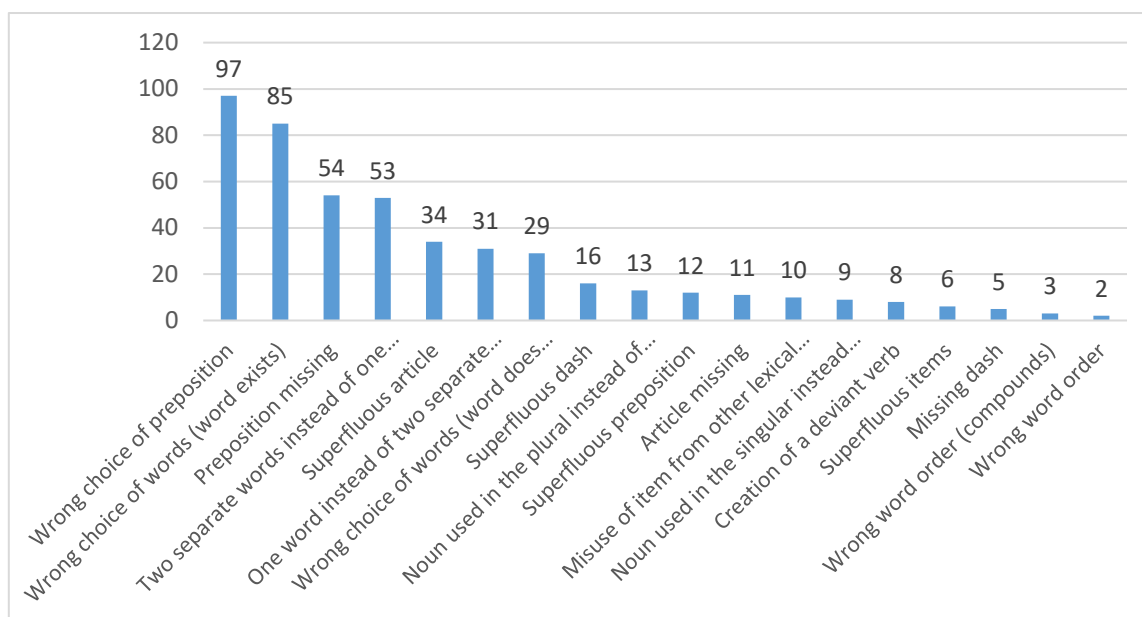


Figure 5.3. Total number of phraseological errors produced by all learners of English per category (raw frequencies)

As shown in Figure 5.3., the most frequent categories of errors concern prepositions (*wrong choice of prepositions* and *preposition missing*) and a wrong selection of words, most often replaced by an existing one (*wrong choice of words (the word exists)*). The categories *two separate words instead of one word* and *one word instead of two separate words* relate to the misspelling of open versus solid compounds³⁵ – learners combined the right elements but made a spelling mistake. We may refer to it as a formal error since it is related to form, not to meaning. Another frequent category of errors is the overuse of articles (*superfluous article*).

Since learners' texts differed in terms of number of words produced, we also looked at frequencies of errors per 300 words, the average text length, per learner (see Figure 5.4. below).

³⁵ In English, solid compounds are spelled as one orthographic word; they do not include a blank space between their constituents and may include hyphens or not (e.g. *seafood*, *heart-warming*). In contrast, the constituents of open compounds are separated by blank spaces (e.g. *ice cream*).

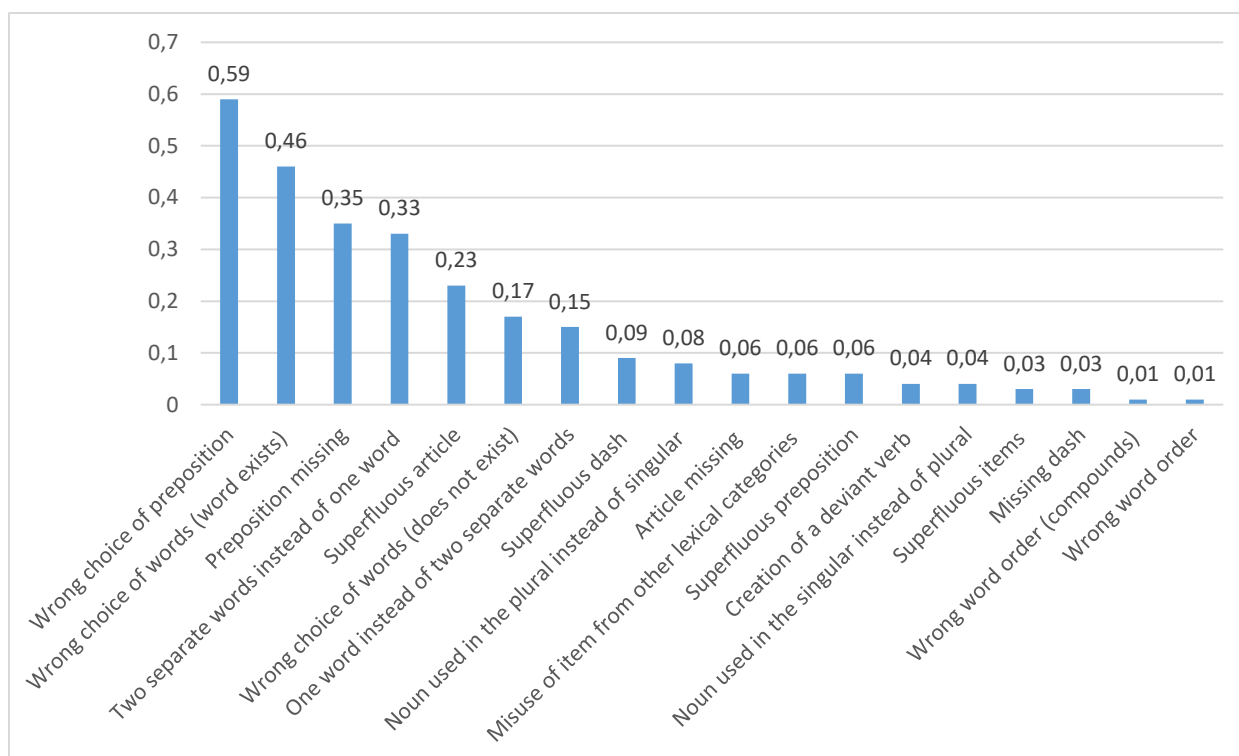


Figure 5.4. Total number of phraseological errors produced by all learners of English per category (relative frequencies per learner)

Figure 5.4. shows more or less the same picture as Figure 5.3. The most frequent categories of errors are still the same.

Dutch

The total number of phraseological errors found in the productions of all the learners of Dutch (both CLIL and NON-CLIL) amounts to 1,087.

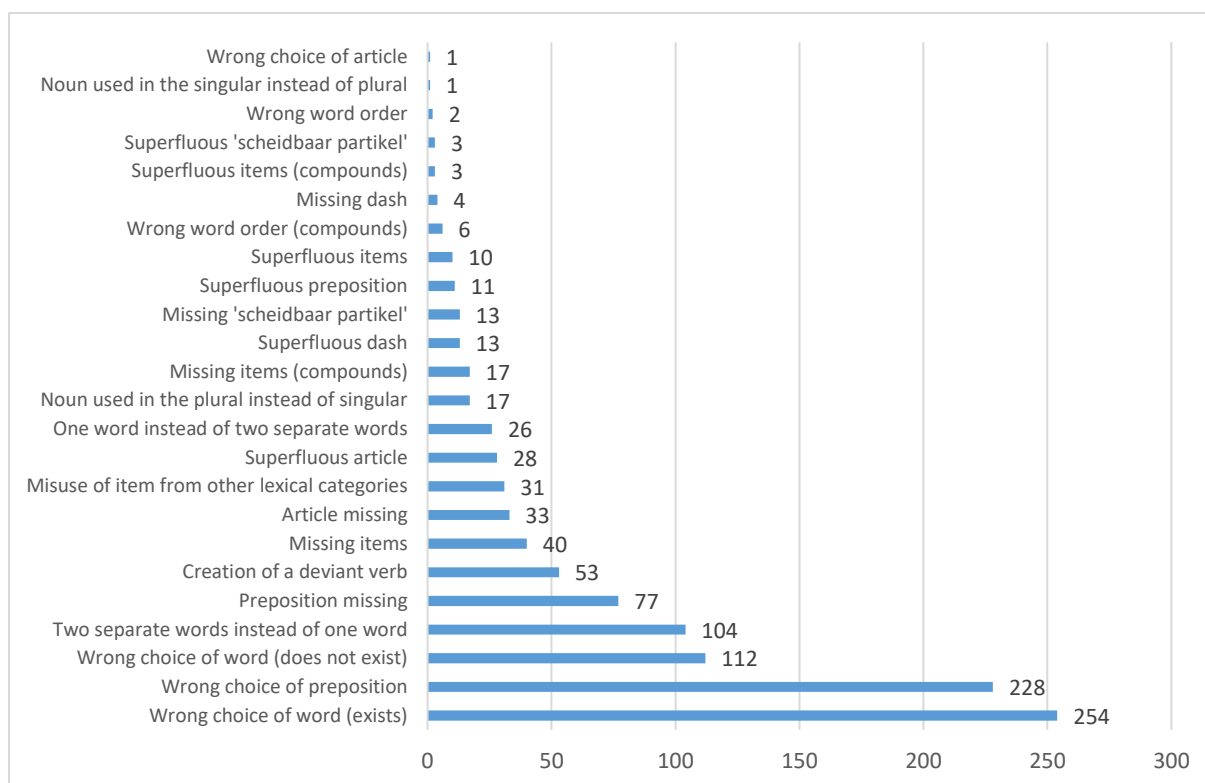


Figure 5.5. Total number of phraseological errors produced by all learners of Dutch per category (raw frequencies)

Figure 5.5. shows that the most frequent categories of errors involve prepositions (*wrong choice of prepositions* and *preposition missing*) and a wrong selection of words, most often replaced by an existing one (*wrong choice of words (the word exists)*). The category *two separate words instead of one word* is also found in the top most frequent categories of phraseological errors. Two other frequent categories of errors are the *creation of a deviant verb* and the *misuse of item from other lexical categories*. The former type of error refers to the fact that learners produced an erroneous verb form such as **spreekten Engels [spraken]* (literally ‘*spoke English [spoke]’). The latter refers to errors such as *volgens *ik [mij]* (literally ‘according to *I [me]’). The use of articles also seems to be problematic since learners sometimes forget – or consciously choose not – to use them (*article missing*, as in *naar *[de] bioscoop gaan*, literally ‘go to *[the] cinema’), and sometimes overuse them (*superfluous articles*, as in *op *de zaterdag*, literally ‘on *the Saturday’). A last prominent category of error is *missing items*, which refers to mistakes such as *jammer [genoeg]* (‘unfortunately’).

The texts differed in terms of number of words produced, hence we also looked at relative frequencies of errors per 250 words (the average text length of the Dutch productions), per learner (see Figure 5.6. below).

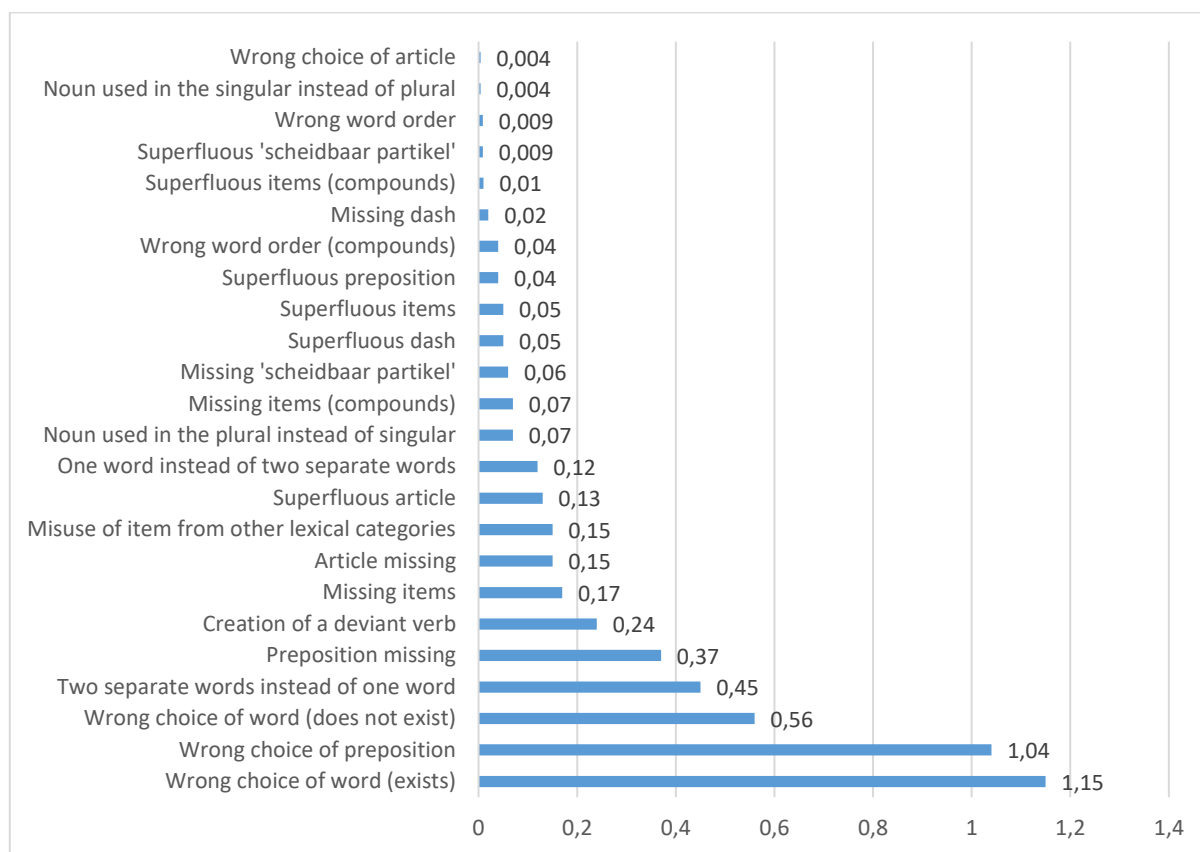


Figure 5.6. Total number of phraseological errors produced by all learners of Dutch per category (relative frequencies per learner)

Figure 5.6. shows a very similar picture to Figure 5.5.; the most frequent categories of errors are still the same.

The following section includes comparative analyses between CLIL and NON-CLIL learners for the total number of errors produced and for the most frequent categories of errors. Relative frequencies are used.

5.2.2.3.2. Frequency: comparison CLIL/NON-CLIL

English

As shown in Table 5.21., CLIL learners produced fewer phraseological errors than NON-CLIL learners. We carried out a Mann-Whitney test in order to see whether this difference was statistically significant. The test indicated that the number of phraseological errors produced was significantly higher in the NON-CLIL group than in the CLIL group: $U = 2839.5$, $p = .001$, $r = -.26$ (small to medium effect size).

Table 5.21. Total number of phraseological errors produced by CLIL and NON-CLIL learners of English (relative frequencies per 300 words per learner)

CLIL	NON-CLIL
2.2	3.5

If we look at the total number of phraseological errors in relation to the total number of phrasemes produced for both groups, we see that the NON-CLIL learners made twice as many mistakes (10%) as the CLIL learners (5%).

We compared the relative frequencies of errors produced by CLIL and NON-CLIL learners (on 300 words) for the most frequent categories of errors using a Mann-Whitney test.

Table 5.22. Mann Whitney U test results per category of errors for each English learner group

	CLIL	NON-CLIL		
	Mean	Mean	U	r
Wrong choice of preposition	0.07	0.56**	3250.5	-.20
Wrong choice of words (word exists)	0.47	0.45	3985	-.02
Preposition missing	0.2	0.55*	3726.5	-.10
Two separate words instead of one word	0.37	0.38	3981	-.02
Superfluous article	0.08	0.37**	3391	-.22
Wrong choice of words (word does not exist)	0.16	0.19	3868.5	-.06
One word instead of two separate words	0.17	0.14	4014	-.01

Notes. * $p < .05$. ** $p < .01$.

We can see from Table 5.22. that NON-CLIL pupils significantly produced more errors in the *wrong choice of preposition*, *superfluous article*, and *preposition missing* categories. They also made slightly more mistakes in the *wrong choice of words (word does not exist)* and *two separate words instead of one word* categories, although not significantly. Interestingly, CLIL learners appeared to make a few additional errors in the *wrong choice of words (word exists)* category – unlike their NON-CLIL peers, when they failed to select the correct word, they more often chose an existing English word. The CLIL group also seemed to make a little bit more errors in the *one word instead of two separate words* category.

A closer look at errors in the category *wrong choice of preposition* revealed that learners mostly selected wrong prepositions in *grammatical collocations* (e.g. **in* the beach) and *complex*

prepositions (e.g. ***at** the top of). However, CLIL learners seemed to make more mistakes when the preposition preceded a noun (e.g. ***in** the airport, ***on** the afternoon/evening), while NON-CLIL learners had more difficulties when the preposition was combined with a verb (e.g. think ***at**, tell ***for**, giving something ***at** somebody).

It seems that the overuse of articles – which was mainly associated with the NON-CLIL group – could possibly be traced back to the learners’ L1 (French). Examples include ‘all ***the** night’ (in French *toute la nuit*), ‘***the** next week’ (in French *la semaine prochaine*), ‘***the** last night’ (in French *la nuit passée*) or ‘eat ***a** barbecue’ (in French *manger un barbecue*).

Regarding the *preposition missing* category, NON-CLIL learners mostly forgot to use a preposition in the combination VERB + PREPOSITION (e.g. talk **[to]**, wait **[for]**, listen **[to]**, live **[in]**, want **[to]**, tell **[about]**). We found more variety in the CLIL group (e.g. put music **[on]**, talked **[to]**, **[in]** front of, **[for]** a while).

Both CLIL and NON-CLIL learners sometimes had difficulties finding the right word. When learners resorted to an incorrect, but existing English word, it mostly occurred with verbs in *lexical collocations* (e.g. ***fished** a fish, ***pass** holidays, ***made** fun, ***did** noise, ***make** shopping), but also with nouns in *compounds* (e.g. ***alcohol** games, ***attraction(s)** park, tennis ***field**, ***aqua** park). When learners used a non-existing word, it mostly occurred with adjectives in *lexical collocations* for the NON-CLIL group (e.g. very ***sympathique**, very ***interessant**, hotel was ***luxious**) and with verbs in *lexical collocations* for the CLIL group (e.g. ***pasted** time, ***repass** an exam). Non-existing words were also found in *compounds* for both groups of learners (e.g. coral ***bareer**, ***bier(-)**pong, night ***boîte**, ***groot**-mother). Once again, it is not unreasonable to think that the learners’ L1 (French) is one of the possible reasons for some of these errors – in French we say *pêcher un poisson*, *passer des vacances*, *parc d’attractions*, *très sympathique*, *très intéressant*, *repasser un examen*, *barrière de corail*, *boîte de nuit*, etc. Moreover, the non-existing words *bier* and *groot* found in ***bier(-)**pong and ***groot**-mother might originate from the learners’ other TL (Dutch).

Learners were also creative in the spelling of compounds; they often wrote open compounds in one word (e.g. **swimmingpool**, **dancefloor**, **icecream**, **waterpark**) and solid compounds in two separate words (e.g. **birth day**, **sea food**, **bed room**, **grand mother**). It is not entirely surprising to find that type of errors since the standard spelling norm of English compounds is not very stable; as Kuperman and Bertram (2013) stated, “the rules that prescribe the correct spelling of

compounds are neither accurate nor exhaustive” (p. 940), which explains why we can come across different variants of a same compound in the language of native speakers too.

Dutch

As indicated in Table 5.23., CLIL learners produced on average fewer phraseological errors than NON-CLIL learners. A Mann-Whitney test revealed that this difference was statistically significant: the number of phraseological errors produced was significantly higher in the NON-CLIL group than in the CLIL group: $U = 4448.5$, $p = .000$, $r = -.28$ (small to medium effect size).

Table 5.23. Total number of phraseological errors produced by CLIL and NON-CLIL learners of Dutch (relative frequencies per 250 words per learner)

CLIL	NON-CLIL
4.01	6.23

When we look at the total number of phraseological errors in relation to the total number of phrasemes produced for both groups, we see that the NON-CLIL learners made nearly twice as many mistakes (22%) as the CLIL learners (14%).

We compared the relative frequencies of errors produced by CLIL and NON-CLIL learners (on 250 words) for the most frequent categories of errors using a Mann-Whitney test.

Table 5.24. Mann Whitney U test results per category of errors for each Dutch learner group

	CLIL	NON-CLIL	U	r
Wrong choice of words (word exists)	0.99	1.36	6261	-.04
Wrong choice of preposition	0.77	1.39**	5261	-.18
Wrong choice of words (word does not exist)	0.24	0.97***	4605.5	-.31
Two separate words instead of one word	0.51*	0.37	5567	-.16
Preposition missing	0.31	0.46	6375	-.04
Creation of a deviant verb	0.22	0.28	6496	-.02
Missing items	0.15	0.18	6435.5	-.03
Article missing	0.07	0.26*	5875	-.16
Misuse of item from other lexical categories	0.13	0.17	6462	-.03
Superfluous article	0.07	0.21	6085.5	-.12
One word instead of two separate words	0.10	0.14	6544	-.01

Notes. * $p < .05$. ** $p < .01$. *** $p < .001$

We can see from Table 5.24. that NON-CLIL pupils significantly produced more errors in the *wrong choice of preposition*, *wrong choice of words (word does not exist)*, and *article missing* categories. NON-CLIL learners also made slightly more mistakes in all other categories although not significantly, with the exception of the *two separate words instead of one word* category where CLIL learners appeared to make significantly more errors.

We have seen that learners very often used incorrect words and that these were existing Dutch words, or not. When learners used an incorrect but existing word, it mostly involved verbs in *lexical collocations* (e.g. vliegtuig ***krijgen** [nemen] ('take the plane'), activiteiten ***maken** [doen] ('do activities'), in het ziekenhuis ***staan** [liggen] ('be in hospital')). The NON-CLIL learners also very frequently used the adverb *veel* instead of *heel* in combination with adjectives (e.g. ***veel** grappig ('very funny'), ***veel** leuk ('very nice'), ***veel** lekker ('very good'), etc.) (see also Hendriks, 2019). When learners used a non-existing word, it mostly involved adjectives in *lexical collocations* (e.g. heel ***plaisant** [plezant] ('very pleasant'), ***last** jaar [vorig] ('last year'), ***zonnelijk** weer [zonnig] ('sunny weather'), heel ***small** [klein] ('very small')) and *compounds* (e.g. ***transats** [strandstoelen] ('beachchairs'), ***football** [voetbal] ('football'), ***hike** [trektocht] ('hike'), tennis***terrain** [veld] ('tennis court'), ***grandvader** [groot] ('grandfather')). A number of non-existing words also resulted from the use of English prepositions in *grammatical collocations*, and this especially concerned the NON-CLIL group (e.g. ***on** de strand [op] ('on the beach'), ***at** de fuif [op] ('at the party'), ***at** 8 uur [om] ('at 8 o'clock')). Some of the phraseological errors are undoubtedly due to the influence of the learners' L1 French (for example *plaisant*, *transats*, *terrain*) and their second TL (English) (for example *last*, *small*, *football*, *hike* and the prepositions *on* and *at*).

A more qualitative look at the phraseological errors made in the category *wrong choice of preposition* revealed that learners mostly used an incorrect preposition in *grammatical collocations*, and more frequently when the preposition was used in combination with a noun (e.g. ***in** school [op] ('at school'), ***aan** een feestje [op] ('at a party'), ***aan** 16 uur [om] ('at 4 pm')) or a verb (e.g. gaan ***in** [naar] ('to go to'), houden ***aan** [van] ('to love'), geven ***naar** [aan] ('to give to'), zorgen ***op** [voor] ('to look after')). As far as the *preposition missing* category is concerned, learners mostly forgot the preposition in VERB + PREPOSITION combinations (e.g. houden **[van]** ('to love'), kijken **[naar]** ('to look at'), spreken **[over]** ('to talk about'), spreken **[met]** ('to talk with')).

CLIL learners seemed particularly creative in the spelling of solid compounds³⁶; they frequently wrote them in two separate words instead of one word (e.g. **dood moe** ('exhausted'), **chocolade taart** ('chocolate cake'), **groot ouders** ('grandparents'), **vakantie plek** ('holiday spot'), etc.). However, this type of errors is related to spelling and the fact that the learners combined the two correct words seems more important from a phraseological point of view.

A significant number of missing articles were found in the phraseological units produced by the NON-CLIL learners and this concerned a variety of word combinations: *speech act formulae* (e.g. tot **[de]** volgende keer ('see you next time')), *linking adverbials* (e.g. in **[het]** algemeen ('in general')), *lexical collocations* (e.g. naar **[de]** bioscoop gaan ('go to the cinema')), or *grammatical collocations* (e.g. in **[het]** hotel ('in the hotel')).

5.2.2.3.3. Grammaticality versus acceptability errors

During the error analysis, we noticed that phraseological units could be erroneous in two different ways. On the one hand, we found what James (2013) calls *grammaticality errors*. They refer to formally inexistent phrasemes, which can be a very close approximation of an existing phraseme (for example 'she is in love ***for** [with] him') or bear no resemblance at all to any existing phraseme (for instance 'he ***made a balad**' [went for a walk]). On the other hand, we found *acceptability errors*, which are formally existing phrasemes but inappropriately used in context, e.g. 'she ***participated in** [attended] the party' (in reference to the fact that the girl came to the party but was not actively involved in its organization).

Erroneously spelled compounds were not taken into consideration in the classification as the standard norm is subject to change.

English

Table 5.25. Number of grammaticality and acceptability errors produced by all learners of English (raw frequencies + mean number of errors per 300 words per learner)

Grammaticality errors	Acceptability errors
312 (= mean nb of 1.85 per learner)	49 (= mean nb of 0.30 per learner)

Table 5.25. shows that learners – no distinction being made between CLIL and NON-CLIL here – produced more grammaticality errors than acceptability errors. In other words, most erroneous

³⁶ Unlike English compounds, Dutch compounds are always spelled as one orthographic word.

phraseological units produced by the learners did not exist, even though some of them were close to the original.

Table 5.26. Number of grammaticality and acceptability errors per group of learners of English (raw frequencies + mean number of errors per 300 words per learner)

	CLIL	NON-CLIL
Grammaticality errors	137 (= mean nb of 1.4) (93.2%)	175 (= mean nb of 2.3) (81.7%)
Acceptability errors	10 (= mean nb of 0.1) (6.8%)	39 (= mean nb of 0.5) (18.3%)
TOTAL	147 (= mean nb of 1.5) (100%)	214 (= mean nb of 2.8) (100%)

We found a significant difference in the number of acceptability errors between CLIL (SD = 0.20) and NON-CLIL (SD = 0.59) learners: $t(31) = -2.726$, $p = .010$, $d = 1.18$ (very large effect size). The NON-CLIL group was also found to include significantly more grammaticality errors than the CLIL group: $U = 2871$, $p = .001$, $r = -0.25$ (small to medium effect size).

Dutch

Table 5.27. Number of grammaticality and acceptability errors produced by all learners of Dutch (raw frequencies + mean number of errors per 250 words per learner)

Grammaticality errors	Acceptability errors
893 (= mean nb of 4.1 per learner)	29 (= mean nb of 0.1 per learner)

It is clear from Table 5.27. that learners (taking into account both CLIL and NON-CLIL learners) produced more formally non-existing phraseological units (grammaticality errors) than formally existing phrasemes used in an inappropriate context (acceptability errors).

Table 5.28. Number of grammaticality and acceptability errors per group of learners of Dutch (raw frequencies + mean number of errors per 250 words per learner)

	CLIL	NON-CLIL
Grammaticality errors	469 (= mean nb of 3.20) (97.3%)	424 (= mean nb of 5.40) (96.4%)
Acceptability errors	13 (= mean nb of 0.08) (2.7%)	16 (= mean nb of 0.20) (3.6%)

TOTAL	482 (= mean nb of 3.28) (100%)	440 (= mean nb of 5.60) (100%)

No significant difference was found between the CLIL and the NON-CLIL learners, neither in terms of grammaticality errors ($U = 6085.5$, $p = .305$, $r = -0.07$), or in terms of acceptability errors ($U = 6438.5$, $p = .553$, $r = -0.04$).

It is important to note that some erroneous phraseological units classified as grammaticality errors remain interesting from a pedagogical point of view. In fact, they demonstrate that both groups of learners are in a stage where they are still making hypotheses about the TL. For example, erroneous phraseological units such as ‘**spend* [have] a good time’ or ‘**naar* muziek beluisteren’ (‘listen to music’) are evidence of the fact that learners have probably learned and acquired the combinations *spend time* and *naar muziek luisteren* and attempted to use them with similar structures without being aware of the collocational restrictions.

5.2.3. Discussion and implications

The present analysis attempted to compare the phraseological language of French-speaking secondary school learners of English or Dutch as a TL in CLIL and NON-CLIL settings. We were especially interested in finding to what extent and in which way their phraseological lexicon differed in terms of range and accuracy.

Regarding RQ1 (range), the results showed that the CLIL learners of TL English produced significantly more phraseological units than their NON-CLIL counterparts, and this in a large number of (sub-)categories. We found no significant difference between the learners regarding *lexical collocations*, *linking adverbials*, and *communicative phrasemes* (both *speech act formulae* and *attitudinal formulae*). In addition to producing more phraseological units than their NON-CLIL peers, CLIL learners also displayed more variety in the categories of phrasemes investigated in this study, namely *lexical collocations*, *grammatical collocations*, and *phrasal verbs*.

As for the learners of TL Dutch, the CLIL group also showed a significantly higher production of phraseological units than the NON-CLIL group, and this in all (sub-)categories except the *speech act formulae* category. More variety was also found in the CLIL group for the categories of phraseological units under examination, namely *lexical collocations*, *grammatical collocations*, and *phrasal verbs*.

As far as RQ2 is concerned (accuracy), the present study concerned with the learners of TL English revealed a significantly higher number of phraseological errors in the NON-CLIL group. NON-CLIL learners significantly made more mistakes in the use of prepositions (*wrong choice of preposition* and *preposition missing*) and articles (overuse). Their texts also contained a higher number of erroneous phraseological units due to the selection of a non-existing word and misspelt solid compounds (*two separate words instead of one word*), although not significantly. As for CLIL learners, they produced a little bit more erroneous phraseological units involving existing words and had more difficulties with the spelling of open compounds (*one word instead of two separate words*) but these differences were not significant. Regarding the grammaticality versus acceptability nature of the learners' errors, we observed that both groups of learners mostly produced grammaticality errors. The NON-CLIL group significantly produced more grammaticality and acceptability errors than the CLIL group.

The NON-CLIL learners of TL Dutch also made significantly more phraseological errors on the whole than their peers in a CLIL context. They made more errors in almost all the most frequent categories, with significant differences found in the *wrong choice of preposition*, *wrong choice of word (word does not exist)*, and *article missing* categories. The CLIL learners significantly made more mistakes in the category involved with the spelling of solid compounds (*two separate words instead of one word*). When distinguishing between grammaticality and acceptability errors, we found that learners mainly produced grammaticality errors, but no significant difference was found between the CLIL and NON-CLIL learners.

Some of the findings presented in this study corroborate previous results mentioned in Chapter 1, Section 1.9.1.1. In fact, the greater variety displayed in the CLIL learners' texts can be related to Mewald's (2007) and Celaya's (2008) studies which showed a higher range of words in the productions of CLIL learners compared to NON-CLIL learners. Our results are also in line with Berendse's findings (2014) in the sense that NON-CLIL learners significantly made more mistakes in their use of prepositions. However, the fact that NON-CLIL learners produced more phraseological errors overall goes against the results of Verspoor and Edelenbos (2011) who found more errors in the CLIL group. Moreover, the productions written by CLIL learners did not seem to include more lexical inventions (as found by Celaya, 2008 and Agustín Llach, 2009) – on the contrary, we found more occurrences of non-existing words in the texts written by NON-CLIL learners. However, comparisons with these previous studies should be made with caution as they all looked at English as a TL (not Dutch) and involved participants who were not L1 French-speakers.

Interestingly, CLIL and NON-CLIL learners did not differ in their use of communicative phrasemes (*speech act formulae* and *attitudinal formulae* for TL English and only the former for TL Dutch). This could be explained by the fact that it is possible and relatively easy to teach routine formulae in the foreign language classroom and that higher proficiency does not seem to be a real advantage in this case (Roever, 2011). Some phraseological errors (e.g. the overuse of articles or the use of non-existing words) might be a result of a negative cross-linguistic influence from their mother tongue (French). This is not surprising since research on the phraseological lexicon of TL learners recognized the L1 influence as one of the most frequent sources of errors (see Chapter 1, Sections 1.5.3. and 1.7.2.).

Given that both types of participants were quite homogeneous in terms of their general characteristics and linguistic background, we may assume that the higher quantity of input provided via the CLIL approach is beneficial for the processing of phraseological units. However, we saw that CLIL learners were having significantly more informal contact with the TL outside school, which may also account for the increased phraseological competence. In fact, various studies (Schmitt & Redwood, 2011 on phrasal verbs; Lin, 2014 on formulaic sequences; Fernández & Schmitt, 2015 on collocations) showed a positive influence of extracurricular activities such as extensive reading, watching movies/TV, internet use or social networking on the acquisition of phraseological units. Van Mensel, Bulon, Hendrikx, Meunier, and Van Goethem (submitted) also demonstrated that informal input actually seems to be a better predictor of learners' writing skills than formal input, at least for Dutch as a TL.

5.2.4. Conclusion

All in all, this study demonstrated a difference in the phraseological competence between the two groups of learners, with the CLIL learners of English or Dutch displaying a better phraseological competence.

The pilot study looking at the learners' global written competence (Chapter 4, Section 4.2.) showed a greater difference between the CLIL and the NON-CLIL learners in TL Dutch than in TL English. This also appeared to be the case in the present study: the CLIL learners of Dutch significantly produced more phraseological units in a larger number of (sub-)categories of phrasemes and only made more phraseological errors in one category.

In the following study, we aim to investigate the evolution of the phraseological competence of both groups of learners thanks to the collection of longitudinal data. The current study has revealed traces of emergence of various phraseological constructions in CLIL as well as in

NON-CLIL productions, with a more advanced level of acquisition in the CLIL group (higher frequency, variety, and accuracy overall). Hence, it would be interesting to see how both groups have been evolving towards acquisition (if any evolution has occurred at all).

5.3. Longitudinal development of the CLIL and NON-CLIL learners' phraseological language over time

The present study examines the development of the phraseological language of CLIL and NON-CLIL learners of English or Dutch as a TL using a longitudinal corpus of written productions. This study contributes to the body of existing research on the TL development of CLIL learners compared to NON-CLIL learners, especially in the field of phraseology since very few studies have tackled that area. There also seems to be a need for more longitudinal studies using larger samples as these allow for the identification of patterns in language development with a high degree of representativeness of the data (Paquot & Granger, 2012). In fact, most of the studies looking at language development are pseudo-longitudinal or quasi-longitudinal, as discussed in Chapter 1, Section 1.5.1.3.

Section 5.3.1. introduces the research questions and hypotheses. Section 5.3.2. provides information about the participants per TL, the study design, the collected data, and the statistical analyses. Section 5.3.3. describes the results per type of analysis and per TL. Section 5.3.4. summarizes the results and discusses them.

5.3.1. Research questions and hypotheses

The aim of the study is to explore the development of phraseological language of the same CLIL and NON-CLIL learners over two consecutive school years.

The research questions are as follows:

1. How does the phraseological language of CLIL and NON-CLIL learners develop longitudinally in terms of frequency of use (RQ1)? ;
2. How does the phraseological language of CLIL and NON-CLIL learners develop longitudinally in terms of accuracy (RQ2)? ;
3. To what extent does the phraseological language of CLIL and NON-CLIL learners differ in their longitudinal development in terms of frequency of use (RQ3)? ;
4. To what extent does the phraseological language of CLIL and NON-CLIL learners differ in their longitudinal development in terms of accuracy (RQ4)?

It is believed that time spent learning a TL and proficiency in that language are positively correlated (Meunier, 2015). Therefore, we expect CLIL and NON-CLIL learners to perform better at Time 2. As for RQ3 and RQ4, we expect the CLIL learners' TL phraseological language to improve more than the one of NON-CLIL learners since they benefit from a larger amount of more authentic input in the TL (see Chapter 3).

5.3.2. Methodology

5.3.2.1. Participants³⁷

English

The participants were 339 French-speaking secondary school learners of English as a TL in a CLIL (78) or a NON-CLIL (66) context. Tables 5.29. and 5.30. describe the participants in terms of background characteristics and (extra)curricular exposure to English. The sample sizes (*n*) in the tables do not always correspond to 144 because some data were missing.

Table 5.29. Background characteristics for the learners of English

	<i>n</i> =144	Gender (%)		Home language (%)		
		<i>n</i> =144		<i>n</i> =143		
		Male	Female	French	French + other	Mostly other
CLIL	78	46.2	53.8	62.8	23.1	14.1
NON-CLIL	66	31.8	68.2	72.3	26.2	1.5
ALL	144	39.6	60.4	67.1	24.5	8.4

87 (60.4%) of the participants were female, while 57 (39.6%) of them were male. The majority of the participants (67.1%) spoke only French at home. Some of them spoke sometimes French, sometimes another language (24.5%). A minority mostly spoke another language (8.4%), and this mostly concerns CLIL learners (14.1%). Among those who answered speaking another language than French at home, only five mentioned English as the 'other language' – the others did not answer or reported a language that was not relevant for this study. They were not excluded from the analysis.

Table 5.30. Learners' (extra)curricular exposure to English

<i>n</i>	Nb of hours/week at school	Extra classes	Stay abroad	Informal contact

³⁷ Data about the participants' home language and linguistic background were collected at Time 1 only, except for the variable regarding the number of hours of TL per week the participants received at school. However, this variable remained unchanged between Grade 11 and 12.

		(%) (n=143)		(%) (n=143)	(%) (n=143)	outside school (Means) (n=142)
		4h/week	+ 4h/week			
CLIL	78	75.6	24.4	3.8	10.3	3.45
NON-CLIL	66	72.3	27.7	0	1.5	2.89
ALL	144	74.1	25.9	2.1	6.3	3.20

In terms of the participants' linguistic background, a large number of learners received four hours of English as a TL per week at school (74.1%), while some of them received more than 4h/week (24.4% in CLIL and 27.7% in NON-CLIL). Only 3.8% of CLIL learners took extra classes of English outside school and only a minority of participants have stayed in an English-speaking country or region for more than three months (10.3% of CLIL learners and 1.5% of NON-CLIL learners). An independent samples t-test showed that CLIL learners (Mean = 3.45, SD = 0.83) had significantly more current informal contact with the TL outside school than NON-CLIL learners (Mean = 2.89, SD = 0.77), $t(140) = 4.127$, $p = .000$, $d = .70$ (medium to large effect size).

Dutch

The participants were 195 French-speaking secondary school learners of Dutch as a TL in a CLIL (123) or a NON-CLIL (72) context. Tables 5.31. and 5.32. give a description of the participants in terms of background characteristics and (extra)curricular exposure to Dutch. The samples sizes (n) in the tables do not always correspond to 195 because some data were missing.

Table 5.31. Background characteristics for the learners of Dutch

	$n=195$	Gender (%) $n=195$		Home language (%) $n=194$		
		Male	Female	French	French + other	Mostly other
CLIL	123	54.5	45.5	83.6	15.6	0.8
NON-CLIL	72	45.8	54.2	70.8	26.4	2.8
ALL	195	51.3	48.7	78.9	19.6	1.5

95 (48.7%) of the participants were female and 100 (51.3%) were male. French was the sole language spoken at home for the majority of the participants (78.9%). 19.6% sometimes spoke

another language than French – among those who specified that ‘other language’, seven of them answered Dutch – and 1.5% mostly spoke another language (not specified).

Table 5.32. Learners’ (extra)curricular exposure to Dutch

	<i>n</i>	Nb of hours/week at		Extra	Stay	Informal
		school		classes	abroad	contact
		(%)		(%)	(%)	outside
		(n=194)		(n=194)	(n=193)	school
						(Means)
						(n=188)
		4h/week	+ 4h/week			
CLIL	123	87.7	12.3	0.8	8.3	2.52
NON-CLIL	72	100	0	8.3	0	2.13
ALL	195	92.3	7.7	3.6	5.2	2.38

Regarding the number of hours of Dutch as a TL given at school, most of the participants (92.3%) received 4h/week – a number of CLIL learners received more than 4h/week. A small minority of learners took extra classes of Dutch outside school (0.8% in CLIL and 8.3% in NON-CLIL) or have stayed in a Dutch-speaking country or region for at least three months (5.2%, only CLIL learners). An independent samples t-test revealed that CLIL learners (Mean = 2.52, SD = 0.78) had significantly more current informal contact with Dutch outside school than NON-CLIL learners (Mean = 2.13, SD = 0.71), $t(186) = 3.392$, $p=.001$, $d=.52$ (medium effect size).

5.3.2.2. Design and data

The study is longitudinal; the data were gathered across two consecutive school years (2015-2016 and 2016-2017). Informal written texts from both CLIL and NON-CLIL learners were collected at two different time points. The first data collection (Time 1) took place between October 19th and November 9th 2015 – the learners were at the beginning of the fifth year of their secondary school education (Grade 11). The second and last data collection (Time 2) was conducted between April 20th and May 15th 2017 when the learners were about to finish their sixth year of secondary school education (Grade 12).

English

The longitudinal TL English corpus consisted of 288 written productions in the form of informal e-mails to a friend. These were collected through the writing task as described in Chapter 4, Section 4.5.1.

Table 5.33. Description of the English corpus

	Time 1				Time 2			
	Number of texts	Number of words	Mean text length	SD	Number of texts	Number of words	Mean text length	SD
CLIL	78	25,537	327.40	65.39	78	24,447	313.42	71.43
NON- CLIL	66	17,341	262.74	82.46	66	17,284	261.88	54.80
TOTAL	144	42,878	297.76	80.24	144	41,731	289.80	69.11

Dutch

The longitudinal TL Dutch corpus was made up of 390 written productions that were gathered in the exact same way.

Table 5.34. Description of the Dutch corpus

	Time 1				Time 2			
	Number of texts	Number of words	Mean text length	SD	Number of texts	Number of words	Mean text length	SD
CLIL	123	34,786	282.89	58.25	123	33,495	272.32	50.28
NON- CLIL	72	14,031	194.57	62.96	72	15,535	215.76	47.53
TOTAL	195	48,817	250.28	73.45	195	49,030	251.43	56.26

Tables 5.33. and 5.34. above show that the corpora from Time 1 are quite similar to the corpora from Time 2 in terms of number of words and mean text length. However, the CLIL and NON-CLIL corpora substantially differ in the length of their texts at both data points and this has been taken into account in the analyses of phraseological accuracy.

The process of extracting and classifying the phraseological units and errors is the same as the one used in the previous cross-sectional study and its description can be found in Chapter 4, Section 4.6.

5.3.2.3. Longitudinal analyses

Researchers investigating the linguistic development of CLIL and NON-CLIL learners over time appear to differ in the way they analyze the impact of CLIL. Juan-Garau, Salazar-Noguera, and Gené-Gil (2013) conducted paired-samples t-tests for each learner group separately in order to measure their writing development between Time 1 and Time 2, and then compared the learners' writing performance at Time 2. Gené-Gil, Juan-Garau, and Salazar-Noguera (2015) carried out an analysis of variance (ANOVA) in order to uncover any significant development in the writing of CLIL and NON-CLIL learners across four different data collection points. These intra-group analyses allowed them to look at the two groups independently and measure their respective progress over time. Roquet i Pugès (2011) performed inter-group analyses using ANOVA tests in order to measure the differential linguistic gains made by CLIL and NON-CLIL learners so that we know whether the groups significantly differed in their evolution. The average performances of the two groups were presented graphically, followed by an interaction plot for each linguistic variable measured.

Both types of analyses were conducted in the present study since they seem relevant and complementary when exploring the effect of CLIL education on linguistic development. Paired-samples t-tests are first used to uncover any significant difference between Time 1 and Time 2 for both groups of learners independently (intra-group analysis). Repeated measures ANOVAs with the educational setting (CLIL/NON-CLIL) used as a between-subjects factor are then conducted to see whether the difference between the CLIL group's progress and the NON-CLIL group's progress is statistically significant (inter-group analysis).

5.3.3. Results

The results are presented in four different sections. The first two sections (5.3.3.1. and 5.3.3.2.) describe the findings of the intra-group analysis, while the last two sections (5.3.3.3. and 5.3.3.4.) focus on the results of the inter-group analysis. Each type of analysis looks at the learners' phraseological performance, first in terms of frequency of use and then in terms of accuracy.

5.3.3.1. Intra-group analysis – frequency of use (RQ1)

Before turning to the longitudinal findings, it seems essential to remember that comparative analyses that were carried out at Time 1 only (see previous study in this chapter) revealed a significantly higher number of phraseological units in the CLIL learners' productions compared to the NON-CLIL learners' ones. Significant differences in favor of the CLIL group were also

found in the majority of the various sub-categories. This holds for English and Dutch as target languages.

English

The means and standard deviations for the total number of words, the total number of phraseological units, the three broad categories of phrasemes and their various sub-categories are shown in Table 5.35. for the CLIL group and in Table 5.36. for the NON-CLIL group.

Table 5.35. Change in scores over time for the CLIL English group (absolute frequencies)

	Time 1		Time 2		
	Mean	SD	Mean	SD	
Number of words	327.40	65.39	313.42	71.43	
Number of phrasemes	31.38	9.07	33.85	8.65	*
Referential phrasemes	27.92	8.42	30.22	8.09	*
Textual phrasemes	1.04	1.26	1.28	1.18	
Communicative phrasemes	2.42	2.21	2.35	1.90	
Lexical collocations	10.37	4.66	10.58	4.85	
Compounds	7.10	3.69	6.24	3.19	
Grammatical collocations	9.72	4.51	12.90	5.19	***
Phrasal verbs	0.54	0.93	0.42	0.67	
Complex prepositions	0.36	0.68	0.49	0.64	
Complex conjunctions	0.26	0.52	0.35	0.62	
Linking adverbials	0.40	0.81	0.41	0.65	
Speech act formulae	1.53	1.28	1.40	1.32	
Attitudinal formulae	0.90	1.70	0.95	1.13	

Notes. * $p < .05$. *** $p < .001$

A statistically significant increase is found in the categories *number of phrasemes* ($t(77) = -2.071$, $p = .042$, $d = .28$), *referential phrasemes* ($t(77) = -2.155$, $p = .034$, $d = .28$), and *grammatical collocations* ($t(77) = -4.969$, $p = .000$, $d = .65$).

Table 5.36. Change in scores over time for the NON-CLIL English group (absolute frequencies)

	Time 1		Time 2	
	Mean	SD	Mean	SD

Number of words	262.74	82.46	261.88	54.80	
Number of phrasemes	22.18	8.90	28.79	8.01	***
Referential phrasemes	19.85	8.12	26.00	7.44	***
Textual phrasemes	0.52	0.83	0.47	0.75	
Communicative phrasemes	1.82	1.55	2.26	1.68	
Lexical collocations	8.97	4.51	10.94	4.00	**
Compounds	4.36	3.25	5.09	2.79	
Grammatical collocations	6.42	3.78	9.94	3.94	***
Phrasal verbs	0.05	0.21	0.06	0.30	
Complex prepositions	0.09	0.29	0.11	0.31	
Complex conjunctions	0.08	0.32	0.12	0.54	
Linking adverbials	0.35	0.67	0.21	0.41	
Speech act formulae	1.35	1.18	1.70	1.41	
Attitudinal formulae	0.47	0.83	0.56	0.79	

Notes. ** $p < .01$. *** $p < .001$

A statistically significant increase is found in the categories *number of phrasemes* ($t(65) = -5.320$, $p=.000$, $d=.78$), *referential phrasemes* ($t(65) = -5.432$, $p=.000$, $d=.79$), *lexical collocations* ($t(65) = -2.829$, $p=.006$, $d=.46$), and *grammatical collocations* ($t(65) = -6.490$, $p=.000$, $d=.91$).

The results from the intra-group analysis in terms of frequency of use show improvement in the production of phraseological units in both learner groups, even though they were less productive in terms of number of words. The NON-CLIL learners increased their production of phraseological units in most of the (sub-)categories (11 out of 14). When statistically analyzed, significant differences are observed in four categories and the effect sizes go from medium to large. The CLIL learners expanded their range of phraseological units in more than half of the categories measured (9 out of 14). A statistically significant difference is found in three categories and the effect sizes go from small to medium. Interestingly, the (sub-)categories that significantly differ between Time 1 and Time 2 are the same for the two groups of learners: *total number of phraseological units*, *referential phrasemes* and *grammatical collocations*.

Dutch

The means and standard deviations for the total number of words, the total number of phraseological units, the three broad categories of phrasemes and their various sub-categories are shown in Table 5.37. for the CLIL group and in Table 5.38. for the NON-CLIL group.

Table 5.37. Change in scores over time for the CLIL Dutch group (absolute frequencies)

	Time 1		Time 2		
	Mean	SD	Mean	SD	
Number of words	282.81	58.25	272.32	50.28	
Number of phrasemes	31.77	8.92	30.28	8.28	
Referential phrasemes	26.22	8.29	24.11	7.00	**
Textual phrasemes	2.91	2.09	3.33	2.33	
Communicative phrasemes	2.64	1.53	2.85	1.71	
Lexical collocations	14.04	5.76	11.37	4.39	****
Compounds	3.17	2.25	3.12	2.38	
Grammatical collocations	7.11	3.57	7.72	3.49	
Phrasal verbs	1.86	1.44	1.90	1.57	
Complex prepositions	0.13	0.34	0.24	0.48	
Complex conjunctions	2.33	1.91	2.54	1.98	
Linking adverbials	0.36	0.77	0.46	0.81	
Speech act formulae	1.46	1.22	1.53	1.18	
Attitudinal formulae	1.19	1.11	1.30	1.12	

Notes. ** $p < .01$. *** $p < .001$

A statistically significant decrease is found in the categories *referential phrasemes* ($t(122) = 2.640$, $p = .009$, $d = .27$) and *lexical collocations* ($t(122) = 4.744$, $p = .000$, $d = .52$) from Time 1 to Time 2.

Table 5.38. Change in scores over time for the NON-CLIL Dutch group (absolute frequencies)

	Time 1		Time 2		
	Mean	SD	Mean	SD	
Number of words	194.88	62.96	215.76	47.53	**
Number of phrasemes	22.08	8.04	24.64	7.68	*
Referential phrasemes	18.69	7.29	20.12	7.25	
Textual phrasemes	1.26	1.60	1.82	1.71	*
Communicative phrasemes	2.12	1.85	2.69	1.95	*

Lexical collocations	10.50	4.26	10.82	4.43	
Compounds	1.86	1.68	2.00	1.93	
Grammatical collocations	5.44	3.56	6.57	3.54	*
Phrasal verbs	0.85	1.08	0.74	1.20	
Complex prepositions	0.04	0.20	0.15	0.40	*
Complex conjunctions	1.04	1.39	1.29	1.32	
Linking adverbials	0.17	0.50	0.37	0.86	
Speech act formulae	1.44	1.46	1.47	1.30	
Attitudinal formulae	0.68	1.11	1.22	1.33	**

Notes. * $p < .05$. ** $p < .01$.

A statistically significant increase is found in the categories *number of words* ($t(71) = -3.008$, $p = .004$, $d = .37$), *number of phrasemes* ($t(71) = -2.214$, $p = .030$, $d = .32$), *textual phrasemes* ($t(71) = -2.066$, $p = .043$, $d = .34$), *communicative phrasemes* ($t(71) = -2.032$, $p = .046$, $d = .30$), *grammatical collocations* ($t(71) = -2.314$, $p = .024$, $d = .32$), *complex prepositions* ($t(71) = -2.194$, $p = .032$, $d = .35$), and *attitudinal formulae* ($t(71) = -2.712$, $p = .008$, $d = .44$).

The results from the intra-group analysis in terms of frequency of use show a clear improvement in the production of phraseological units in the NON-CLIL group. In fact, NON-CLIL learners produced more phraseological units in all categories and sub-categories, except one (*phrasal verbs*). They also wrote longer productions. When statistically analyzed, significant differences are observed in half of the categories (7 out of 14). The effect sizes go from small to medium. Improvement is less visible in the CLIL group; we notice an increase in phraseological production for 9 (sub-)categories out of 14 and a decrease for the remaining 5 other categories (including *text length* and *total number of phraseological units*). More importantly, a statistically significant decrease is found in two categories of phrasemes, namely *referential phrasemes* and *lexical collocations* (effect sizes being small or medium).

5.3.3.2. Intra-group analysis – accuracy (RQ2)

The comparative analyses of the learners' phraseological production at Time 1 (see cross-sectional study in this chapter) showed that accuracy tended to be better in the CLIL group (significantly less phraseological errors produced overall). This was true for TL English and TL Dutch.

Potential changes in the accuracy of the CLIL and the NON-CLIL learners' phraseological productions between the two time points are first represented graphically. Statistical tests are then conducted for the most frequent categories of phraseological errors. Both the absolute and

relative frequencies (per 300 words for TL English and per 250 words for TL Dutch) are taken into account since the learner groups differ in text length.

English

Figure 5.7. shows the absolute frequencies of phraseological errors produced in the CLIL group at Time 1 and Time 2. We see that the accuracy of the CLIL learners' phraseological lexicon tends to decrease over time. In fact, the frequencies of more than half of the categories of phraseological errors (10 out of 18) seem to increase (more orange, less blue). A drop in the number of errors is seen in only five categories of errors (more blue, less orange), while three categories show identical frequencies.

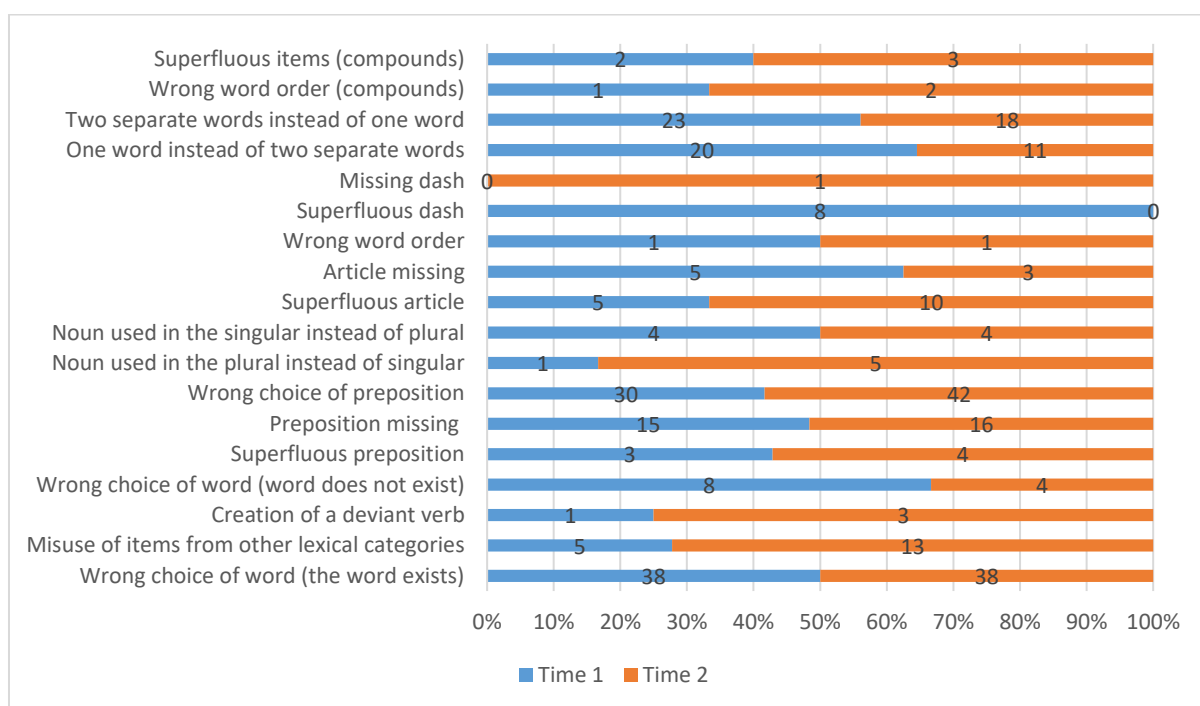


Figure 5.7. Evolution in the number of phraseological errors over time per category for the CLIL English group (absolute frequencies)

Figure 5.8. shows the relative frequencies of errors for the CLIL group and displays a similar picture than Figure 5.7. 12 categories out of 18 show an increase in erroneous phraseological units (more orange, less blue), while six categories see their frequencies decrease over time (more blue, less orange).

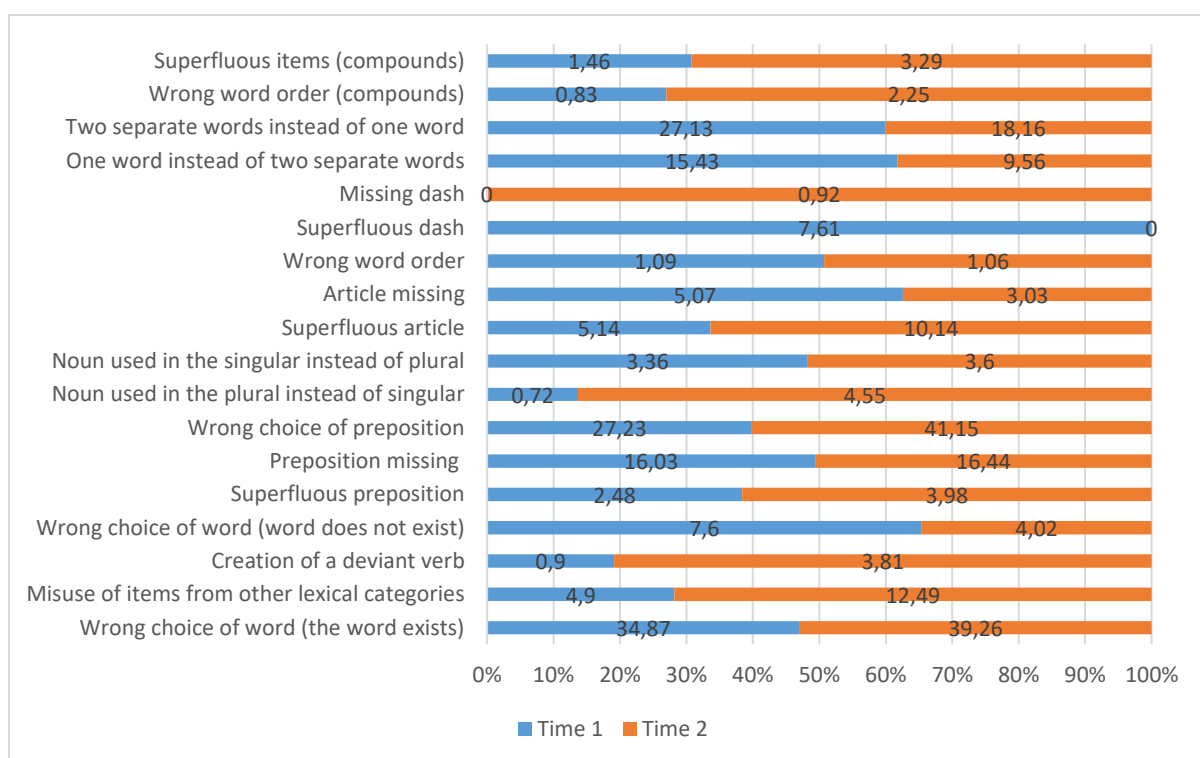


Figure 5.8. Evolution in the number of phraseological errors over time per category for the CLIL English group (relative frequencies per 300 words)

Table 5.39. compares the absolute frequencies (and relative frequencies) between the two data collections for the total number of phraseological errors and the most frequent categories of errors.

Table 5.39. Change in scores over time for the CLIL English group

	Time 1		Time 2	
	Mean	SD	Mean	SD
Number of phraseological errors	2.18 (2.07)	2.00 (1.98)	2.28 (2.28)	2.11 (2.17)
Wrong choice of existing word	0.49 (0.48)	0.86 (0.76)	0.49 (0.50)	0.78 (0.85)
Wrong choice of preposition	0.38 (0.35)	0.63 (0.56)	0.54 (0.53)	0.83 (0.88)
Two separate words instead of one word	0.29 (0.35)	0.84 (1.17)	0.23 (0.23)	0.72 (0.81)
Preposition missing	0.19 (0.20)	0.48 (0.51)	0.21 (0.21)	0.49 (0.53)

One word instead of two separate words	0.26 (0.20)	0.84 (0.60)	0.14 (0.12)	0.53 (0.44)	
Misuse of items from other lexical categories	0.06 (0.06)	0.25 (0.24)	0.17 (0.16)	0.41 (0.39)	
Superfluous article	0.06 (0.06)	0.25 (0.26)	0.13 (0.13)	0.41 (0.47)	
Wrong choice of word (does not exist)	0.10 (0.10)	0.34 (0.33)	0.05 (0.05)	0.27 (0.30)	
Superfluous dash	0.10 (0.10)	0.30 (0.30)	0 (0)	0 (0)	** **

Notes. ** $p < .01$.

A statistically significant decrease in the number of phraseological errors is found for the category *superfluous dash* (absolute frequencies: $t(77) = 2.966$, $p = .004$, $d = .47$) and relative frequencies: $t(77) = 2.920$, $p = .005$, $d = .47$).

Figure 5.9. shows the absolute frequencies at Time 1 and Time 2 for the NON-CLIL group. The accuracy of the NON-CLIL learners' phraseological lexicon seems to improve for eight categories of errors (out of 18) (more blue, less orange) and to decrease for 10 of them (more orange, less blue).

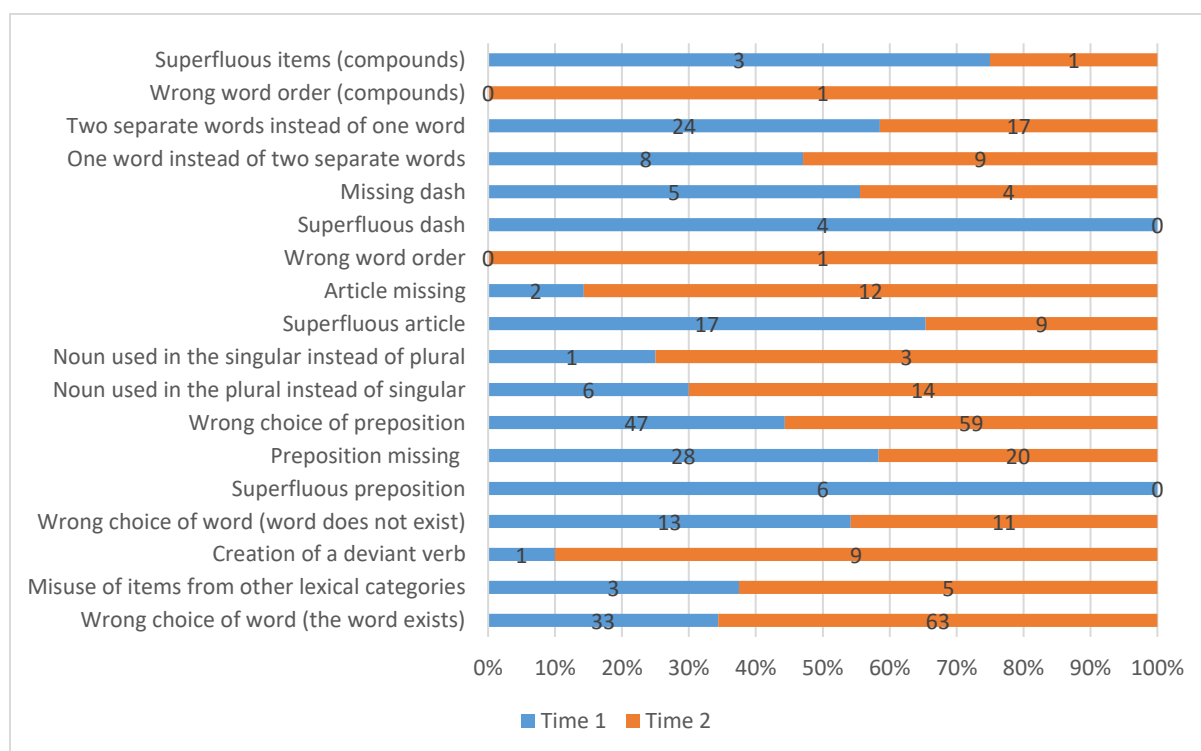


Figure 5.9. Evolution in the number of phraseological errors per category over time for the NON-CLIL English group (absolute frequencies)

Figure 5.10. shows the relative frequencies at Time 1 and Time 2 for the NON-CLIL group and displays the same results as Figure 5.9.

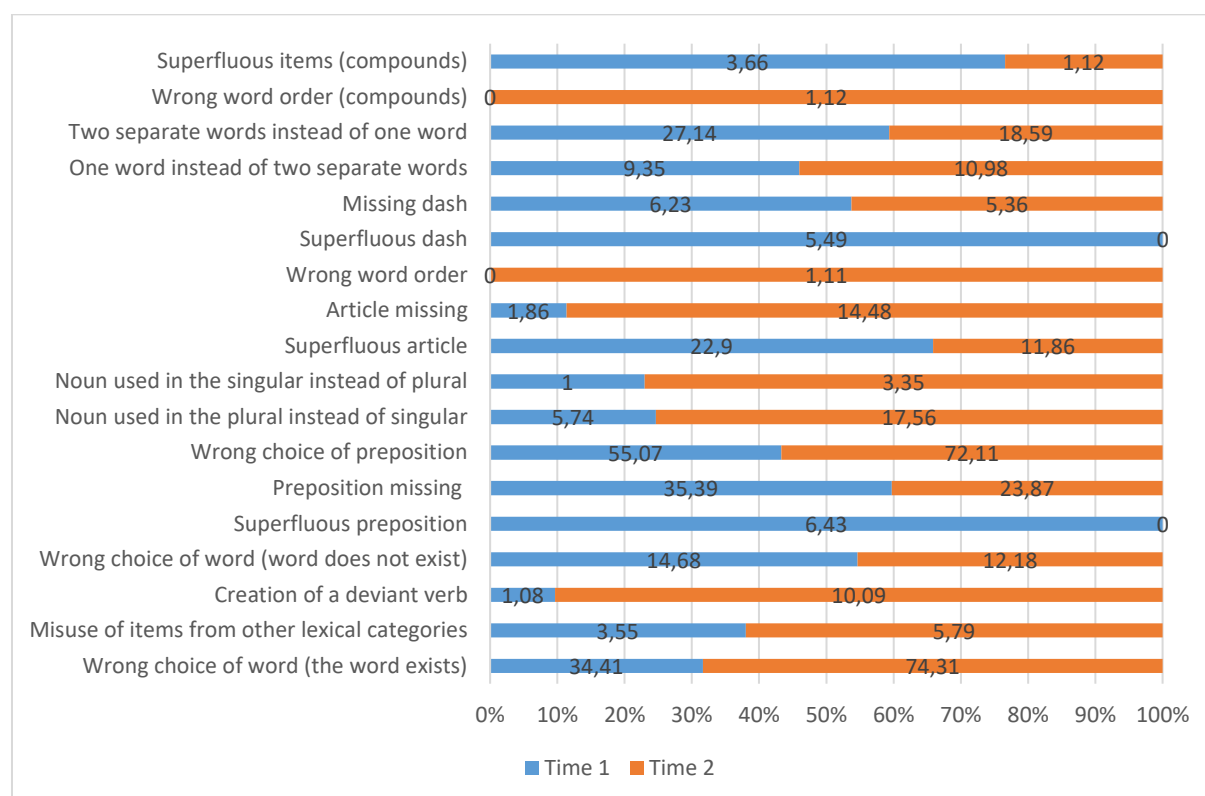


Figure 5.10. Evolution in the number of phraseological errors over time per category for the NON-CLIL English group (relative frequencies per 300 words)

A comparison of the absolute frequencies (and relative frequencies) between Time 1 and Time 2 in the total number of phraseological errors and the most frequent categories of errors are shown in Table 5.40. below.

Table 5.40. Change in scores over time for the NON-CLIL English group

	Time 1		Time 2	
	Mean	SD	Mean	SD
Number of phraseological errors	3.05 (3.56)	2.54 (2.78)	3.61 (4.30)	2.69 (3.17)
Wrong choice of preposition	0.71 (0.83)	1.12 (1.29)	0.89 (1.09)	1.08 (1.32)

Wrong choice of existing word	0.50 (0.52)	0.77 (0.76)	0.95 (1.12)	1.23 (1.50)	* **
Preposition missing	0.42 (0.54)	0.88 (1.20)	0.30 (0.36)	0.61 (0.69)	
Two separate words instead of one word	0.36 (0.41)	0.69 (0.77)	0.26 (0.28)	0.66 (0.72)	
Superfluous article	0.26 (0.35)	0.53 (0.76)	0.14 (0.18)	0.35 (0.48)	
Wrong choice of word (does not exist)	0.20 (0.22)	0.47 (0.56)	0.17 (0.18)	0.45 (0.49)	
Noun used in the plural instead of singular	0.09 (0.09)	0.34 (0.32)	0.21 (0.27)	0.54 (0.71)	
Article missing	0.03 (0.03)	0.17 (0.16)	0.18 (0.22)	0.52 (0.63)	* *
Creation of a deviant verb	0.02 (0.02)	0.12 (0.13)	0.14 (0.15)	0.46 (0.54)	*
Superfluous preposition	0.09 (0.10)	0.34 (0.39)	0 (0)	0 (0)	* *

A statistically significant increase in the number of phraseological errors is found in the categories *wrong choice of existing word* (absolute frequencies: $t(65) = -2.495$, $p=.015$, $d=.44$ and relative frequencies: $(t(65) = -2.885$, $p=.005$, $d=.50)$, *article missing* (absolute frequencies: $(t(65) = -2.193$, $p=.032$, $d=.95$ and relative frequencies: $(t(65) = -2.358$, $p=.021$, $d=.41)$, and *creation of a deviant verb*, but for the absolute frequencies only ($t(65) = -2.048$, $p=.045$, $d=.36$ (the p-value for the relative frequencies was almost significant: $p=.051$). However, the number of errors significantly decreased for the category *superfluous preposition* (absolute frequencies: $t(65) = 2.181$, $p=.033$, $d=.37$ and relative frequencies: $t(65) = 2.029$, $p=.047$, $d=.36$).

The overall trends show a decrease in the accuracy of the learners' phraseological language in the CLIL and the NON-CLIL group. In fact, the total number of phraseological errors tends to increase in both groups (although not significantly) and we find an increase in the frequency of errors for more than half of the categories measured (see Figures 5.7. to 5.10.). However, when we look at the significant results, we see that the CLIL group improved over time (significant decrease of errors in one category with a medium effect size) while the NON-CLIL group

became less accurate (significant increase of errors in three categories with effect sizes going from medium to large and significant decrease in one category with a small to medium effect size).

Dutch

Figure 5.11. shows the absolute frequencies of phraseological errors produced in the CLIL group at Time 1 and Time 2. We see that the accuracy of the CLIL learners' phraseological lexicon tends to improve over time. In fact, the frequencies of more than half of the categories of phraseological errors (14 out of 24) seem to drop (more blue, less orange). An increase in the number of errors is seen in only eight categories of errors (more orange, less blue), while two categories show identical frequencies.

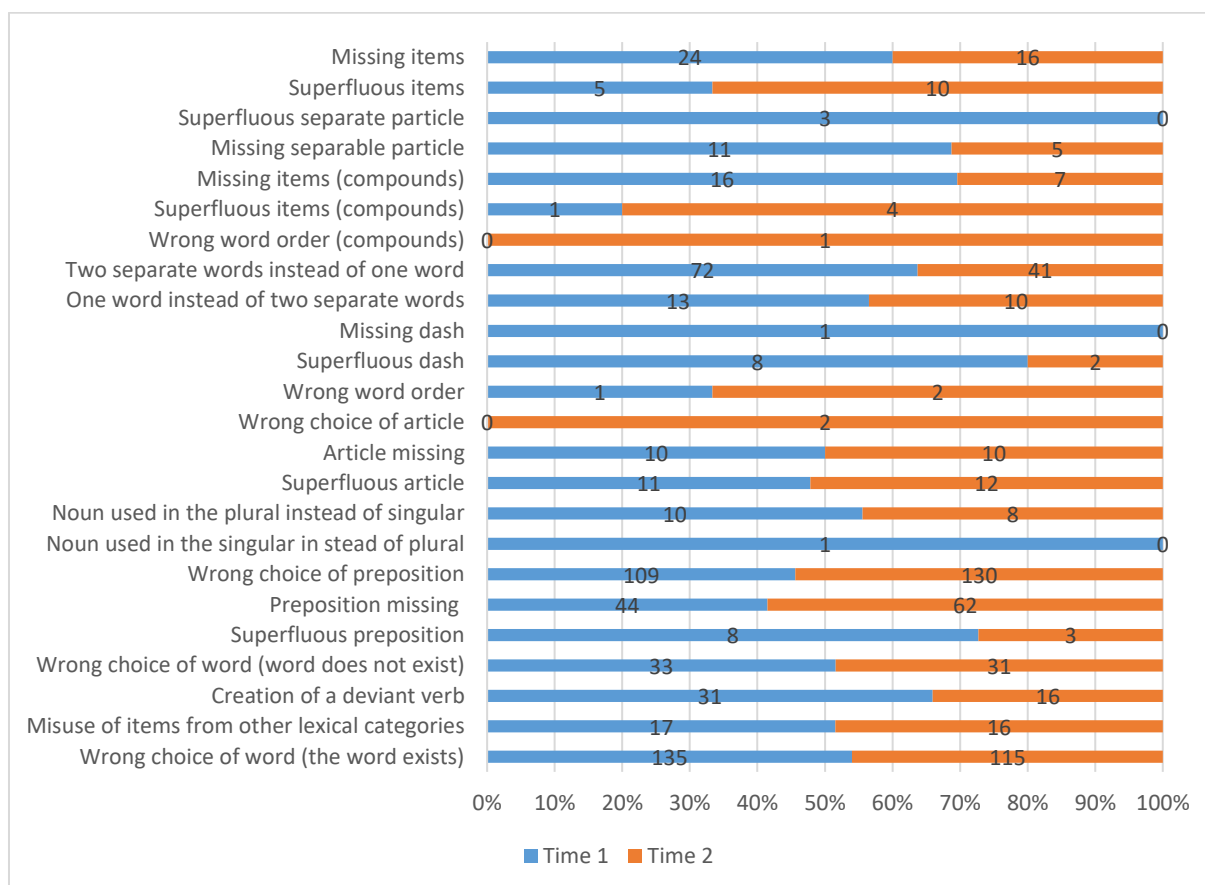


Figure 5.11. Evolution in the number of phraseological errors over time per category for the CLIL Dutch group (absolute frequencies)

Figure 5.12. shows the relative frequencies of errors for the CLIL group and displays a similar picture than Figure 5.11.: 14 categories of phraseological errors (out of 24) show a decrease (more blue, less orange), while 10 categories see their frequencies increase over time (more orange, less blue).

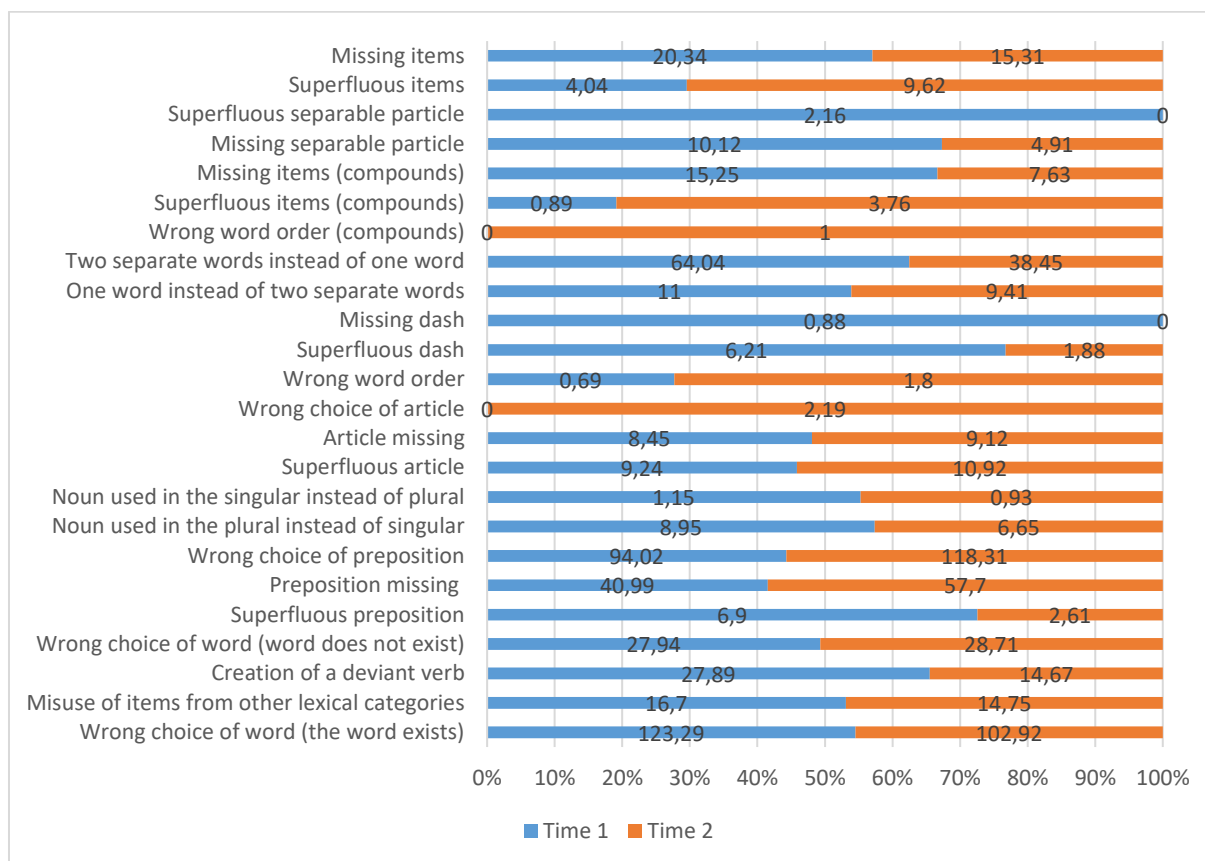


Figure 5.12. Evolution in the number of phraseological errors over time per category for the CLIL Dutch group (relative frequencies per 250 words)

Table 5.41. compares the absolute frequencies (and relative frequencies) between the two data collections for the total number of phraseological errors and the most frequent categories of errors.

Table 5.41. Change in scores over time for the CLIL Dutch group

	Time 1		Time 2	
	Mean	SD	Mean	SD
Number of phraseological errors	4.59 (4.07)	2.63 (2.24)	4.10 (3.77)	2.53 (2.30)
Wrong choice of existing word	1.10 (1.00)	1.11 (1.05)	0.93 (0.84)	1.20 (1.07)
Wrong choice of preposition	0.89 (0.76)	1.07 (0.88)	1.06 (0.96)	1.10 (1.04)

Two separate words instead of one word	0.59 (0.52)	0.86 (0.80)	0.33 (0.31)	0.64 (0.61)	** **
Preposition missing	0.36 (0.33)	0.67 (0.62)	0.50 (0.47)	0.69 (0.64)	
Wrong choice of word (does not exist)	0.27 (0.23)	0.53 (0.45)	0.25 (0.23)	0.52 (0.49)	
Creation of a deviant verb	0.25 (0.23)	0.57 (0.52)	0.13 (0.12)	0.38 (0.37)	* *
Missing items	0.20 (0.16)	0.47 (0.41)	0.13 (0.12)	0.42 (0.41)	
Misuse of item from other lexical categories	0.14 (0.13)	0.37 (0.37)	0.13 (0.12)	0.36 (0.33)	
Missing items (compounds)	0.13 (0.12)	0.44 (0.43)	0.06 (0.06)	0.27 (0.30)	
One word instead of two separate words	0.11 (0.09)	0.33 (0.28)	0.08 (0.08)	0.30 (0.29)	
Superfluous article	0.09 (0.07)	0.29 (0.24)	0.10 (0.09)	0.35 (0.32)	
Article missing	0.80 (0.69)	0.27 (0.23)	0.80 (0.74)	0.30 (0.27)	
Missing separable particle	0.09 (0.08)	0.36 (0.36)	0.04 (0.04)	0.20 (0.20)	

A statistically significant decrease in the number of phraseological errors is found for the categories *two separate words instead of one word* (absolute frequencies: $t(122) = 3.039$, $p=.003$, $d=.34$ and relative frequencies: $t(122) = 2.745$, $p=.007$, $d=.29$) and *creation of a deviant verb* (absolute frequencies: $t(122) = 2.130$, $p=.035$, $d=.25$ and relative frequencies: $t(122) = 2.072$, $p=.040$, $d=.24$).

Figure 5.13. shows the absolute frequencies at Time 1 and Time 2 for the NON-CLIL group. The accuracy of the NON-CLIL learners' phraseological lexicon seems to improve for 10

categories of errors (out of 22³⁸) (more blue, less orange) and to decrease for nine of them (more orange, less blue). The frequencies of three categories do not appear to change over time.

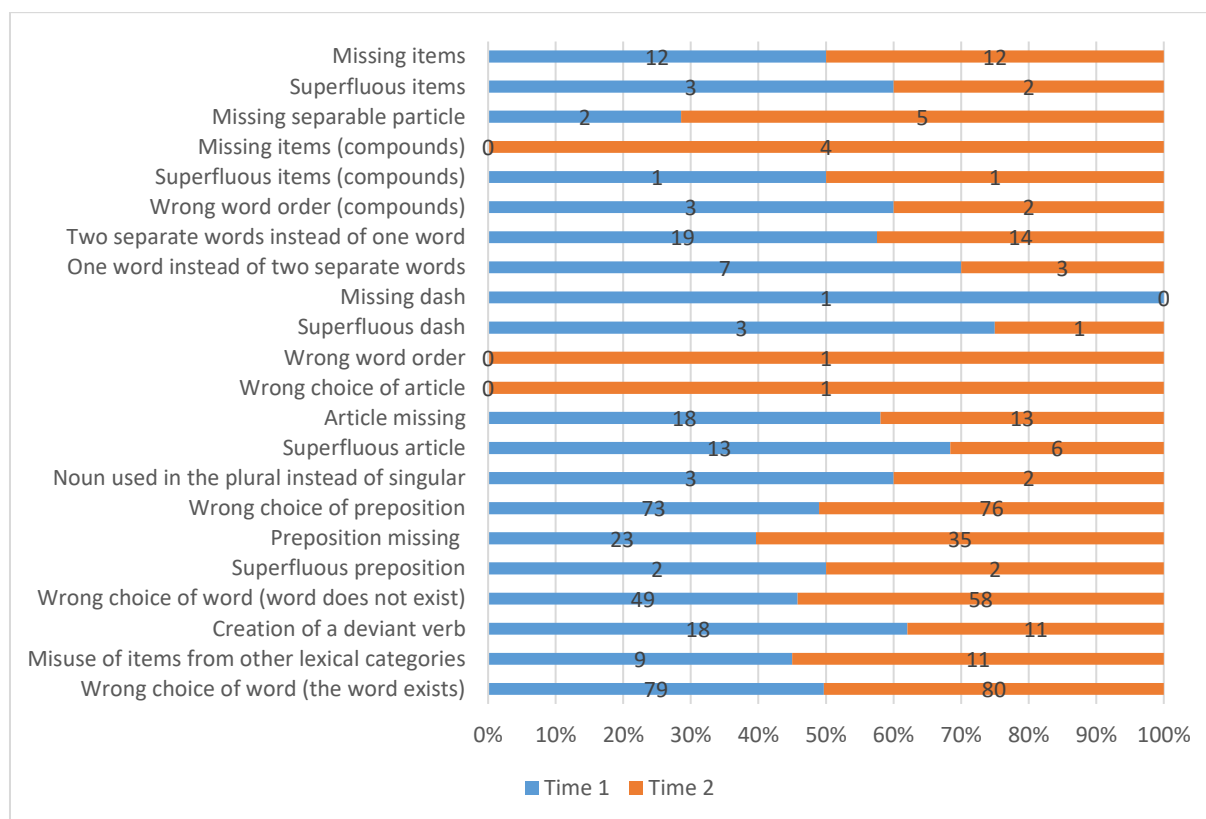


Figure 5.13. Evolution in the number of phraseological errors per category over time for the NON-CLIL Dutch group (absolute frequencies)

Figure 5.14. shows the relative frequencies at Time 1 and Time 2 for the NON-CLIL group. Once again, we can say that Figure 5.13. and Figure 5.14. display similar results. The slight changes lie in the fact that small differences now exist for the three categories whose absolute frequencies were identical, which results in the following observations: a decrease in 12 categories (out of 22) (more blue, less orange) and an increase in 10 of them (more orange, less blue).

³⁸ No phraseological errors were found for the categories *superfluous separable particle* and *noun used in the singular instead of plural*.

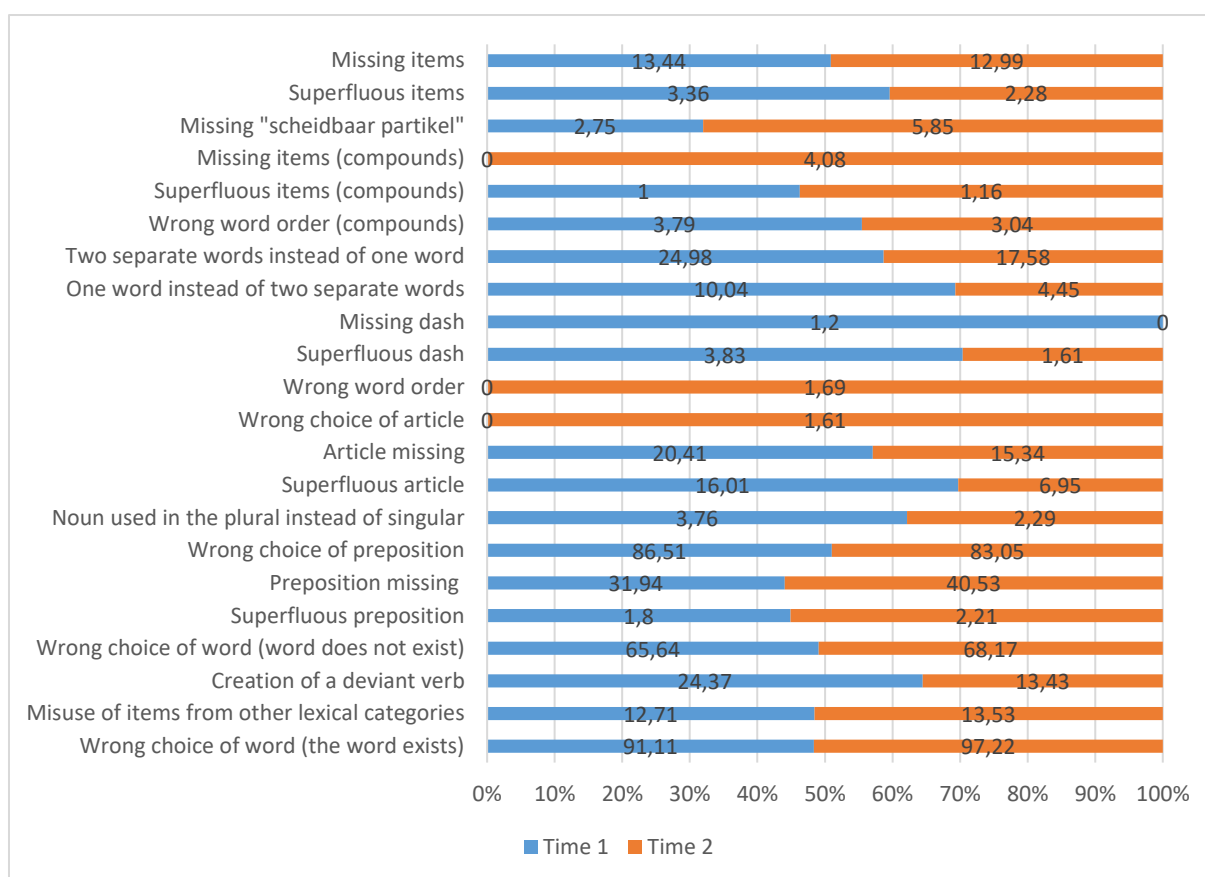


Figure 5.14. Evolution in the number of phraseological errors over time per category for the NON-CLIL Dutch group (relative frequencies per 250 words)

A comparison of the absolute frequencies (and relative frequencies) between Time 1 and Time 2 in the total number of phraseological errors and the most frequent categories of errors is shown in Table 5.42. below.

Table 5.42. Change in scores over time for the NON-CLIL Dutch group

	Time 1		Time 2	
	Mean	SD	Mean	SD
Number of phraseological errors	4.69 (5.81)	3.52 (3.81)	4.72 (5.54)	3.68 (4.06)
Wrong choice of existing word	1.10 (1.26)	1.59 (1.73)	1.11 (1.35)	1.39 (1.73)
Wrong choice of preposition	1.01 (1.20)	1.18 (1.34)	1.06 (1.15)	1.31 (1.43)
Wrong choice of word (does not exist)	0.68 (0.91)	0.98 (1.38)	0.81 (0.95)	1.31 (1.48)

Preposition missing	0.32 (0.44)	0.67 (0.93)	0.49 (0.56)	0.90 (0.96)	
Two separate words instead of one word	0.26 (0.35)	0.63 (0.86)	0.19 (0.24)	0.46 (0.59)	
Creation of a deviant verb	0.25 (0.34)	0.50 (0.72)	0.15 (0.19)	0.36 (0.45)	
Article missing	0.25 (0.28)	0.55 (0.60)	0.18 (0.21)	0.39 (0.47)	
Missing items	0.17 (0.19)	0.50 (0.56)	0.17 (0.18)	0.58 (0.60)	
Misuse of item from other lexical categories	0.13 (0.18)	0.41 (0.59)	0.15 (0.19)	0.36 (0.46)	
Superfluous article	0.18 (0.22)	0.45 (0.55)	0.08 (0.10)	0.32 (0.39)	*
One word instead of two separate words	0.10 (0.14)	0.30 (0.45)	0.04 (0.06)	0.26 (0.42)	
Missing items (compounds)	0 (0)	0 (0)	0.06 (0.06)	0.23 (0.24)	* *

A statistically significant decrease in the number of phraseological errors is found in the category *superfluous article*, but for the relative frequencies only: $t(71) = 2.003$, $p=.049$, $d=.25$ (the p-value for the absolute frequencies was almost significant: $p=.052$). However, the number of errors significantly increases for the category *missing items (compounds)* (absolute frequencies: $t(71) = -2.044$, $p=.045$, $d=.37$ and relative frequencies: $t(71) = -2.035$, $p=.046$, $d=.35$).

The results show that both groups of learners made positive progress in the accuracy of their phraseological productions, but it seems greater in the CLIL group. In fact, the mean for the total number of phraseological errors decreases in the CLIL group, while it increases in the NON-CLIL group (although not significantly). More specifically, both groups of learners improve their accuracy in more than half of the categories of phraseological errors, with a significant improvement found in two categories for the CLIL group and one category for the NON-CLIL group. A significant increase in the production of one specific type of phraseological errors is found in the NON-CLIL group. The effect sizes are rather small.

5.3.3.3. Inter-group analysis – frequency of use (RQ3)

The intra-group analysis looking at TL English revealed a significant development in a number of categories of phrasemes for the CLIL and the NON-CLIL group between Time 1 and Time 2. For TL Dutch, the intra-group analysis showed a significant improvement in a large number of categories of phrasemes for the NON-CLIL group between Time 1 and Time 2 and a significant decline in two categories for the CLIL group, which seems to indicate that CLIL learners of Dutch are gradually losing their advantage over their NON-CLIL peers.

Inter-group analyses were subsequently carried out for each measure (the number of words, the number of phraseological units and the various (sub-)categories). Only the categories that yielded significant results are discussed here.

Note that the scales used below may vary from figure to figure.

English

Figure 5.15. shows an increase of the total number of phraseological units in both learner groups, with a more substantial progress in the NON-CLIL group; there is a 2.47 improvement in the CLIL group versus a 6.61 improvement in the NON-CLIL group. This difference turns out to be statistically significant ($F(1, 142) = 5.77$, $p=.018$, $\eta^2=0.039$). Note that the intra-group variability (standard deviation) shows a larger decrease in the NON-CLIL learners (loss of 0.89) than in the CLIL learners (loss of 0.42).

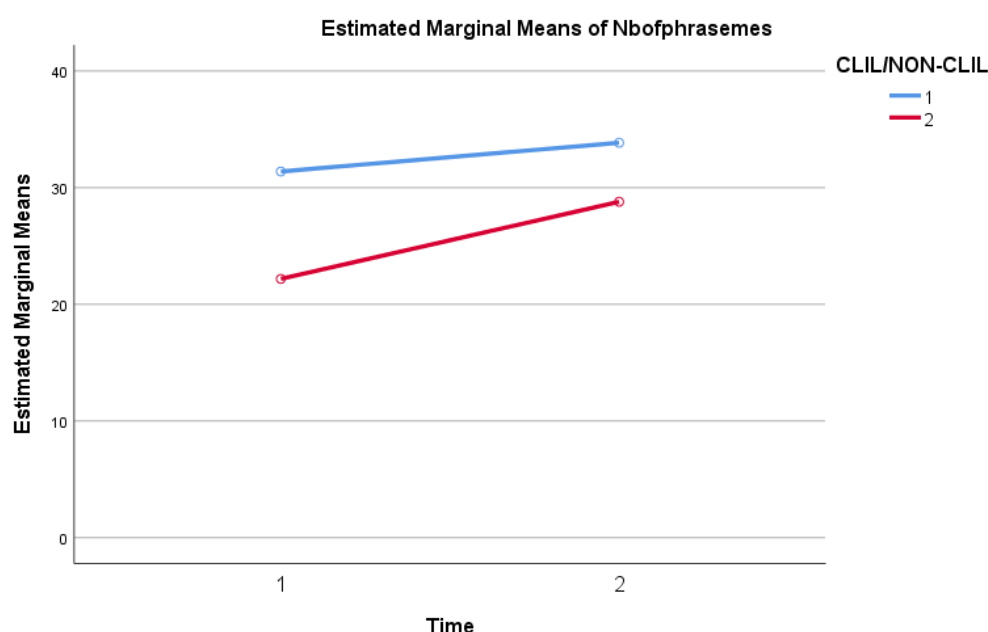


Figure 5.15. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *number of phrasemes*

Figure 5.16. shows an increase of *referential phrasemes* in the productions of the CLIL and the NON-CLIL learners. Progress is once again greater in the NON-CLIL group (6.15 improvement) than in the CLIL group (2.30 improvement) and this difference is statistically significant ($F(1, 142) = 6.27, p=.013, \eta p^2=0.042$). We also notice a sharper decrease in intra-group variability for NON-CLIL learners (-0.68) compared to CLIL learners (-0.33).

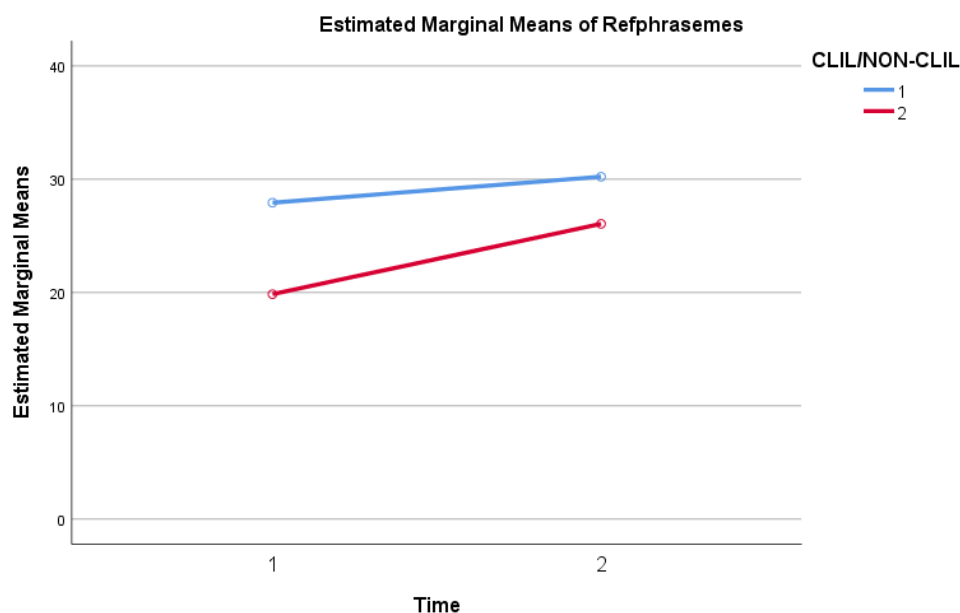


Figure 5.16. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *referential phrasemes*

As shown in Figure 5.17., CLIL learners decrease their production of *compounds* while NON-CLIL learners increase theirs. This difference between the groups is significant ($F(1, 142) = 4.84, p=.029, \eta p^2=0.033$).

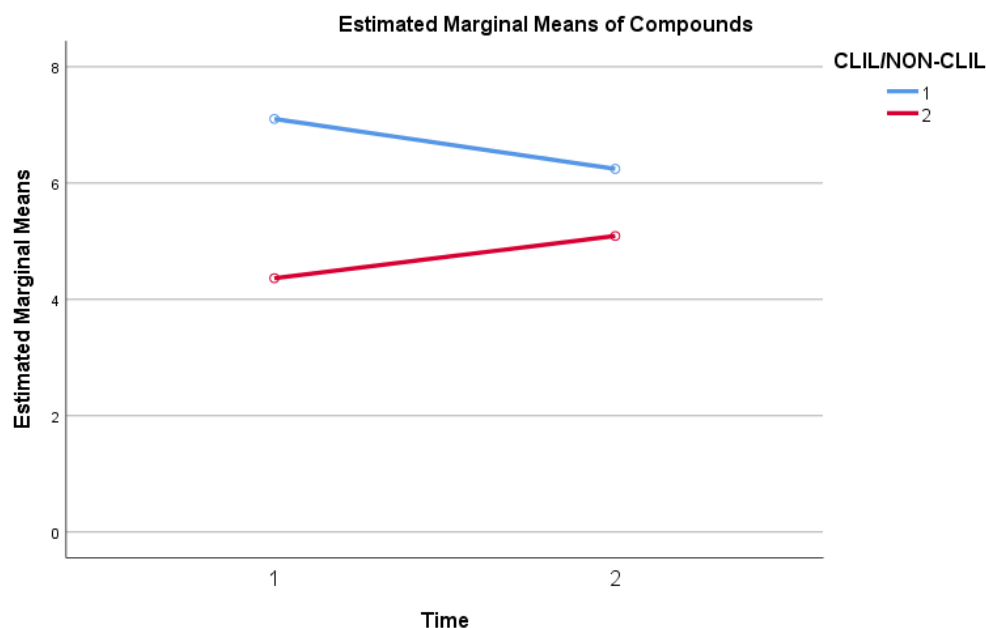


Figure 5.17. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *compounds*

The results show a statistically significant difference between the CLIL group's progress and the NON-CLIL group's progress in the total number of phraseological units and two (sub-)categories of phrasemes. The NON-CLIL learners outperform the CLIL learners in each case, either because the CLIL group progresses to a lesser extent or because they decline.

Dutch

A significant difference is found between the CLIL group's progress and the NON-CLIL group's progress for the category *number of words* ($F(1, 193) = 11.27, p=.001, \eta^2=0.055$). Figure 5.18. shows a reverse tendency between the groups: the CLIL group's word production slightly drops while the NON-CLIL group's productions become longer.

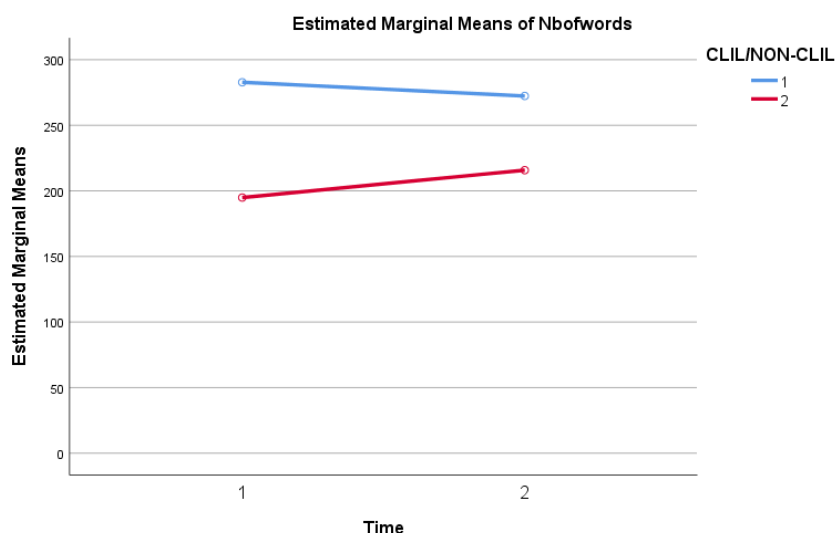


Figure 5.18. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *number of words*

Figure 5.19. shows an increase in the overall production of phraseological units for the NON-CLIL group and a slight decrease for the CLIL group. This difference is statistically significant ($F(1, 193) = 7.48, p=.007, \eta p^2=0.037$).

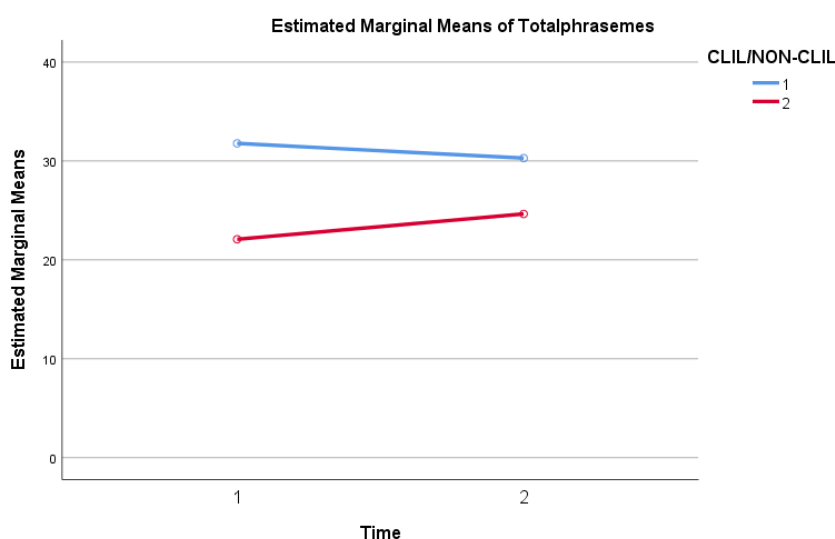


Figure 5.19. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *number of phrasemes*

Figure 5.20. shows that NON-CLIL learners slightly increase their production of *referential phrasemes* while CLIL learners decrease their use of this type of phraseological units. A significant difference is found between the two learner groups ($F(1, 193) = 7.53, p=.007, \eta p^2=0.038$).

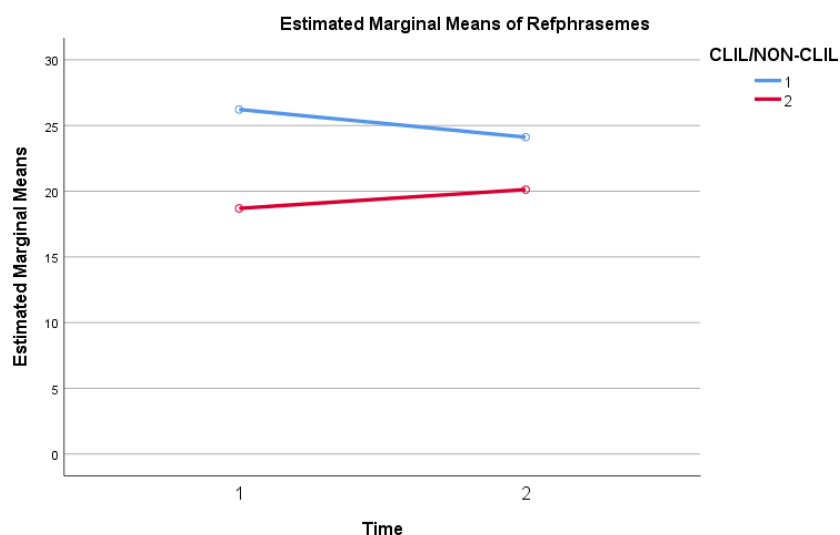


Figure 5.20. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *referential phrasemes*

Figure 5.21. shows that the CLIL group sharply decreases their production of *lexical collocations*. A slight improvement is seen in the NON-CLIL group. This difference between the groups is significant ($F(1, 193) = 11.20, p=.001, \eta^2=0.055$).

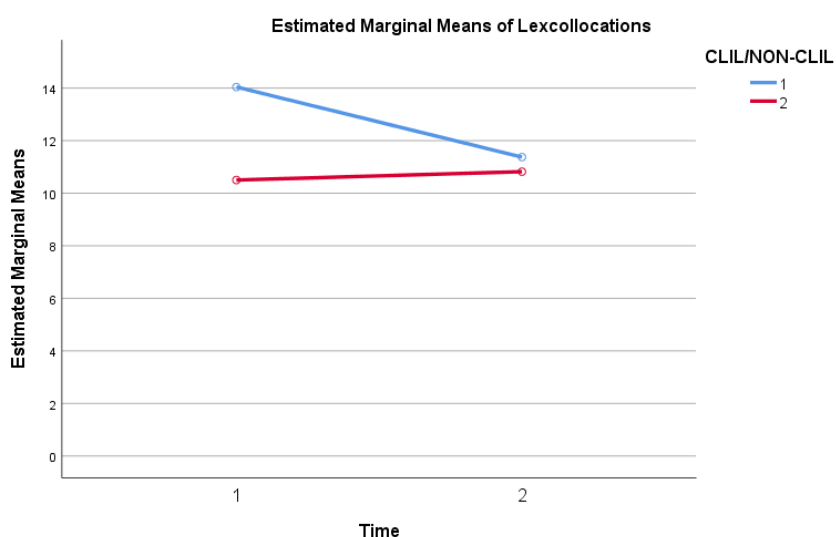


Figure 5.21. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *lexical collocations*

The results show that the two learner groups significantly differ in their phraseological development in terms of text length, overall phraseological production and two categories of phrasemes. The interaction plots clearly indicate opposing trends, namely an increase in the frequency of use in the NON-CLIL group and a decrease of use in the CLIL group.

These findings show that the CLIL group's progress stands in contrast with the NON-CLIL group's progress. When we look at the interaction plots for the non-significant results, we note a more similar progress between the groups in the majority of the categories, as in *textual phrasemes* for example (see Figure 5.22. below).

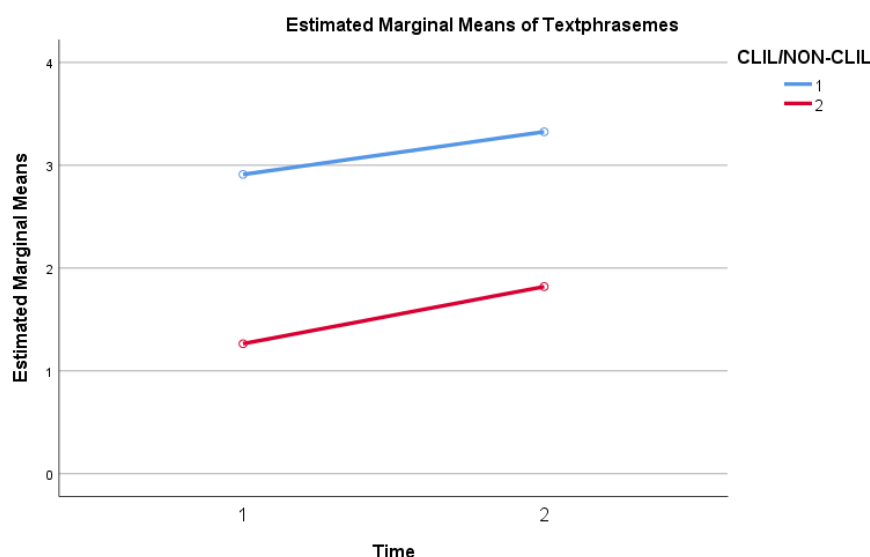


Figure 5.22. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *textual phrasemes*

5.3.3.4. Inter-group analysis – accuracy (RQ4)

English

The intra-group analysis showed a significant improvement of accuracy over time in one category of phraseological errors for the CLIL group. Accuracy significantly decreased in three categories of errors and significantly increased in one category for the NON-CLIL group.

Inter-group analyses were conducted for the frequent categories of phraseological errors found in Tables 5.39. and 5.40. We only consider the categories of errors that yielded significant results.

Figure 5.23. shows that, while the CLIL group equally performs at Time 1 and Time 2, the NON-CLIL group increases their production of phraseological errors in the category *word choice of existing word* (when looking at absolute frequencies) ($F(1, 142) = 4.76, p=.031, \eta^2=0.032$).

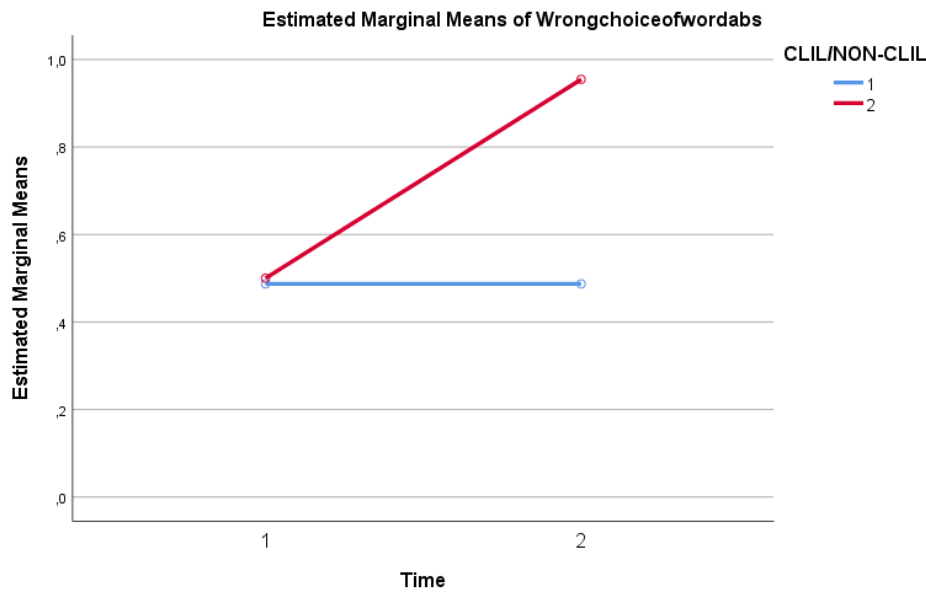


Figure 5.23. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *wrong choice of existing word* (absolute frequencies)

A statistically significant difference in the groups' progress is also found when looking at relative frequencies ($F(1, 142) = 5.59, p=.019, \eta^2=0.038$). Figure 6.24. shows the same trend as Figure 5.23. regarding the NON-CLIL group (increase in the number of errors). However, a slight rise in the number of mistakes is now found in the CLIL group.

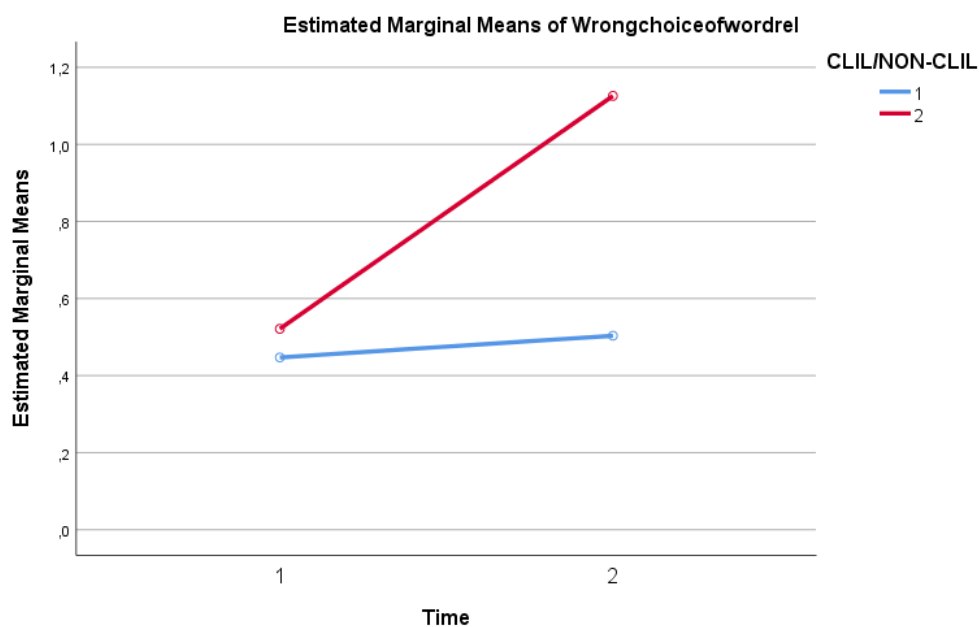


Figure 5.24. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *wrong choice of existing word* (relative frequencies)

Figure 5.25. shows opposing trends with regard to the category *superfluous article* (absolute frequencies): the CLIL learners made more phraseological errors over time while the NON-CLIL learners reduced the number of errors ($F(1, 142) = 4.62, p=.033, \eta^2=0.032$).

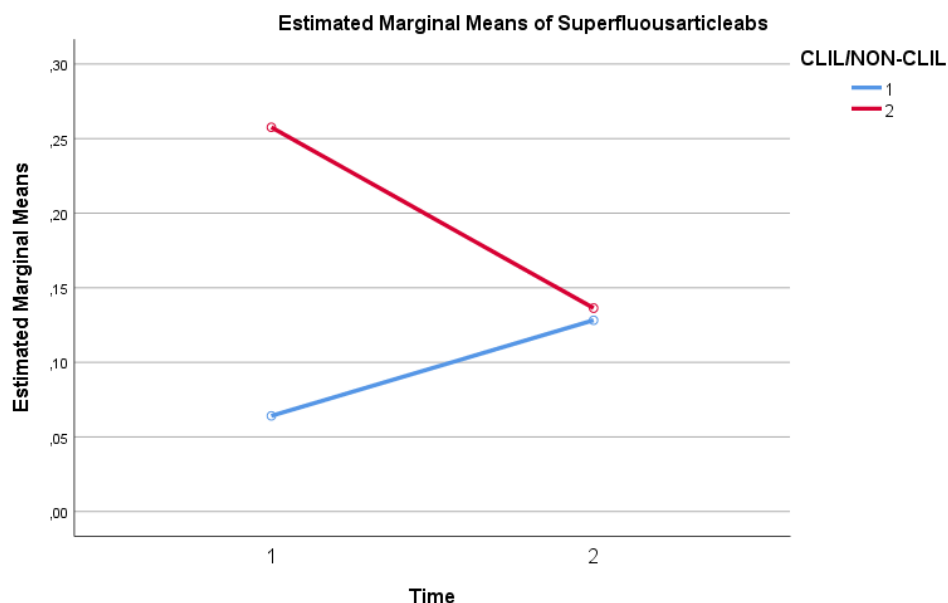


Figure 5.25. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *superfluous article* (absolute frequencies)

Figure 5.26. shows the same trends using relative frequencies. The difference between the groups is also statistically significant ($F(1, 142) = 4.83, p=.030, \eta^2=0.033$).

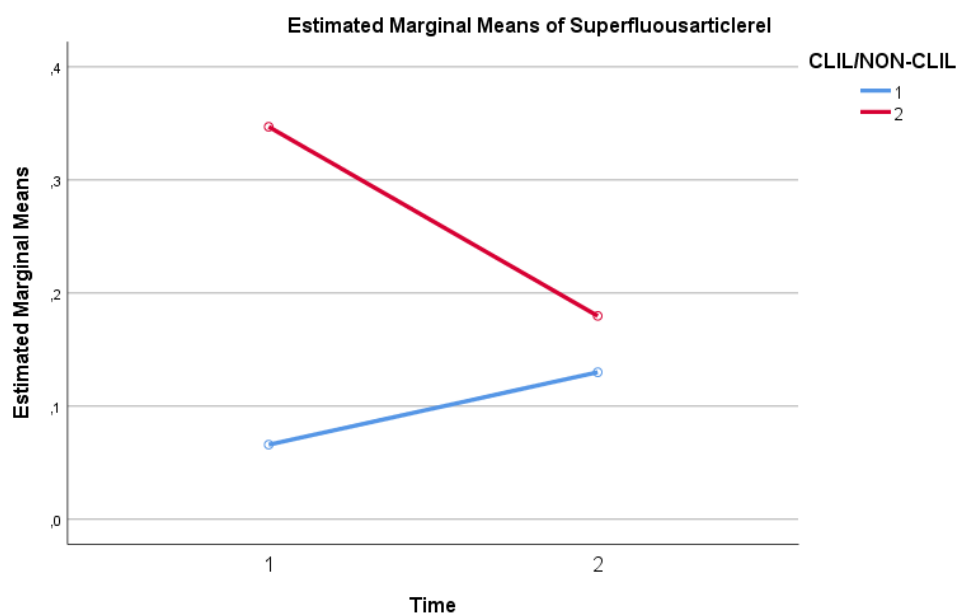


Figure 5.26. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *superfluous article* (relative frequencies)

Figure 5.27. shows a major increase in the category of errors *article missing* (absolute frequencies) for the NON-CLIL group versus a small drop of errors for the CLIL group ($F(1, 142) = 4.37, p=.038, \eta p^2=0.030$).

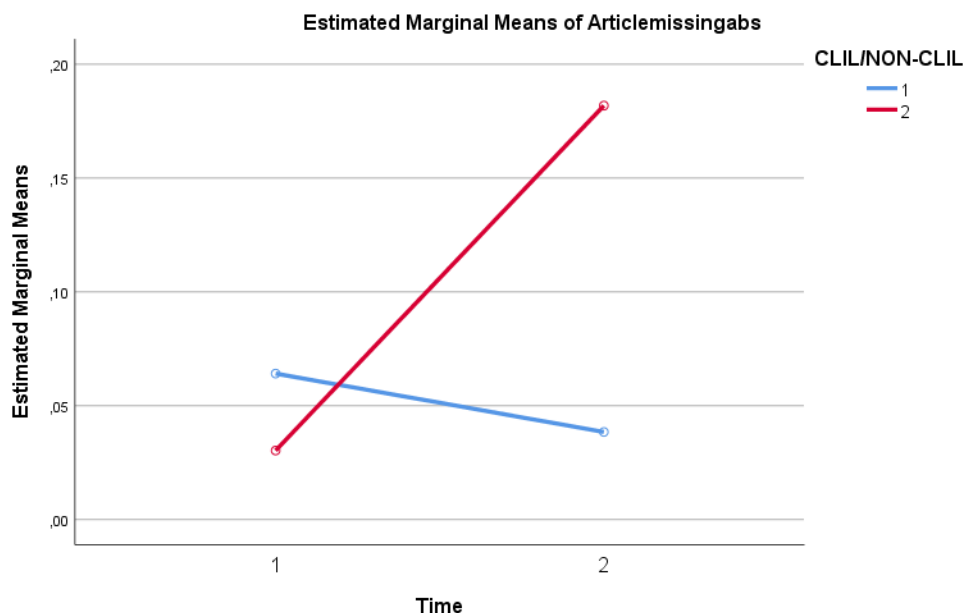


Figure 5.27. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *article missing* (absolute frequencies)

Figure 5.28. shows the same significant results when looking at relative frequencies $F(1, 142) = 5.45, p=.021, \eta p^2=0.037$.

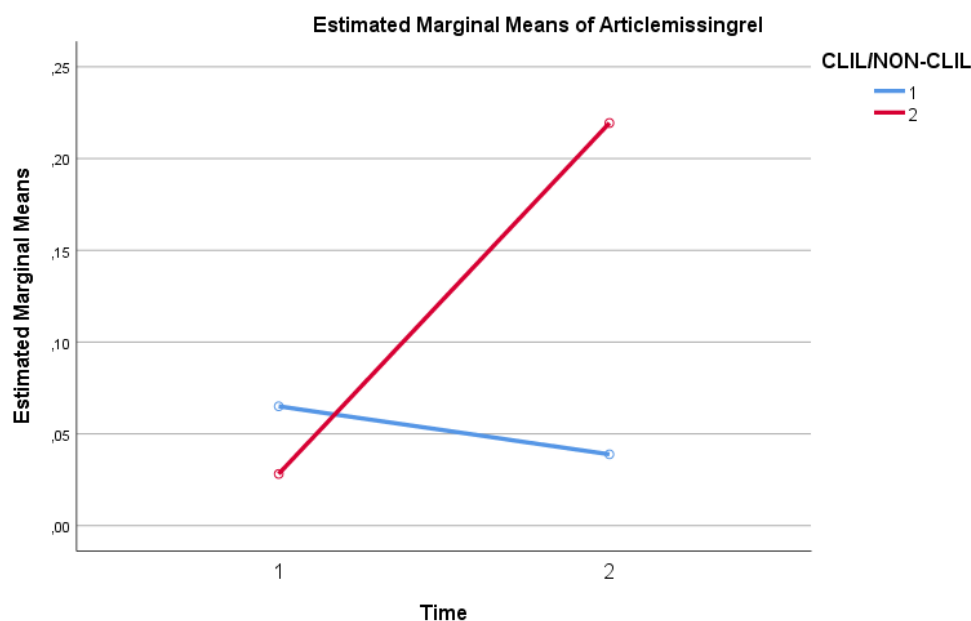


Figure 5.28. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *article missing* (relative frequencies)

A statistically significant difference between the groups is found for the category *superfluous preposition* (when looking at relative frequencies) ($F(1, 142) = 3.95$, $p=.049$, $\eta p^2=0.027$). As shown in Figure 5.29., the NON-CLIL group clearly improves their accuracy between the two data collections while the CLIL group slightly increases their production of errors.

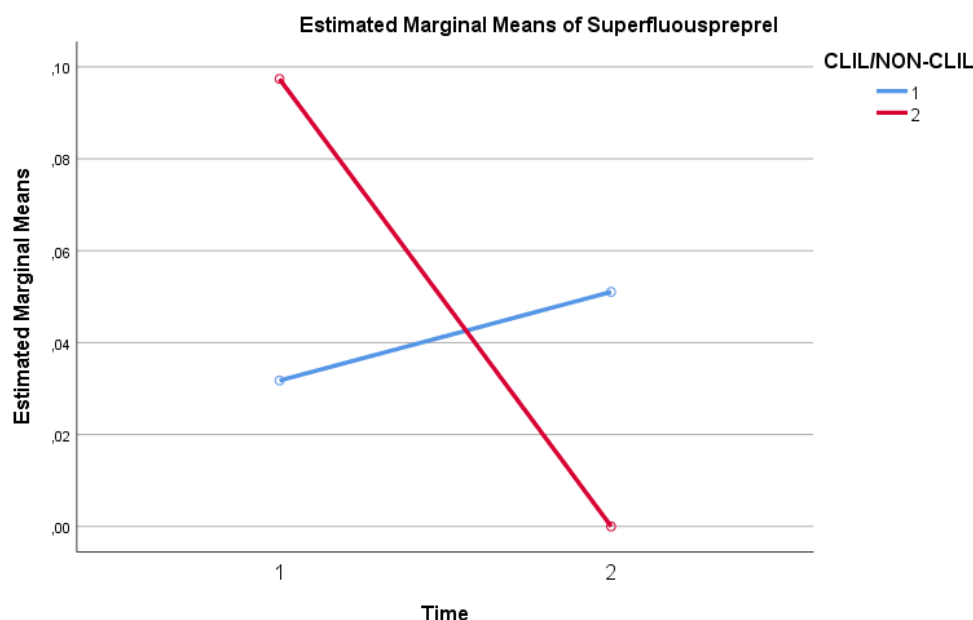


Figure 5.29. Progress in the CLIL English group (blue line 1) and the NON-CLIL English group (red line 2) for the category *superfluous preposition* (relative frequencies)

The results show that CLIL learners and NON-CLIL learners significantly differ in their production of phraseological errors over time, and this in four different categories of phraseological errors (including three where both absolute and relative frequencies yielded significant results). Overall, we can say that the CLIL learners' progress stands in contrast with the NON-CLIL learners' progress and that the trends differ according to the type of phraseological errors.

Dutch

The intra-group analysis showed a significant improvement of accuracy over time in two categories of phraseological errors for the CLIL group. Accuracy significantly improved in one category of errors but significantly decreased in another for the NON-CLIL group.

Inter-group analyses were conducted for the 13 frequent categories of phraseological errors shared by Tables 5.41. and 5.42. We only consider the categories of errors that yielded significant results.

A significant difference is found between the CLIL group's and the NON-CLIL group's progress for the category of error *superfluous article* (when looking at relative frequencies) ($F(1, 193) = 4.82, p=.029, \eta^2=0.024$). Figure 5.30. shows a clear drop in the production of superfluous articles for the NON-CLIL group while a slight increase can be seen for the CLIL group.

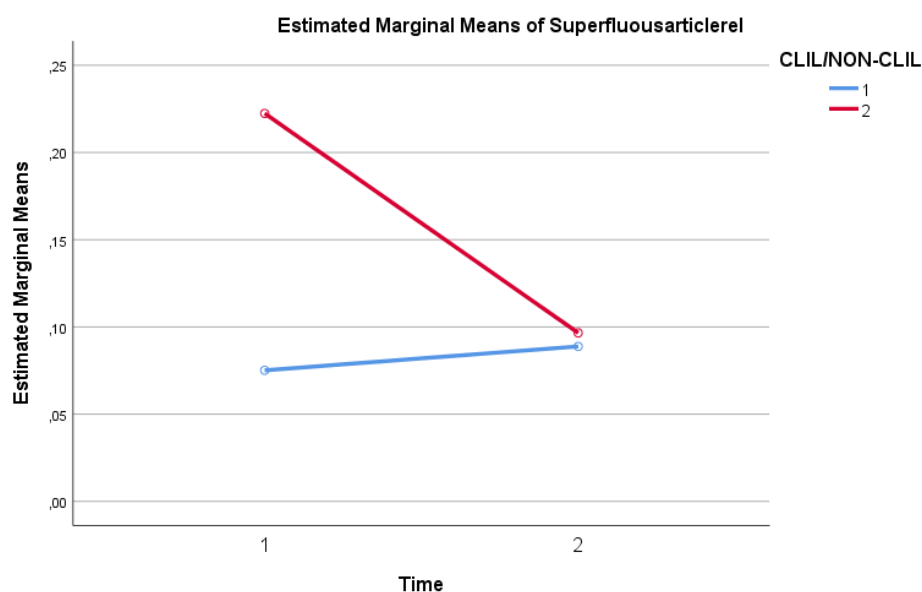


Figure 5.30. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *superfluous article* (relative frequencies)

Figure 5.31. shows that NON-CLIL learners make more mistakes over time in the category of error *missing items (compounds)* (when looking at absolute frequencies), while CLIL learners reduce the number of errors in that category ($F(1, 193) = 4.52, p=.035, \eta^2=0.023$).

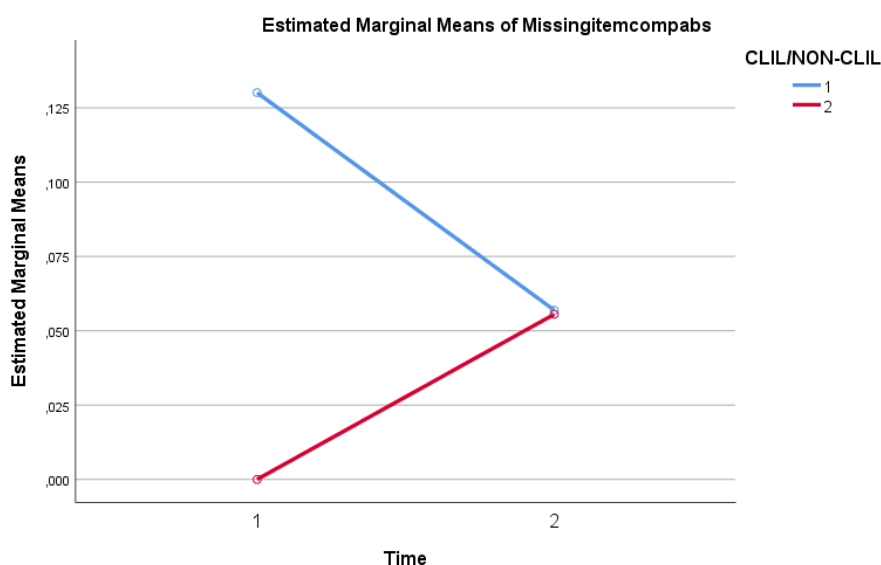


Figure 5.31. Progress in the CLIL Dutch group (blue line 1) and the NON-CLIL Dutch group (red line 2) for the category *missing items (compounds)* (absolute frequencies)

The results show that CLIL learners and NON-CLIL learners significantly differ in their production of phraseological errors over time, and this in two categories only. The trends are not the same according to the type of phraseological error, with the CLIL group improving in one category and decreasing in the other and the other way round for the NON-CLIL group.

5.3.4. Concluding remarks and discussion

The aim of the present study was to look at the development of phraseological language in CLIL and NON-CLIL contexts over the course of two consecutive school years and analyze if and to what extent the two learner groups' progress differed longitudinally. Phraseological competence was both analyzed in terms of frequency of use and accuracy.

Regarding the frequency of production of phraseological units over time within each group of learners (RQ1), the results differed according to the TL. The findings related to the learners of English confirmed our hypothesis; the CLIL and NON-CLIL learners significantly extended their range of phraseological units in a number of categories (three in CLIL and four in NON-CLIL) that turned out to be similar to both groups. The results for Dutch were rather unexpected; while we found a significant progress in the NON-CLIL group – they increased their range of phraseological units in a large number of (sub-)categories – a significant decline was found in the CLIL group – they decreased their use of phraseological units in two categories.

As for the accuracy of phraseological productions over time within each group (RQ2), the English CLIL group made significant improvement in one category of errors while the English

NON-CLIL group significantly increased their number of errors in three categories and decreased in one category. As for the CLIL learners of Dutch, they made significant improvement in two categories of phraseological errors while their NON-CLIL peers improved in one category but decreased in another.

The inter-group analysis related to frequency of use (RQ3) revealed a significant difference between the English CLIL learners' and the English NON-CLIL learners' development in three of the various categories measured, including the total number of phraseological units. The difference lay in the fact that NON-CLIL progressed better than their CLIL counterparts. A significant difference between the Dutch CLIL learners' and the Dutch NON-CLIL learners' development was found in four of the categories measured, with the NON-CLIL group systematically improving and the CLIL group systematically decreasing.

In terms of accuracy (RQ4), the groups significantly differed in four categories of phraseological errors for TL English and in two categories of phrasemes for TL Dutch, but no systematic trend was found.

What emerges from the analyses carried out for TL English is that both the CLIL and the NON-CLIL groups significantly improved their phraseological language over time in terms of frequency of use (but only in a small number of categories) and that larger gains were found in the NON-CLIL group compared to the CLIL group. As for TL Dutch, a marked improvement was seen in a large number of categories for the NON-CLIL learners, while their CLIL peers regressed. The results are less obvious when it comes to accuracy. Looking at the groups independently revealed a decrease in accuracy in the NON-CLIL learners only while the inter-group comparisons showed the CLIL and the NON-CLIL group improving and decreasing in turns according to the type of phraseological error.

All these findings seem to reject our hypothesis that CLIL learners would progress better over time; although CLIL learners outperformed their NON-CLIL peers at a given point in time, the extra exposure they received did not allow them to attain higher longitudinal development as far as phraseological competence is concerned. Therefore, as far as CLIL learners are concerned, the present study does not seem to support the fact that time spent learning a TL and proficiency in that language always positively correlate. This lack of substantial improvement of CLIL learners may be related to a general proficiency plateauing; CLIL learners were already better at Time 1, so the scope for improvement was more limited in that group. Moreover, the task may have been less challenging for CLIL learners at Time 2 since it was the same as the

one used at Time 1; a different evolution may have been observed with a different and/or more difficult task. On another note, no clear progress was noticed regarding accuracy of use, which is in line with the literature (Laufer & Waldman, 2011; Thewissen, 2013).

Results from the one-time comparison between the groups (see previous study in this chapter) revealed a substantial difference between the CLIL and the NON-CLIL learners in their use of phrasal verbs in terms of frequency and variety, with the CLIL group outperforming by far the NON-CLIL group. However, we saw no significant longitudinal progress in CLIL learners for that category of phraseological units; a slight decrease was even found for the CLIL learners of English. We can therefore not talk about a regular and constant use of that language structure yet. It has nevertheless been shown that interlanguage development is far from being stable – it is usually made of peaks and dips before stabilizing (Verspoor & Smiskova, 2012).

One of this study's limitations is the fact that it only has two data collection points. As reported by Singer and Willett (2003, p. 9) and mentioned in Chapter 1, Section 1.5.1.3., longitudinal studies should have “three or more waves of data” in order to study a precise form of change; “in statistical terms, two-wave studies cannot describe individual trajectories of change” (2003, p. 9). However, the time elapsed between the two data collection points (a year and a half) seemed long enough to expect clear evolution. Besides, other data collected in the framework of the project were also analyzed longitudinally and seem to confirm the present results. In fact, Hendrikx's study (2019) investigated the longitudinal development of intensifying constructions over time using the same sample and revealed similar patterns; as far as morphological constructions are concerned, “in the learners of Dutch, the difference between groups is significant in the fifth year, but in the sixth year, the CLIL students lose their advantage over the non-CLIL learners of Dutch” (Hendrikx, Van Goethem & Wulff, 2019, p. 81).

Finally, a recurring finding is the somewhat different results observed according to the TL (English or Dutch). This corroborates this thesis' previous empirical findings and those of another study carried out in the framework of this project (Van Mensel, Bulon, Hendrikx, Meunier & Van Goethem, submitted), which highlight the importance of taking into account the target language's status in research on CLIL.

CHAPTER 6: Receptive phraseological competence

6.1. Introduction

The learners' productive use of phraseological units and the development of their phraseological language over time was analyzed in Chapter 5. We now turn to the learners' receptive phraseological knowledge in order to further investigate on which linguistic aspects the two groups of learners differ.

Research on vocabulary knowledge describes the lexical competence construct as encompassing two main dimensions, i.e. productive knowledge and receptive knowledge (e.g. Nation, 2001). Investigating both dimensions contributes to a better understanding of the learners' lexical (here phraseological) knowledge. Whilst the previous analysis of the learners' phraseological productive knowledge (Chapter 5) looked at a large variety of phraseological units (collocations, idioms, routine formulae, compounds, complex prepositions, etc.), the present chapter only tackles the receptive knowledge of verb-noun collocations. Verb-noun collocations are frequent in the language (Cowie, 1991; Howarth, 1996; Nesselhauf, 2005; Siepmann, 2005a) and were reported as particularly problematic for the learners (Biskup, 1992; Bahns & Eldaw, 1993; Liu, 1999; Chen, 2002). Yan's study (2010) showed that, among all types of lexical collocations, verb-noun collocations accounted for 50% of lexical collocation mistakes. One of the reasons accounting for this large number of errors may lie in the fact that the verbs in verb-noun collocations are often highly frequent and that learners have great difficulty using these high-frequency verbs (e.g. Altenberg & Granger, 2001, see Chapter 1, Section 1.9.1.1. for further references). In addition, these verbs are often used in their delexicalized meaning, i.e. a meaning that is not usually supplied by introspection and that can often only be explained in terms of usage (Sinclair, 1991); such verbs do not often add much meaning to the combination, as for example in *take care* and *make progress* (Gyllstad, 2009). The choice of these verbs is mostly arbitrary (Allerton, 1984); choosing *take a step* instead of **make a step*, for instance, is semantically unmotivated and largely language-specific.

After describing the research questions and hypotheses in Section 6.2., Section 6.3. tackles the methodological aspects of the study (the participants and the receptive tests). The results are presented in Section 6.4., before concluding in Section 6.5.

6.2. Research questions and hypotheses

This study's main objective is to compare the receptive verb-noun collocational knowledge of French-speaking learners of English or Dutch as a TL in CLIL and NON-CLIL settings. Our research questions are the following:

1. Do CLIL and NON-CLIL learners differ in their receptive knowledge of verb-noun collocations? (RQ1);
2. What is the impact of the collocation itself on the results, first in terms of its frequency and second in terms of its degree of similarity with the L1 (congruence)? (RQ2).

As far as RQ1 is concerned, it can be hypothesized that the CLIL group will perform better at the receptive test than the NON-CLIL group. To our knowledge, the receptive collocational knowledge of CLIL learners has not been investigated yet, but previous research comparing the receptive collocation competence of learners in lower input (EFL) and higher input (ESL) conditions showed better results in the ESL groups (e.g. Shehata, 2008; Alsakran, 2011; Bueraheng & Laohawiriyanon, 2014; see Chapter 1, Section 1.9.2. for further references). Regarding RQ2, it seems interesting to look at other variables than the learning context as research shows that verb-noun collocations may be more or less difficult to acquire according to, for instance, their degree of similarity with the learners' L1 (Yamashita & Jiang, 2010; see Sections 1.7.2. and 1.9.2. in Chapter 1). The frequency of the collocation and its single-word items may also have an influence (Nguyen & Webb, 2017); (highly) frequent words are usually the first ones to be acquired by learners since they will be the first words learners will be exposed to (Tomasello & Stahl, 2004), hence increasing their degree of familiarity. The degree of delexicalization of the verb also seems to play a role (Källkvist, 1999; Altenberg & Granger, 2001; Nesselhauf, 2005). The present study gives prominence to two aspects that are inherent characteristics of the collocation itself, that is its frequency of occurrence and its degree of similarity with the learners' L1, while considering the degree of restriction of the verb.

6.3. Methodology

6.3.1. Participants

English

The sample of French-speaking secondary school learners of English consisted of 166³⁹ participants. They were in their last year of upper-secondary school when the data were

³⁹ The reason why the sample sizes (*n*) for the different variables in the tables do not correspond to 166 is that data were missing for some participants.

collected. They learned English as a TL either in a CLIL or in a NON-CLIL context. The tables below provide information about the participants' background characteristics (Table 6.1.) and about their curricular and extracurricular exposure to English (Table 6.2.).

Table 6.1. Description of the English sample – background characteristics

	<i>n</i>	Gender (%) (<i>n</i> =165)		Home language (%) (<i>n</i> =160)		
		Male	Female	French	French + other	Mostly other
CLIL	87	47.1	52.9	65.1	22.1	12.8
NON-CLIL	79	38.5	61.5	74.3	23	2.7
ALL	166	57	43	69.4	22.5	8.1

57% of the participants were female, while 43% were male. The majority of the participants (69.4%) spoke only French at home. 22.5% spoke sometimes French and sometimes another language (only three of them specified and said that the other language was English). The rest of the participants (8.1%) mostly spoke another language than French at home (amongst those who specified, only two mentioned English). They were not excluded from the analyses. We note that the difference between CLIL and NON-CLIL learners is more marked for the 'mostly other' category of answer: 12.8% of learners in CLIL mostly spoke another language at home versus 2.7% in NON-CLIL.

Table 6.2. Description of the English sample – (extra)curricular exposure to the TL

	<i>n</i>	Nb of hours/week at school (%) (<i>n</i> =161)		Extra classes (%) (<i>n</i> =159)	Stay abroad (%) (<i>n</i> =159)	Informal contact outside school (Means) (<i>n</i> =157)
		4h/week	+ 4h/week			
CLIL	87	74.4	25.6	3.5	9.3	3.37
NON-CLIL	79	72	28	1.4	1.4	2.91
ALL	166	73.3	26.7	2.5	5.7	3.16

In terms of the participants' linguistic exposure at school, we know that most of them (74.4% in CLIL and 72% in NON-CLIL) received 4 hours of English as a TL per week, while some of

them (25.6% in CLIL and 28% in NON-CLIL) received more than 4h/week. As far as extracurricular exposure is concerned, only 3.5% of CLIL and 1.4% of NON-CLIL pupils took extra classes of English outside school. 9.3% of CLIL learners and 1.4% of NON-CLIL learners have stayed in an English-speaking country or region for more than three months. CLIL learners (Mean = 3.37, SD = 0.838) seemed to benefit from more current informal contact with the TL outside school than NON-CLIL learners (Mean = 2.91, SD = 0.793), $t(155) = 3.481$, $p = .001$, $d = .56$ (medium effect size).

Dutch

The sample of French-speaking secondary school learners of Dutch consisted of 214⁴⁰ participants. They were also in their last year of upper-secondary school at the time of the data collection. They learned Dutch as a TL either in a CLIL or in a NON-CLIL context. The tables below give a description of the participants' background characteristics (Table 6.3.) and of their (extra)curricular exposure to Dutch (Table 6.4.).

Table 6.3. Description of the Dutch sample – background characteristics

	<i>n</i>	Gender (%) (<i>n</i> =214)		Home language (%) (<i>n</i> =197)		
		Male	Female	French	French + other	Mostly other
CLIL	133	55.6	44.4	83.9	15.3	0.8
NON-CLIL	81	48.1	51.9	68.5	28.8	2.7
ALL	214	47.2	52.8	78.2	20.3	1.5

47.2% of the participants were female, while 52.8% were male. Most of the participants (78.2%) spoke only French at home. A number of learners (20.3%) answered speaking sometimes French and sometimes another language (Dutch was mentioned by seven of them). Only 1.5% said speaking mostly another language than French, but did not specify.

Table 6.4. Description of the Dutch sample – (extra)curricular exposure to the TL

<i>n</i>	Nb of hours/week at school (%) (<i>n</i> =197)	Extra classes (%) (<i>n</i> =197)	Stay abroad (%) (<i>n</i> =196)	Informal contact outside school

⁴⁰ The reason why the sample sizes (*n*) for the different variables in the tables do not correspond to 214 is that data were missing for some participants.

(Means)						
(n=188)						
		4h/week	+ 4h/week			
CLIL	133	87.9	12.1	0.8	8.1	2.52
NON-CLIL	81	100	0	8.2	0	2.15
ALL	214	92.4	7.6	3.6	5.1	2.39

As far as the participants' curricular exposure is concerned, 92.4% received 4 hours of Dutch as a TL per week (the remaining 7.6% are CLIL learners receiving more than 4h/week). Regarding extracurricular exposure, only 0.8% of CLIL learners and 8.2% of NON-CLIL learners took extra classes of Dutch outside school. Only a small number of participants (5.1%, only CLIL learners) have stayed in a Dutch-speaking country or region for more than three months. An independent samples t-test revealed a significantly higher amount of current informal contact with the TL outside school in the CLIL group (Mean = 2.52, SD = 0.783) than in the NON-CLIL group (Mean = 2.15, SD = 0.720), $t(186) = 3.253$, $p = .001$, $d = .49$.

6.3.2. Receptive collocation test

English

The receptive test used for English is based on Gyllstad's (2007) work. In the framework of his doctoral dissertation, Gyllstad developed a test called COLLEX 5 (last version) which proved to yield reliable and valid scores of receptive knowledge of English verb-noun collocations used with advanced learners of English. His main criteria for the selection of items is frequency. In fact, Gyllstad chose his items on the basis of word lists developed by Nation (2006), which "cover the 1000 most frequent and the second 1000 most frequent word families of English [and] 570 word families that are frequent in upper secondary school and university texts from a wide range of subjects" (Gyllstad, 2007, p. 76). The nouns from these word lists were then checked for frequent verb collocations using the BNC.

Gyllstad's version consisted of 50 items, a number that is considered ideal in terms of reliability (Nation, 2001). Due to time constraints and concentration issues – our participants had to perform more than 10 different tasks in one day – we reduced the number of items to 18. To do so, we made use of the English Vocabulary Profile list. This online tool shows the words and phrases learners are ought to know at each level of the CEFR. We checked all 50 collocations from Gyllstad's test and selected the ones corresponding to B1-B2 level, which is the expected

level at upper secondary school in Belgium. Five collocations from the C1-C2 level were also included to better discriminate between learners.

The items were presented in a multiple-choice exercise where learners had to choose between three options: one frequent and conventionalized English lexical collocation (the ‘target collocation’) and two combinations that are not frequent, nor conventionalized in English functioning as distractors (the ‘pseudo-collocations’). Since in many constructions, the noun selects the verb element and not the other way round (e.g. the noun ‘bath’ selects the verb ‘have’ in the collocation ‘have a bath’; the verb ‘have’ can collocate with a large number of other nouns) (Mel’čuk, 1998), we only alternated the verb. The test was context-independent – the collocations were not contextualized in sentences – and there was no control for guessing. The test items were ordered randomly and so were the three possible answers; there were thus different versions of questionnaires depending on the order of the items. Examples are presented in (3) and (4) below (see Appendix 3 for the complete version of the exercise):

- (3) a. do damage
b. make damage
c. run damage

- (4) a. commit a crime
b. comply a crime
c. conduct a crime

Regarding the scoring system, a correct answer was awarded 1 point and an incorrect answer 0 point. If the participant did not answer, it was recoded as an incorrect answer.

Tables 6.5. and 6.6. below show the word frequencies of each single-word item used in the test (collocations and distractors) and the CEFR level corresponding to each target collocation.

Table 6.5. Frequency levels of selected English verbs and nouns⁴¹

1K ⁴²	bring, catch, create, do, draw, drop, exercise, give, hold, house, keep, last, lay, lose, make, pass, pay, press, pull,	damage, fire, discussion, visit, tear, right, opportunity, speech, count, purpose, friend, secret

⁴¹ Partly taken from Gyllstad (2007). All the nouns used in the test belonged to the first two frequency bands.

⁴² Gyllstad (2007) used the JACET 8000 word list (Ishikawa, Uemura, Kaneda, Shimizu, Sugimori & Tono, 2003), which is based on the BNC, to check the frequencies of the individual words. These were organized in different frequency bands according to that list: 1K refers to the frequency band containing the 1,000 most frequent words of English, 2K to the 2,000 most frequent words of English, etc.

	push, put, raise, run, serve, set, take, turn, win	
2K	conduct, earn, gain, perform	charge, conclusion, crime, breath, birth, harm
3K	seize, shed	
4K	commit, comply, employ	

Table 6.6. CEFR levels of selected English verb-noun collocations

B1	put out a fire, hold discussions, commit a crime, keep a secret, make friends
B2	do damage, pay a visit, seize an opportunity, draw a conclusion, give a speech, hold one's breath, give birth, do harm
C1	shed tears
C2	exercise rights, press charges, lose count, serve a purpose

Dutch

The receptive test for Dutch was developed on the basis of the book KONNEX⁴³, which includes various types of exercises (including multiple choice questions) designed to help learners acquire Dutch word combinations. The book aims at B2 level learners of Dutch, which corresponds to the level expected at upper secondary school in French-speaking Belgium. A number of 22 verb-noun collocations⁴⁴ (along with their ‘pseudo-collocations’ and the context sentence) were selected from the book – we are aware of the fact that the number of items is relatively small (see Nation, 2001) but our participants had to perform more than 10 different tasks in one day. The context sentences were sometimes a bit modified in order to make sure the learners would clearly understand it. The test was checked by different professors of Dutch as a TL and was pre-tested with first year bachelor’s French-speaking students learning Dutch at university.

The items were presented in a multiple-choice exercise where learners had to identify the frequent and conventionalized Dutch lexical collocation (the ‘target collocation’) among three options. Two other combinations that are not frequent, nor conventionalized in Dutch were functioning as distractors (the ‘pseudo-collocations’). Only the verb changed and the test was

⁴³ VanParys, J., Stumpel, R., Boers, T., & de Bot, V. (2008). *KONNEX: Uitdrukkingen en vaste verbindingen in de Nederlandse taal*. Utrecht: ThiemeMeulenhoff.

⁴⁴ However, two Dutch collocations are not verb-noun collocations: *openbaar maken* (‘make public’) and *gewond raken* (‘get injured’) combine verbs with adjectives.

context-dependent – the verb-noun collocations were contextualized in sentences. There was no control for guessing. Both the test items and the three possible answers were presented to the learners in a random order, so that there were different versions of exercises depending on the order of the items. Examples are given in (5) and (6) below (see Appendix 3 for the complete test):

(5) Zo kan ik me niet concentreren. Dat lawaai buiten ... op mijn zenuwen.

- a. loopt
- b. werkt
- c. drukt

(6) De mensen achter mij in de bioscoop zaten de hele tijd commentaar te ... op de acteurs.

Heel irritant!

- a. leveren
- b. praten
- c. bezorgen

The reason why the receptive test used for Dutch differed from the one used for English is that we could not find information on its validity as we did for the one used for English. We decided therefore to stick to the original format of the material. The context provided in the Dutch test was however minimal.

Regarding the scoring system, a correct answer was awarded 1 point and an incorrect answer 0 point. If the participant did not answer, it was recoded as an incorrect answer.

6.4. Results

6.4.1. Research question 1: do CLIL and NON-CLIL learners differ in their receptive knowledge of verb-noun collocations?

This section is devoted to the comparative analysis of the CLIL and NON-CLIL learners' performance to the receptive collocation test.

English

Figure 6.1. shows that the number of correct (54.8%) and incorrect (45.2%) answers (on a total of 2,988 answers) is rather evenly distributed, with a slightly higher frequency of correct answers.

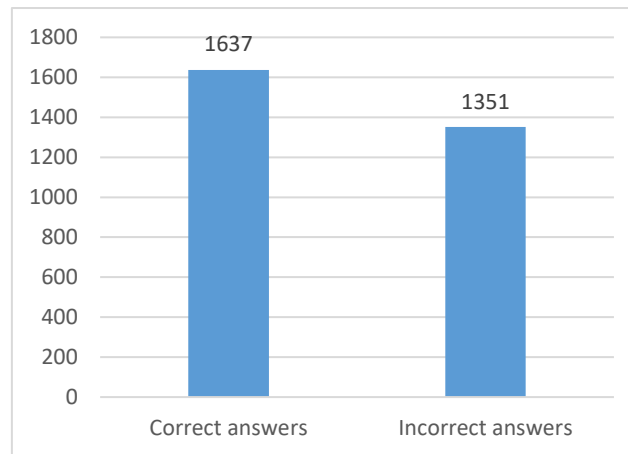


Figure 6.1. Frequency of correct and incorrect answers given by the learners of English
(Total=2,988)

If we look at the distribution of correct and incorrect answers per learner group (see Figure 6.2. below), we see that CLIL learners gave more correct answers (61.4%) than their NON-CLIL peers (47.5%), and this difference is significant (Mann-Whitney U test: Mean (CLIL) = 11.06 and Mean (NON-CLIL) = 8.54, $U = 1898$, $p = .000$, $r = -.39$).

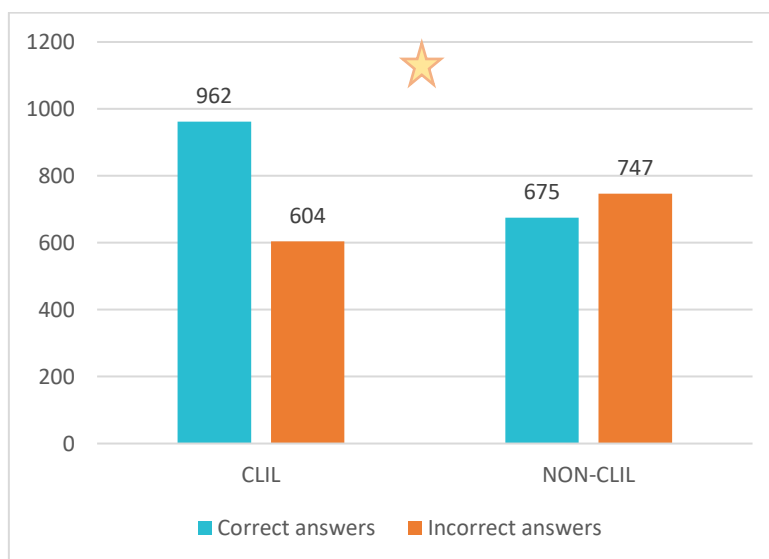
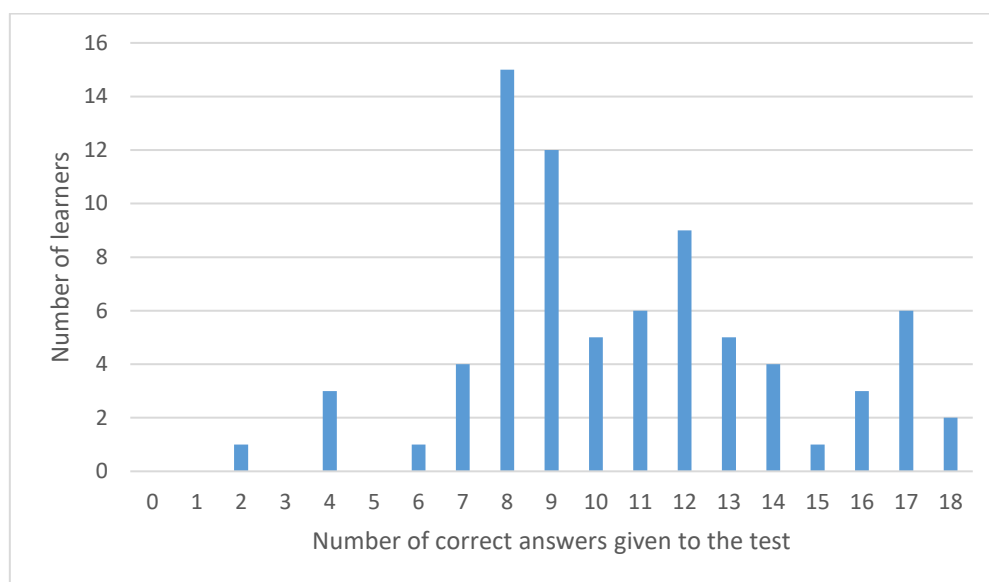
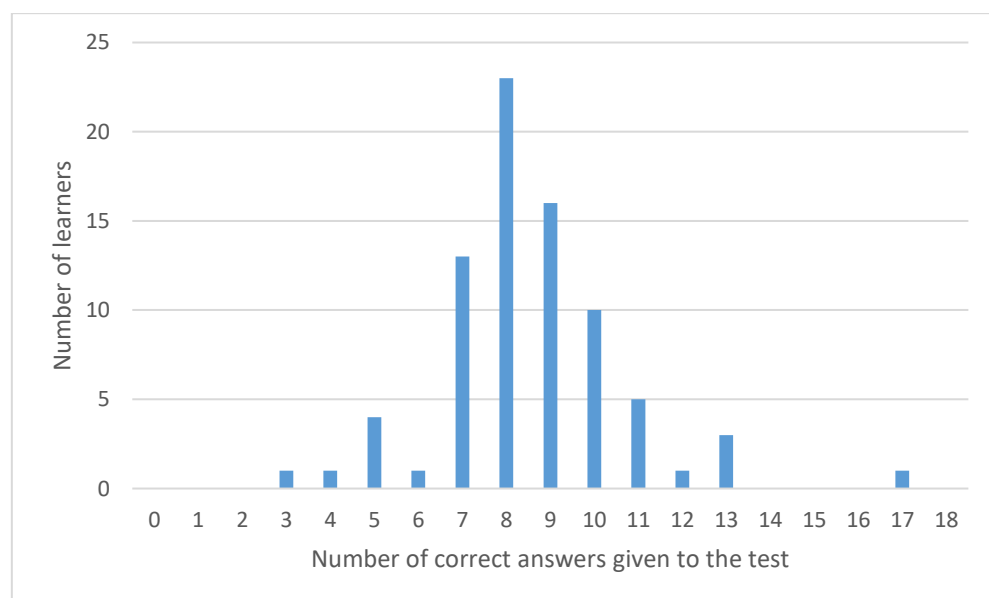


Figure 6.2. Frequency of correct and incorrect answers per English learner group

Table 6.7 and Figures 6.3. and 6.4. below give an overview of the data distribution across the CLIL and the NON-CLIL group. The horizontal axis of the histograms corresponds to the number of correct answers given to the test on a scale from 0 to 18 (18 indicating that the learner gave a correct answer for all items). The vertical axis corresponds to the number of learners who gave a certain number of correct answers. For example, 6 CLIL learners correctly identified 17 verb-noun collocations out of 18 and 10 NON-CLIL learners correctly identified 10 verb-noun collocations out of 18.

Table 6.7. Measures of global tendency in each English learner group

	<i>n</i>	Mean	Median	Minimum	Maximum	Range	SD	Variance
CLIL	87	11.06	11	2	18	16	3.499	12.241
NON-CLIL	79	8.54	8	3	17	14	2.099	4.405

**Figure 6.3.** Distribution of correct answers given to the English receptive test (CLIL)**Figure 6.4.** Distribution of correct answers given to the English receptive test (NON-CLIL)

The distributions are not normal, neither in the CLIL or the NON-CLIL group (significant Shapiro-Wilk tests). We see that the results are more dispersed for CLIL learners (median absolute deviation of 2) than for NON-CLIL learners (median absolute deviation of 1).

Dutch

As shown in Figure 6.5., the number of correct answers (57.2%) and incorrect answers (42.8%) (on a total of 4,730 answers) is rather evenly distributed, with a slightly higher frequency of correct answers.

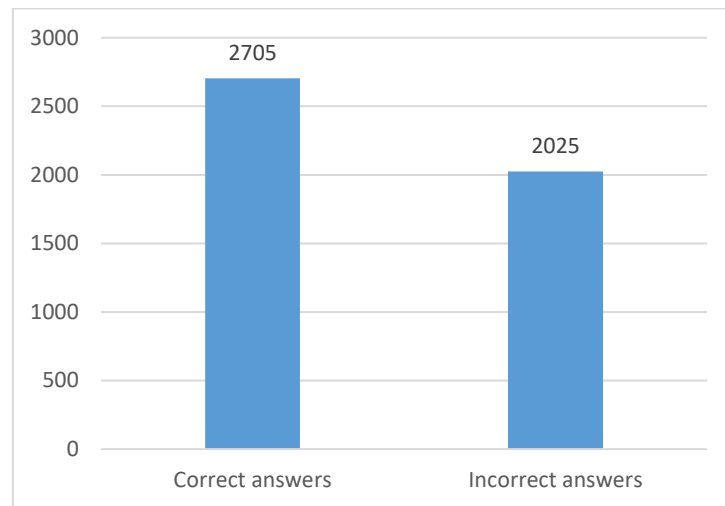


Figure 6.5. Frequency of correct and incorrect answers given by the learners of Dutch
(Total=4,730)

When looking at the distribution of correct and incorrect answers per learning condition (see Figure 6.6. below), we see that CLIL learners gave more correct answers (65.4%) than their NON-CLIL peers (43.9%), and this difference is significant (Mann-Whitney U test: Mean (CLIL) = 14.38 and Mean (NON-CLIL) = 9.66, $U = 1526$, $p = .000$, $r = -.60$).

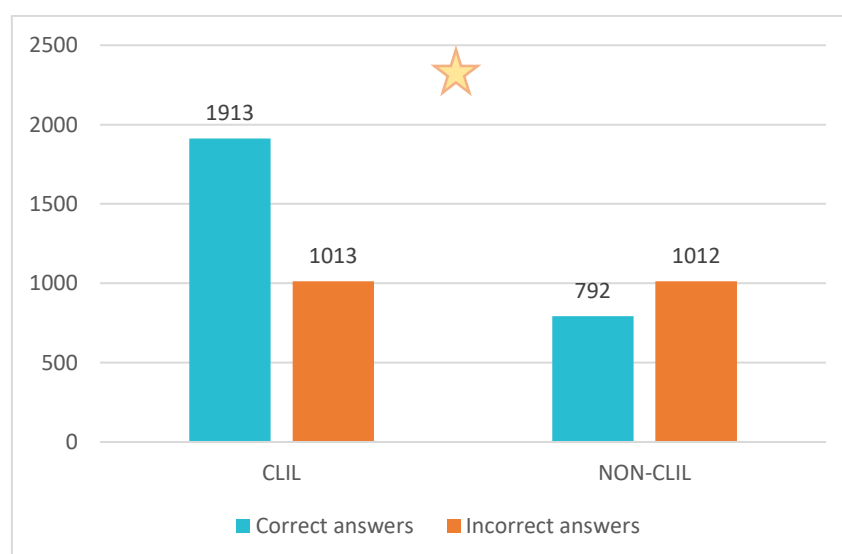


Figure 6.6. Frequency of correct and incorrect answers per Dutch learner group

Table 6.8. and Figures 6.7. and 6.8. give an overview of the data distribution per learner group.

Table 6.8. Measures of global tendency in each Dutch learner group

	<i>n</i>	Mean	Median	Minimum	Maximum	Range	SD	Variance
CLIL	133	14.38	15	5	22	17	3.133	9.814
NON-CLIL	82	9.66	9.50	3	18	15	3.147	9.907

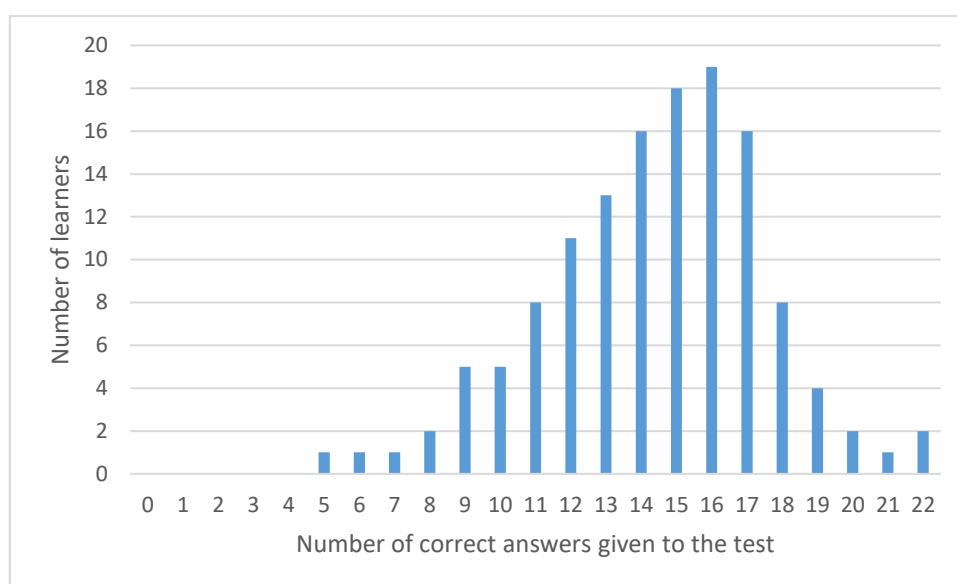


Figure 6.7. Distribution of correct answers given to the Dutch receptive test (CLIL)

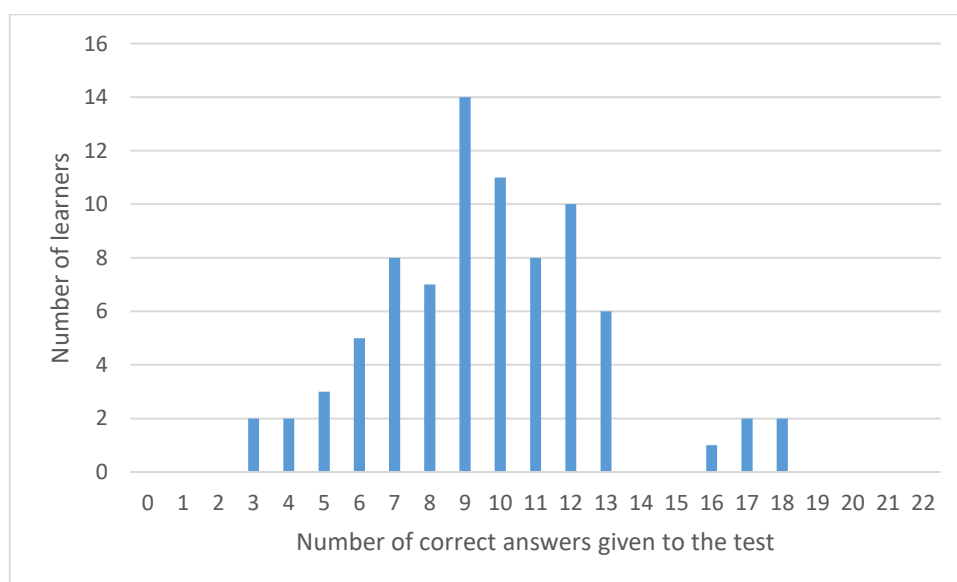


Figure 6.8. Distribution of correct answers given to the Dutch receptive test (NON-CLIL)

The distribution is normal in the CLIL group (non-significant Shapiro-Wilk test) and not normal in the NON-CLIL group (significant Shapiro-Wilk test). The data are quite dispersed in both groups.

6.4.2. Research question 2: Impact of the collocation

In this section, we first have a closer look at the results per collocation. Then, we try to identify the aspects related to the collocation that have an impact on the type of answer given by the learners: Section 6.4.2.1. analyzes the influence of the frequency of the collocation, while Section 6.4.2.2. qualitatively analyzes the influence of the degree of similarity of the target collocation with the L1 (congruence).

English

Figure 6.9. below shows the number of correct and incorrect answers per item.

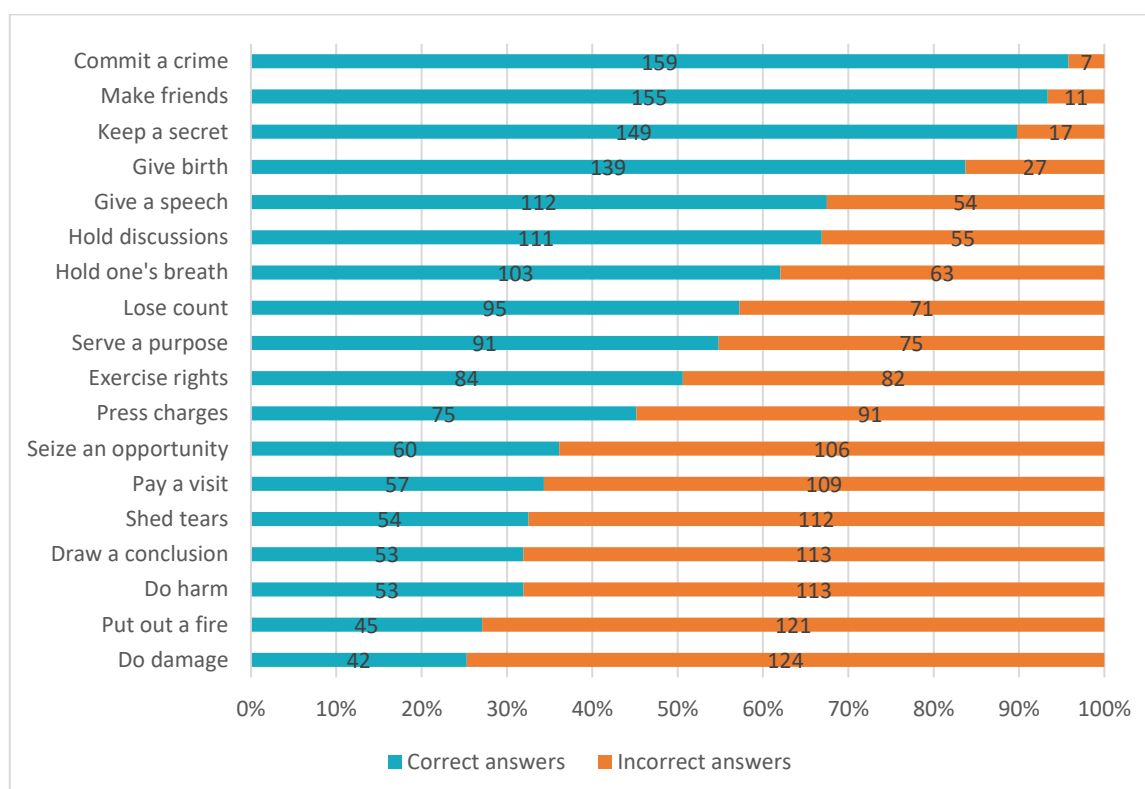


Figure 6.9. Frequency of correct and incorrect answers per collocation – English

There seems to be an effect of the collocation on the test's results. In fact, some items seem to be easier to retrieve than others. Examples are *commit a crime*, *make friends*, *keep a secret* or *give birth*. Others like *do damage*, *put out a fire*, *do harm*, *draw a conclusion*, *shed tears* or *pay a visit* seem to be more difficult.

A Chi-square test of independence showed that the correctness of the answer (correct or incorrect) significantly depends on the collocation (χ^2 (17, N= 2988) = 646.52, $p=.000$) and that the association between the two is moderate (Phi coefficient = .465).

When we cross both variables (educational setting and collocation) in order to see how CLIL and NON-CLIL learners' answers differ according to the collocation (see Figure 6.10. below), we notice that the difference between the groups for all items does not seem crucial. It is more salient for some collocations such as *draw a conclusion*, *hold one's breath*, *lose count*, *seize an opportunity* or *serve a purpose*. CLIL learners gave correct answers for all items, except *put out a fire* and *press charges* but the proportions are close.



Figure 6.10. Proportion of correct and incorrect answers per educational setting and collocation – English

Dutch

Figure 6.11. below shows the number of correct and incorrect answers per item.

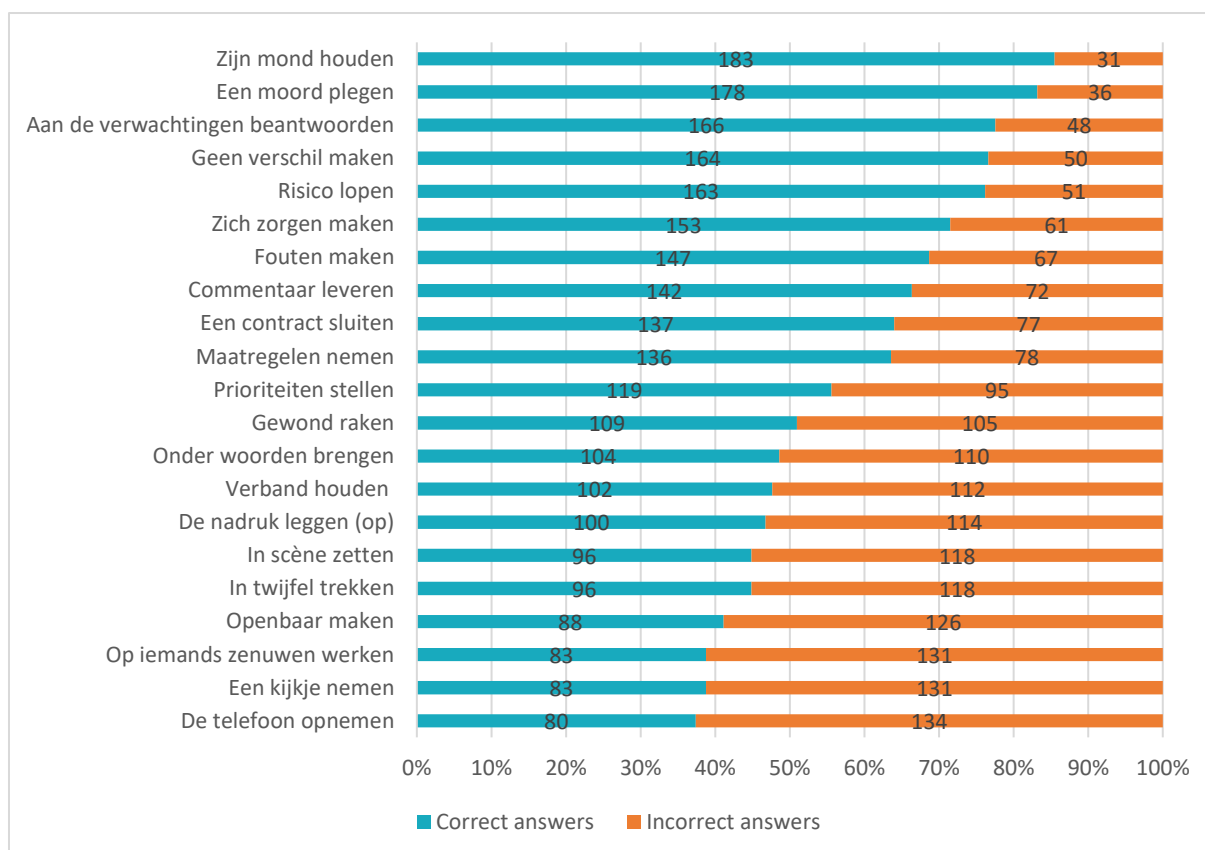


Figure 6.11. Frequency of correct and incorrect answers per collocation – Dutch

The collocation itself appears to have an influence on the results as some collocations seem to be more easily retrieved than others by the learners. Examples of items that are easily recognized are *zijn mond houden* and *een moord plegen*. Other collocations like *een kijkje nemen*, *aan tafel gaan*, *openbaar maken*, *de telefoon opnemen* en *op iemands zenuwen werken* seem to cause more trouble to learners.

A Chi-square test of independence showed that the correctness of the answer (correct or incorrect) is significantly related to the collocation ($\chi^2(21, N=4730) = 513.40, p=.000$) and that the association between the two is moderate (Phi coefficient = .329).

When we cross the variables ‘educational setting’ and ‘collocation’ in order to see how CLIL and NON-CLIL learners’ answers differ according to the collocation (see Figure 6.12. below), we notice that CLIL learners gave correct answers for all items, except for *in twijfel trekken* and *in scène zetten* but the proportions are close. The difference between CLIL/NON-CLIL is more salient for collocations such as *een contract sluiten*, *onder woorden brengen*, *commentaar leveren*, *risico lopen* or *geen verschil maken*, while it is much subtle for items such as *in scène zetten*, *aan tafel gaan* or *op iemands zenuwen werken*.



Figure 6.12. Proportion of correct and incorrect answers per educational setting and collocation – Dutch

The difference between CLIL and NON-CLIL learners seems less crucial than the difference between items. This is why a statistical analysis of the impact of the frequency the collocation is carried out in Section 6.4.2.1., followed by a qualitative analysis of each verb-noun collocation in Section 6.4.2.2.

Preliminary conclusions are the following: 1) CLIL and NON-CLIL learners significantly differ in their ability to retrieve the right collocation – CLIL learners scored better at choosing the correct verb in the verb-noun collocation and 2) the scores also considerably differ according to the collocation – some collocations being easier to retrieve than others.

6.4.2.1. Analysis of the impact of collocation frequency on receptive collocation knowledge

In order to investigate the relation between the frequency of the collocations used in the test and the number of learners being able to recognize the collocations correctly, the normalized frequency of each collocation was calculated using the following reference corpora: the English Web 2015⁴⁵ (enTenTen15) for the English items and the Dutch Web 2014⁴⁶ (nlTenTen14) for the Dutch items. We work on the assumption that the normalized frequency of the collocation in the reference corpus is a proxy for its frequency of occurrence in the learners' TL input.

English

Table 6.9. below shows the normalized frequency (per million words) for each English verb-noun collocation, from the most frequent to the less frequent one.

Table 6.9. Normalized frequency of the collocations used in the English receptive test

Verb + noun collocation	Normalized frequency (per million words) ⁴⁷
Give birth	6.24
Make friends	2.67
Pay a visit	1.33
Hold one's breath	1.18
Give a speech	0.80
Commit a crime	0.53
Do damage	0.37
Do harm	0.36
Press charges	0.33
Hold discussions	0.29
Lose count	0.28
Shed tears	0.26

⁴⁵ Available via Sketch Engine

https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fententen15_tt21 (web corpus of 15,703,895,409 words).

⁴⁶ Available in Sketch Engine

https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fnlntenten14_tt3_1 (web corpus of 2,253,777,579 words).

⁴⁷ The searches were made using the lemma of the verb so that all different forms were taken into account (e.g. *give birth*, *gave birth*, *giving birth*, *gives birth*, etc.).

Serve a purpose	0.22
Keep a secret	0.16
Draw a conclusion	0.08
Seize an opportunity	0.05
Put out a fire	0.05
Exercise rights	0.05

Using a Spearman's rank correlation coefficient that was calculated on the basis of the number of correct answers given to the test (correct=1; incorrect=0) and the normalized frequency of the particular verb-noun collocation in the reference corpus, we found a significant correlation between the learners' knowledge of the English collocations and their normalized frequency ($\rho=0.199^{**}$, $p=.000$, $n=2988$; separated per condition: CLIL $\rho=0.187^{**}$, $p=.000$, $n=1566$ and NON-CLIL $\rho=0.216^{**}$, $p=.000$, $n=1422$). This seems to show that the frequency of the collocations used in the test (calculated on the basis of a reference corpus and used as a proxy for their frequency of occurrence in the language input) has an impact on their receptive acquisition by our participants.

Dutch

Table 6.10. below shows the normalized frequency (per million words) for each Dutch verb-noun collocation from the most frequent to the less frequent one.

Table 6.10. Normalized frequency of the collocations used in the Dutch receptive test

Verb + noun collocation	Normalized frequency (per million words)⁴⁸
Een kijkje nemen	9.05
Zich zorgen maken	6.47
Maatregelen nemen	6.21
Verband houden met	4.31
Fouten maken	4.04
Risico lopen	3.42
Gewond raken	3.41
Openbaar maken	2.63

⁴⁸ The searches were made using the lemma of the verb and taking into account its different positions in the sentence (e.g. *levert commentaar*, but also *commentaar leveren*, *commentaar geleverd*, *commentaar levert*, *commentaar leverde*, etc.).

De nadruk leggen (op)	2.50
Zijn mond houden	1.02
Prioriteiten stellen	0.89
In twijfel trekken	0.87
Onder woorden brengen	0.64
Aan tafel gaan	0.53
Geen verschil maken	0.48
Commentaar leveren	0.37
Een moord plegen	0.33
De telefoon opnemen	0.28
Een contract sluiten	0.28
In scène zetten	0.15
Op iemands zenuwen werken	0.13
Aan de verwachtingen beantwoorden	0.08

In contrast to the results found for English, no significant relation was found between the normalized frequency of the Dutch verb-noun collocations in the reference corpus and the number of correct answers given to the test (neither when taking all the learners together or when looking at the CLIL group and the NON-CLIL group separately). There are probably other factors accounting for the fact that some Dutch collocations are more difficult to retrieve than others, and transfer from the L1 is probably one of them (see next section). One possible explanation might be related to the fact that, contrary to the English test, the items included in the Dutch test were not selected on the basis of their frequency in the language. There is a longer tradition of working with corpora for the establishment of English vocabulary lists. Therefore, there may be more alignment between the input provided in the classroom and the representativeness of the language in English than in Dutch.

6.4.2.2. Qualitative analysis of the collocation in terms of degree of similarity with the L1 (congruence)

We have seen that the level of difficulty with which collocations were retrieved by the learners varies from one collocation to the other and that on the whole, participation in CLIL and the tested items do not seem to correlate. However, some items give rise to more marked contrasts between the two groups.

As mentioned before, various factors may explain why a collocation is easy or difficult to retrieve. Regarding the frequency of the collocation as a whole, the correlations carried out in the previous section showed that it was positively related to the learners' receptive collocation knowledge for English, but not for Dutch. The influence of the mother tongue (in this case, French) is also an important factor (see Chapter 1, Section 1.7.2. for issues of L1 transfer). In fact, if some equivalence exists between the L1 and the TL collocation, the level of difficulty will probably decrease (see Nesselhauf, 2003; Laufer & Waldman, 2011, and Wolter & Gyllstad, 2011, 2013 for their work on congruence).

A qualitative analysis of the results per learner group would help in the interpretation of the results. In order to do so, every tested collocation was assigned a different color according to its apparent level of difficulty (measured using the total number of correct answers).

Color scheme:

item = very easy item (more than 75% of correct answers)

item = easy item (between 50% and 75% of correct answers)

item = difficult item (between 25% and 50% of correct answers)

item = very difficult item (between 0% and 25% of correct answers)

English

Table 6.11. Classification of English items according to the number of correct answers

	Correct answer	Incorrect answer
Commit a crime	159 (95.8%)	7 (4.2%)
Make friends	155 (93.4%)	11 (6.6%)
Keep a secret	149 (89.8%)	17 (10.2%)
Give birth	139 (83.7%)	27 (16.3%)
Give a speech	112 (67.5%)	54 (32.5%)
Hold discussions	111 (66.9%)	55 (33.1%)
Hold one's breath	103 (62%)	63 (38%)
Lose count	95 (57.2%)	71 (42.8%)
Serve a purpose	91 (54.8%)	75 (45.2%)
Exercise rights	84 (50.6%)	82 (49.4%)
Press charges	75 (45.2%)	91 (54.8%)

Seize an opportunity	60 (36.1%)	106 (63.9%)
Pay a visit	57 (34.3%)	109 (65.7%)
Shed tears	54 (32.5%)	112 (67.5%)
Do harm	53 (31.9%)	113 (68.1%)
Draw a conclusion	53 (31.9%)	113 (68.1%)
Put out a fire	45 (27.1%)	121 (72.9%)
Do damage	42 (25.3%)	124 (74.7%)

Very easy items

	CLIL	NON-CLIL	TOTAL
Commit a crime	83 (95%)	76 (97%)	159 (96%)
*Comply a crime	1 (1%)	0 (0%)	1 (1%)
*Conduct a crime	3 (3%)	2 (3%)	5 (3%)
TOTAL	87 (100%)	78 (100%)	165 ⁴⁹ (100%)

	CLIL	NON-CLIL	TOTAL
*Lay birth	3 (4%)	6 (8%)	9 (5%)
Give birth	81 (94%)	58 (74%)	139 (85%)
*Bring birth	2 (2%)	14 (18%)	16 (10%)
TOTAL	86 (100%)	78 (100%)	164 (100%)

	CLIL	NON-CLIL	TOTAL
*Hold a secret	6 (7%)	6 (8%)	12 (7%)
Keep a secret	78 (90%)	71 (90%)	149 (90%)
*Last a secret	3 (3%)	2 (2%)	5 (3%)
TOTAL	87 (100%)	79 (100%)	166 (100%)

	CLIL	NON-CLIL	TOTAL
Make friends	82 (94%)	73 (94%)	155 (94%)
*Create friends	3 (4%)	3 (4%)	6 (4%)

⁴⁹ The total number of answers does not always amount to 166 because some learners gave no answer to one or more items.

*Gain friends	2 (2%)	2 (2%)	4 (2%)
TOTAL	87 (100%)	78 (100%)	165 (100%)

The fact that these collocations were very easily recognized by most learners is probably due to the fact that they have a direct equivalent in French (L1): *commettre un crime, donner naissance, garder un secret, se faire des amis*. On a side note, the collocation *make friends* was accurately produced by a number of CLIL and NON-CLIL learners in their semi-spontaneous written productions. If learners were able to use it actively, it is not surprising to have them recognize it correctly at the receptive test since “evidence shows that receptive knowledge typically precedes productive mastery” (Alsakran, 2011, p. 68).

Easy items

	CLIL	NON-CLIL	TOTAL
*Employ rights	12 (14%)	21 (27%)	33 (20%)
Exercise rights	50 (58%)	34 (44%)	84 (51%)
*Conduct rights	24 (28%)	23 (29%)	47 (29%)
TOTAL	86 (100%)	78 (100%)	164 (100%)

The majority of the learners in both groups chose the correct answer, but the item was more easily retrieved in the CLIL group.

	CLIL	NON-CLIL	TOTAL
Give a speech	58 (69%)	54 (68%)	112 (69%)
*Hold a speech	20 (24%)	9 (12%)	29 (18%)
*Perform a speech	6 (7%)	16 (20%)	22 (13%)
TOTAL	84 (100%)	79 (100%)	163 (100%)

Both groups of learners easily retrieved the target collocation.

	CLIL	NON-CLIL	TOTAL
Hold discussions	66 (77%)	45 (57%)	111 (67%)
*Do discussions	1 (1%)	15 (19%)	16 (10%)

*Set discussions	19 (22%)	19 (24%)	38 (23%)
TOTAL	86 (100%)	79 (100%)	165 (100%)

The collocation proved to be easy for both groups of learners and all the more so for the CLIL ones.

	CLIL	NON-CLIL	TOTAL
*Keep one's breath	14 (16%)	38 (48%)	52 (31%)
*House one's breath	5 (6%)	6 (8%)	11 (7%)
Hold one's breath	68 (78%)	35 (44%)	103 (62%)
TOTAL	87 (100%)	79 (100%)	166 (100%)

The relatively high number of correct answers found for this item is mostly due to the CLIL learners. In fact, most of the NON-CLIL learners chose the pseudo-collocation 'keep one's breath' (not entirely surprising since 'keep' and 'hold' have similar meanings).

	CLIL	NON-CLIL	TOTAL
Lose count	62 (72%)	33 (42%)	95 (58%)
*Drop count	14 (16%)	24 (31%)	38 (23%)
*Pass count	10 (12%)	21 (27%)	31 (19%)
TOTAL	86 (100%)	78 (100%)	164 (100%)

While the item appeared to be relatively easy for the CLIL group, it revealed itself difficult for the NON-CLIL group.

	CLIL	NON-CLIL	TOTAL
*Earn a purpose	16 (19%)	28 (36%)	44 (27%)
*Win a purpose	12 (14%)	16 (20%)	28 (17%)
Serve a purpose	57 (67%)	34 (44%)	91 (56%)
TOTAL	85 (100%)	78 (100%)	163 (100%)

This item seemed particularly easy for CLIL learners, while the NON-CLIL ones had more difficulties.

Most of these ‘easy’ items are equivalent or at least very similar to French collocations: *exercer des droits*, *donner un discours* (although *faire un discours* is more common), *tenir des discussions*, *retenir sa respiration* (very close to *tenir*, literal translation of the verb ‘hold’), *perdre le compte*, *servir un but*. Most of their normalized frequencies are relatively high. Only the collocation *exercise rights* is less frequent (frequency of 0.05), but the fact that it is a congruent collocation may have facilitated its retrieval anyway.

Although the correct collocation was systematically chosen by the majority of learners in both groups (at the exception of ‘hold one’s breath’ where the majority of NON-CLIL learners gave an incorrect answer), the number of learners giving a correct answer was always higher in the CLIL group.

Difficult items

The items labelled as ‘difficult’ can be subdivided into two different categories: a) the cases where the majority of the learners selected a collocation with an incorrect verb collocate, and b) the cases where all three possible answers were chosen in similar proportions. Closer attention is paid to the distractors since more mistakes have been made.

We will first look at the cases where the majority of the learners chose a distractor/pseudo-collocation over the correct verb (when looking at the total proportions).

	CLIL	NON-CLIL	TOTAL
Do damage	23 (26%)	19 (24%)	42 (25%)
*Make damage	51 (59%)	50 (63%)	101 (61%)
*Run damage	13 (15%)	10 (13%)	23 (14%)
TOTAL	87 (100%)	79 (100%)	166 (100%)

The results showed a clear preference for the incorrect collocation ‘make damage’. The direct equivalent in French is *faire des dégâts*. The difficulty undoubtedly lies in the fact that *faire* in French covers the two different meanings of the English verbs ‘do’ and ‘make’, which results in learners mixing up the two (Nehls, 1991), even though the target collocation is relatively frequent. The CLIL and the NON-CLIL learners do not differ much in the way they answered.

	CLIL	NON-CLIL	TOTAL

*Make a conclusion	34 (39%)	60 (77%)	94 (57%)
*Pull a conclusion	4 (5%)	13 (17%)	17 (11%)
Draw a conclusion	48 (56%)	5 (6%)	53 (32%)
TOTAL	86 (100%)	78 (100%)	164 (100%)

When looking at the total proportions, the majority of the learners chose the pseudo-collocation ‘make a conclusion’ over the other two options. This is a bit surprising since the correct English collocation ‘draw a conclusion’ has a direct equivalent in French (*tirer une conclusion*). The learners might be more familiar with the first and more frequent meaning of ‘draw’, that is ‘sketch’ (literally *dessiner* in French) and would therefore reject this option. They may have chosen the verb ‘make’ as a lexical teddy bear whenever they do not know the correct answer (Hasselgren, 1994). This mostly concerns NON-CLIL learners since most of the CLIL learners gave a correct answer.

	CLIL	NON-CLIL	TOTAL
*Do a visit	38 (44%)	53 (67%)	91 (55%)
*Lay a visit	10 (11%)	8 (10%)	18 (11%)
Pay a visit	39 (45%)	18 (23%)	57 (34%)
TOTAL	87 (100%)	79 (100%)	166 (100%)

When looking at the results of the CLIL and NON-CLIL learners taken together, the incorrect collocation ‘do a visit’ was chosen by most learners. There is no direct equivalent in French for ‘pay a visit’ (literally **payer une visite*, the correct collocation being *rendre visite*) but it is possible to say *faire une visite* (literal translation of English ‘do a visit’), although it does not refer to the same activity. This may have confused them, especially as there was no context sentence. Another explanation would be that learners choose the verb ‘do’, which is a high frequency verb like ‘make’, when they do not know the answer. The high proportion of answers found for the distractor is mostly due to the NON-CLIL learners.

	CLIL	NON-CLIL	TOTAL
*Pull an opportunity	4 (4%)	6 (8%)	10 (6%)
Seize an opportunity	45 (52%)	15 (19%)	60 (36%)
*Catch an opportunity	38 (44%)	58 (73%)	96 (58%)

TOTAL	87 (100%)	79 (100%)	166 (100%)

The learners showed a clear preference for the pseudo-collocation ‘catch an opportunity’ although the target collocation ‘seize an opportunity’ has a direct equivalent in French: *saisir une opportunité*. However, the English verbs ‘seize’ and ‘catch’ have very similar meanings and can both be used in figurative sense, but ‘seize’ is a low frequency verb. Therefore, learners may have opted for the most frequent one. This is mostly the case for NON-CLIL learners, the majority of CLIL learners having chosen the correct collocation.

In sum, we can say that except for the collocation ‘pay a visit’ where there is no direct equivalent in French, all other verbs could have correctly been chosen thanks to a positive transfer from the learners’ mother tongue. However, learners may have been less familiar with the more restricted senses of the verb collocates ‘draw’ and ‘pay’ whose frequency is lower and fell back on the high frequency verbs ‘do’ and ‘make’. They also preferred the more frequent verb ‘catch’ over the less frequent ‘seize’, which is all the more understandable that the two verbs have synonymous meanings. These explanations especially concern NON-CLIL learners since most of them chose pseudo-collocations over the correct answer.

We will now turn to the cases where the three possible options were chosen in similar proportions (when looking at the total proportions).

	CLIL	NON-CLIL	TOTAL
*Take harm	29 (35%)	40 (51%)	69 (42%)
Do harm	34 (41%)	19 (24%)	53 (33%)
*Make harm	20 (24%)	20 (25%)	40 (25%)
TOTAL	83 (100%)	79 (100%)	162 (100%)

The English target collocation ‘do harm’ has a direct equivalent in French (*faire du mal*) but since the verbs ‘do’ and ‘make’ are both translated into *faire* in French, learners probably hesitated between the two. The word ‘harm’ is not very frequent and learners may have found it difficult to understand the collocation without context. The majority of the CLIL learners correctly retrieved the item.

	CLIL	NON-CLIL	TOTAL
Press charges	35 (40.5%)	40 (51%)	75 (46%)

*Run charges	35 (40.5%)	16 (21%)	51 (31%)
*Push charges	16 (19%)	22 (28%)	38 (23%)
TOTAL	86 (100%)	78 (100%)	164 (100%)

Overall, the learners did not express a clear choice, even though a higher number of answers was given for the correct collocation. No direct equivalent can be found in French (*porter plainte*) and the meaning of the verb collocate is far from being transparent since it is used in a more technical sense. Moreover, the target collocation corresponds to a C2 CEFR level and the absence of context may also have hindered the interpretation of the collocation.

	CLIL	NON-CLIL	TOTAL
*Turn out a fire	18 (21%)	29 (37%)	47 (29%)
Put out a fire	17 (19%)	28 (36%)	45 (27%)
*Set out a fire	52 (60%)	21 (27%)	73 (44%)
TOTAL	87 (100%)	78 (100%)	165 (100%)

The correct collocation ‘put out a fire’ has a direct French equivalent: *éteindre un feu*. However, the verbs ‘put out’, ‘turn out’ and ‘set out’ are all phrasal verbs, which makes their meaning harder to grasp for non-native speakers (Gilquin, 2009; Gilquin, 2015). Also, they may have confused ‘turn out’ with ‘turn off’ in the existing collocation ‘turn off the light’ (literally *éteindre la lumière* in French). The hesitation with ‘set out’ may be due to the existence of the verb-noun combination ‘set on fire’, an option chosen by most of the CLIL learners.

	CLIL	NON-CLIL	TOTAL
*Fell tears	24 (28%)	26 (33%)	50 (30%)
Shed tears	36 (41%)	18 (23%)	54 (33%)
*Raise tears	27 (31%)	34 (44%)	61 (37%)
TOTAL	87 (100%)	78 (100%)	165 (100%)

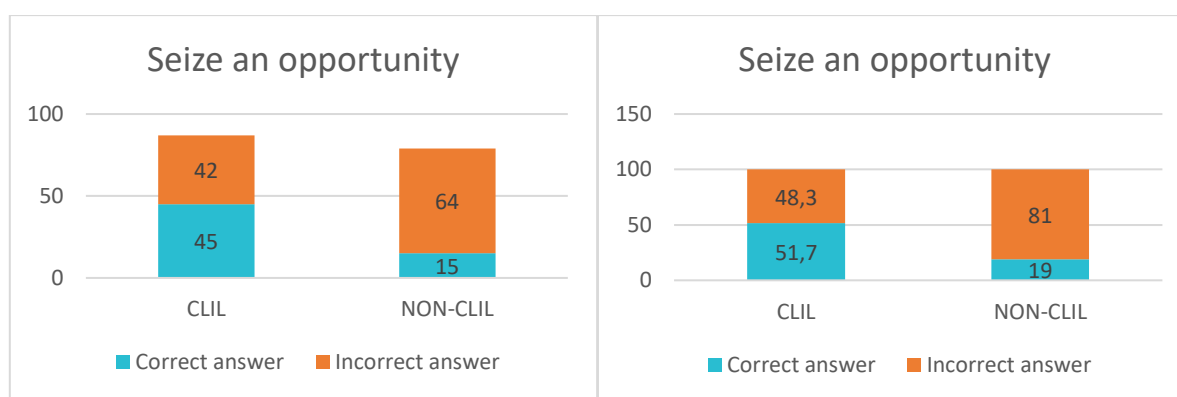
The English collocation ‘shed tears’ and the French equivalent *verser des larmes* (literally ‘pour tears’) do not correspond. The learners probably did not know the verb ‘shed’ – which is a low frequency verb – and some of them may have been more familiar with the meaning of the verb ‘raise’ in the sense of ‘provoke’. They also hesitated with ‘fell tears’, even though the verb ‘fell’ has an even lower frequency than ‘shed’. It may be that learners thought ‘fell’ was the past tense

of the verb ‘fall’ and knew that ‘tears’ collocates with ‘fall’ in the noun-verb collocation ‘tears fall’. We did not expect the learners to have acquired the target collocation anyway, since it corresponds to a C1 CEFR level. Most of the CLIL learners chose the correct item, though.

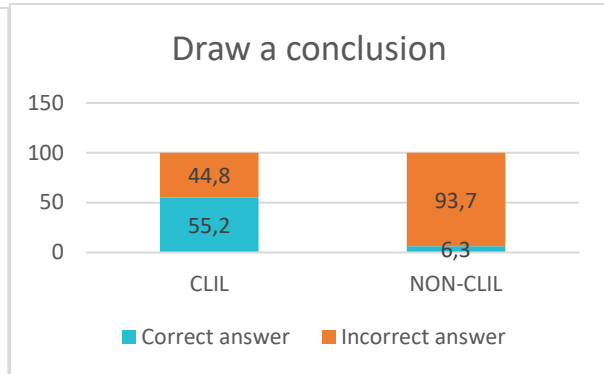
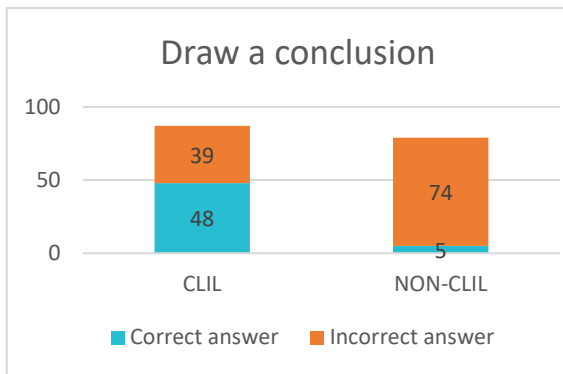
In all these cases, equivalences between English and French were less transparent, except for *put out a fire* and *do harm*. The collocation *put out a fire* has a low frequency (0.05) and literature has shown that learners seem to have trouble dealing with phrasal verbs. The collocation *do harm* is relatively frequent (0.36), but delexical verbs such as ‘do’ – which does not carry a specific meaning – are difficult for learners, especially for French-speaking ones since French does not make the distinction between ‘do’ and ‘make’. Learners probably did not know the collocations *press charges* and *shed tears*, which is not very surprising since they correspond to higher levels of proficiency and include less frequent items.

Comparing the CLIL and the NON-CLIL learners’ answers showed that these items were (very) difficult for both groups and that the case of *press charges* is a peculiar one since over the half of NON-CLIL learners correctly retrieved it.

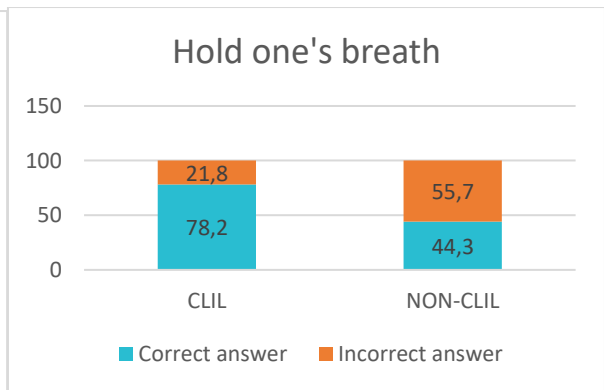
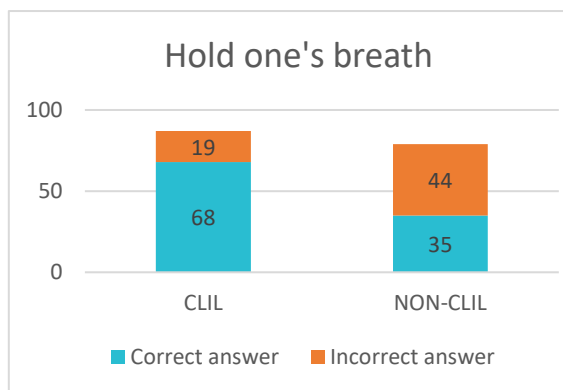
As previously mentioned in the quantitative analysis, the contrast between CLIL and NON-CLIL learners seemed more salient for the following collocations: *seize an opportunity*, *draw a conclusion*, *hold one’s breath*, *lose count* and *serve a purpose*. In fact, CLIL learners gave a higher number of correct answers for these items compared to NON-CLIL learners. The figures below show the distribution of correct and incorrect answers (in absolute frequencies on the left and in proportions on the right) per learner group for each of these items.



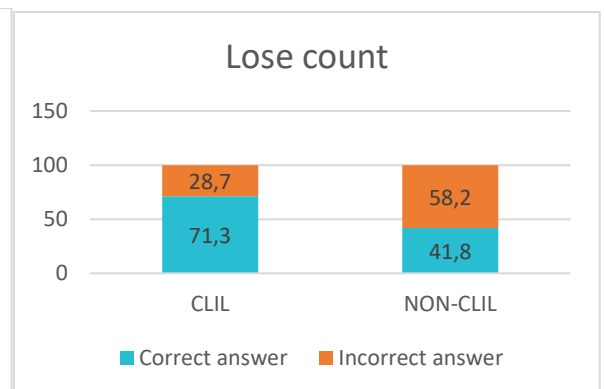
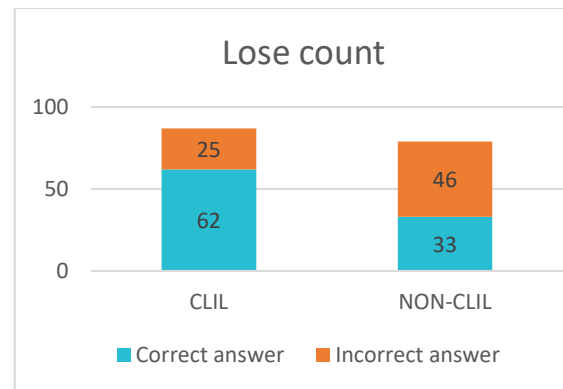
Difficult item



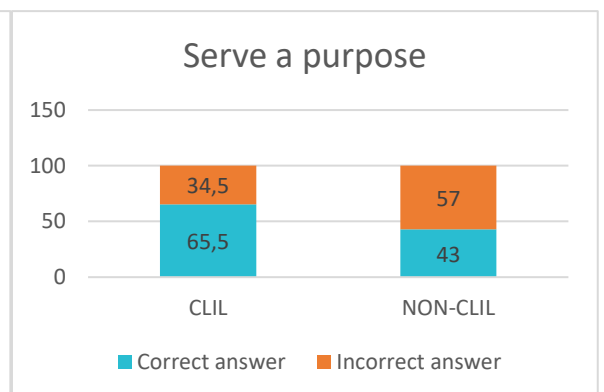
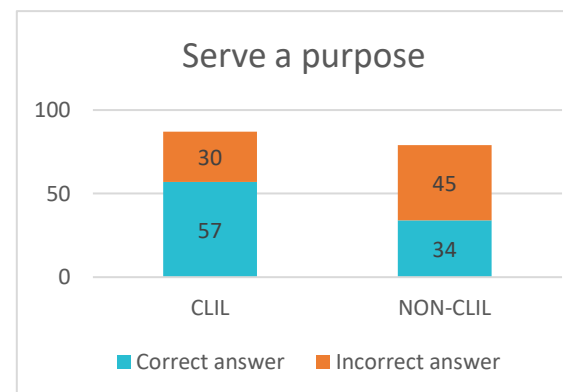
Difficult item



Easy item



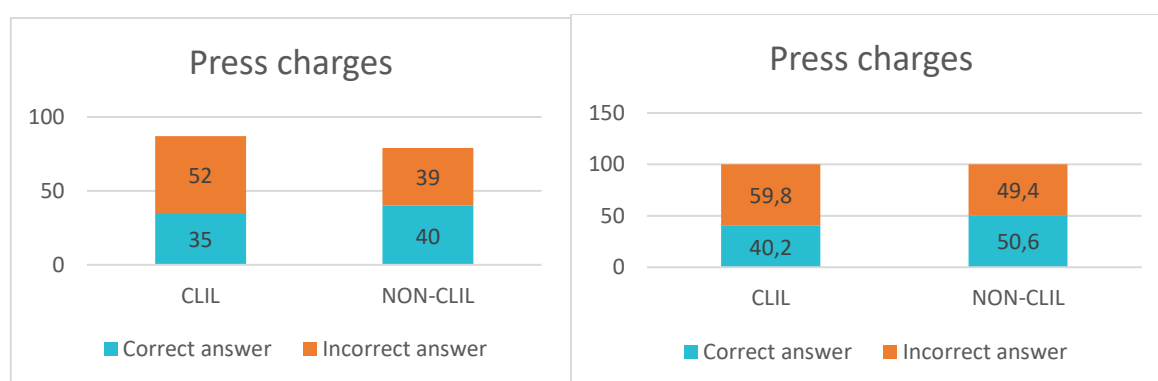
Easy item



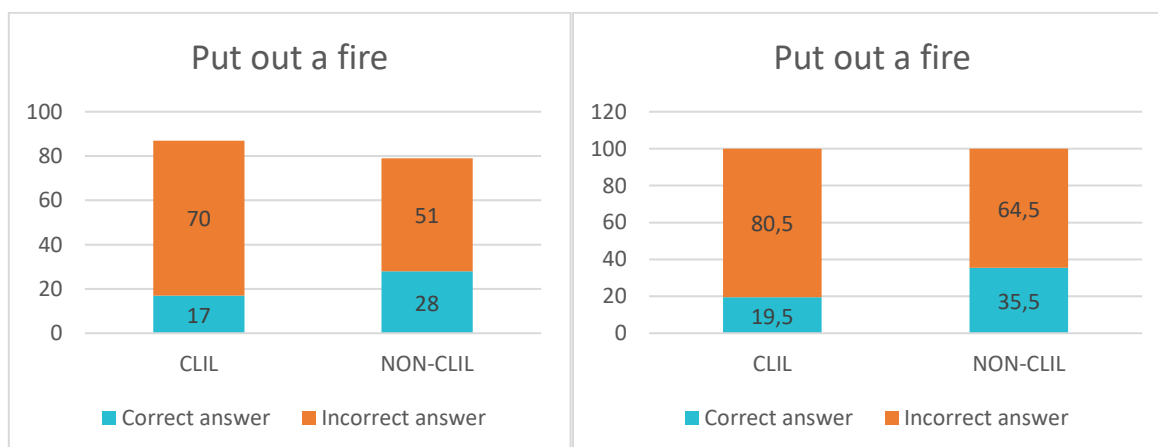
Easy item

As far as the ‘difficult’ items are concerned, CLIL learners may be more aware of the various and less frequent meanings of the verbs ‘seize’ and ‘draw’ thanks to the greater amount of input they received, especially since these collocations were found to be infrequent (normalized frequency of 0.05 for *seize an opportunity* and 0.08 for *draw a conclusion*). Regarding the ‘easy’ items, CLIL learners seem to make an even more positive use of their mother tongue when a direct, transparent equivalent in French exists.

The two cases below concern collocations for which NON-CLIL learners gave a higher number of correct answers: *press charges* and *put out a fire*. The difference is not that large, though, as can be seen in the figures below.



Difficult item



Difficult item

These results are rather peculiar because they involve ‘difficult’ items. This may be due to the fact that specific topics related to these collocations were covered in class with NON-CLIL learners or that the items were explicitly taught to them. NON-CLIL learners may also have accidentally guessed more often correctly than the CLIL learners.

Table 6.12. Classification of Dutch items according to the number of correct answers

	Correct answer	Incorrect answer
Zijn mond houden	183 (85.5%)	31 (14.5%)
Een moord plegen	178 (83.2%)	36 (16.8%)
Aan de verwachtingen beantwoorden	166 (77.6%)	48 (22.4%)
Geen verschil maken	164 (76.6%)	50 (23.4%)
Risico lopen	163 (76.2%)	51 (23.8%)
Zich zorgen maken	153 (71.5%)	61 (28.5%)
Fouten maken	147 (68.7%)	67 (31.3%)
Commentaar leveren	142 (66.4%)	72 (33.6%)
Een contract sluiten	137 (64%)	77 (36%)
Maatregelen nemen	136 (63.6%)	78 (36.4%)
Prioriteiten stellen	119 (55.6%)	95 (44.4%)
Gewond raken	109 (50.9%)	105 (49.1%)
De nadruk leggen (op)	100 (46.7%)	114 (53.3%)
Onder woorden brengen	104 (48.6%)	110 (51.4%)
Verband houden	102 (47.7%)	112 (52.3%)
In twijfel trekken	96 (44.9%)	118 (55.1%)
In scène zetten	96 (44.9%)	118 (55.1%)
Openbaar maken	88 (41.1%)	126 (58.9%)
Een kijkje nemen	83 (38.8%)	131 (61.2%)
Op iemands zenuwen werken	83 (38.8%)	131 (61.2%)
De telefoon opnemen	80 (37.4%)	134 (62.6%)
Aan tafel gaan	63 (29.4%)	151 (70.6%)

Very easy items

	CLIL	NON-CLIL	TOTAL
*Risico hebben	4 (3%)	15 (19%)	19 (9%)
Risico lopen	120 (91%)	43 (54%)	163 (77%)

*Risico krijgen	8 (6%)	21 (27%)	29 (14%)
TOTAL	132 (100%)	79 (100%)	211 (100%)

The collocation was particularly very easy for the CLIL learners, in contrast to NON-CLIL learners who had a bit more difficulty.

	CLIL	NON-CLIL	TOTAL
Zijn mond houden	129 (98%)	54 (67%)	183 (86%)
*Zijn mond doen	2 (1%)	15 (18%)	17 (8%)
*Zijn mond maken	1 (1%)	12 (15%)	13 (6%)
TOTAL	132 (100%)	81 (100%)	213 (100%)

An incredibly high number of correct answers was observed in the CLIL group. The item was also easy for the NON-CLIL learners but to a lesser extent.

	CLIL	NON-CLIL	TOTAL
*Aan de verwachtingen denken	4 (3%)	16 (20%)	20 (10%)
Aan de verwachtingen beantwoorden	112 (85%)	54 (67.5%)	166 (78%)
*Aan de verwachtingen dienen	16 (12%)	10 (12.5%)	26 (12%)
TOTAL	132 (100%)	80 (100%)	212 (100%)

The target collocation was easy to retrieve for both groups of learners, but all the more so for the CLIL participants.

	CLIL	NON-CLIL	TOTAL
*Geen verschil hebben	7 (5%)	22 (27%)	29 (13%)
*Geen verschil raken	6 (5%)	15 (19%)	21 (10%)
Geen verschil maken	120 (90%)	44 (54%)	164 (77%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

The item was very easily retrieved in the CLIL group, which partly explains the high number of correct answers when taking the two groups together.

	CLIL	NON-CLIL	TOTAL
Een moord plegen	126 (95%)	52 (65%)	178 (84%)
*Een moord maken	3 (2%)	14 (17.5%)	17 (8%)
*Een moord doen	4 (3%)	14 (17.5%)	18 (8%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

The collocation *een moord plegen* proved to be very easy for CLIL learners and fairly easy for NON-CLIL learners.

We observed a large number of correct answers for these items, which is probably explained by the fact that most of these collocations have a direct equivalent in French, the learners' mother tongue: *courir un risque*, *répondre aux attentes*, *ne pas faire de différence* and *commettre un meurtre*. The French equivalent of *zijn mond houden* is less transparent, although the verb 'houden' is a direct translation of *tenir* as in the French expression *tenir sa langue* (which is one possible translation of 'mond houden'). It is also possible that the options 'zijn mond doen' and 'zijn mond maken' were excluded by the learners because the verbs 'doen' and 'maken' are too unspecific to render the signification.

As for the comparison between the CLIL and the NON-CLIL group, we saw that these items turned out to be very easy for the learners in CLIL and rather easy for the learners in NON-CLIL.

Easy items

	CLIL	NON-CLIL	TOTAL
*Commentaar bezorgen	8 (6%)	18 (22%)	26 (12%)
Commentaar leveren	117 (88%)	25 (31%)	142 (66%)
*Commentaar praten	8 (6%)	38 (47%)	46 (22%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

The relatively high number of correct answers found for this item is mostly due to the CLIL learners. Almost half of the NON-CLIL learners selected the distractor 'commentaar praten'.

	CLIL	NON-CLIL	TOTAL

Gewond raken	80 (61%)	29 (37%)	109 (52%)
*Gewond worden	36 (27%)	26 (33%)	62 (29%)
*Gewond komen	16 (12%)	24 (30%)	40 (19%)
TOTAL	132 (100%)	79 (100%)	211 (100%)

While the item appeared to be relatively easy for the CLIL learners, it revealed itself difficult for the NON-CLIL learners, their answers being similarly distributed across the three possible answers.

	CLIL	NON-CLIL	TOTAL
*Fouten zetten	1 (1%)	1 (1%)	2 (1%)
Fouten maken	98 (74%)	49 (61%)	147 (69%)
*Fouten doen	34 (25%)	31 (38%)	65 (30%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

More than 50% of the learners in both groups chose the target collocation. This collocation is probably frequently used in the school context and therefore familiar to the learners. However, some learners – and especially NON-CLIL ones – went for ‘fouten doen’, a mistake probably induced by the learners’ L1 which only has one verb (*faire*) for ‘maken’ and ‘doen’.

	CLIL	NON-CLIL	TOTAL
*Een contract dichten	13 (10%)	19 (23%)	32 (15%)
Een contract sluiten	104 (78%)	33 (41%)	137 (64%)
*Een contract maken	16 (12%)	29 (36%)	45 (21%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

This item revealed to be very easy for the CLIL learners, but difficult for their NON-CLIL counterparts.

	CLIL	NON-CLIL	TOTAL
*Zich zorgen doen	10 (7%)	13 (17.5%)	23 (11%)
Zich zorgen maken	105 (79%)	48 (60%)	153 (72%)
*Zich zorgen raken	18 (14%)	18 (22.5%)	36 (17%)

TOTAL	133 (100%)	79 (100%)	212 (100%)

The target collocation was very easily retrieved by the CLIL group and seemed to be an easy item for NON-CLIL learners. Note that a higher number of NON-CLIL learners chose the pseudo-collocation ‘zich zorgen doen’, which again suggests that they have more trouble distinguishing between the verbs ‘maken’ and ‘doen’ due to L1 influence.

	CLIL	NON-CLIL	TOTAL
Maatregelen nemen	94 (71%)	42 (52%)	136 (64%)
*Maatregelen maken	3 (2%)	9 (11%)	12 (6%)
*Maatregelen stellen	36 (27%)	29 (37%)	65 (30.5%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

The majority of the learners in both groups chose the correct verb collocate, although the CLIL group seemed to be more familiar with the collocation.

	CLIL	NON-CLIL	TOTAL
*Prioriteiten maken	15 (11%)	19 (24%)	34 (16%)
*Prioriteiten geven	33 (25%)	27 (33%)	60 (28%)
Prioriteiten stellen	85 (64%)	34 (43%)	119 (56%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

This item turned out to be easy for the CLIL learners only, since not even the half of NON-CLIL learners correctly retrieved the collocation.

Most of these items have a direct equivalent in French : *faire des erreurs*, *se faire du souci*, *prendre des mesures*, *définir les priorités*. This is however not the case for *een contract sluiten* (literally **fermer un contrat*, which does not exist), *commentaar leveren* (literally **fournir un/des commentaire(s)*, while the French equivalent is usually made of one single verb, namely *critiquer* or *commenter*) and *gewond raken* (literally **devenir blessé*, while the corresponding collocation is *être blessé*).

Even though the majority of the learners in both groups opted for the correct collocation (except in the case of *maatregelen nemen* where most of the NON-CLIL learners chose a pseudo-collocation), a higher ability was observed in the CLIL group.

Difficult items

The items labelled as ‘difficult’ can be subdivided into two different categories: a) the cases where most learners chose the target collocation but not in a prominent way, and b) the cases where most learners chose a pseudo-collocation over the collocation with the correct verb. Closer attention is paid to the distractors since more mistakes have been made.

We will start with the cases where learners chose the correct answer among the three options, although not prominently (when looking at the total proportions).

	CLIL	NON-CLIL	TOTAL
*De nadruk zeggen (op)	4 (3%)	36 (45%)	40 (19%)
*De nadruk zetten (op)	54 (41%)	18 (24%)	72 (34%)
De nadruk leggen (op)	75 (56%)	25 (31%)	100 (47%)
TOTAL	133 (100%)	79 (100%)	212 (100%)

The collocations in French and Dutch are equivalent (*mettre l’accent (sur)* is a direct translation of *nadruk leggen (op)*). However, the Dutch verbs ‘zetten’ and ‘leggen’ are closely related and both correspond to the verb ‘mettre’ in French (see e.g. Lemmens, 2002 on the difficulty of ‘posture’ verbs for French-speaking learners of Dutch) which might have confused some learners, especially the CLIL ones. Most of the NON-CLIL learners selected the distractor ‘nadruk zeggen’ (literally ‘say the emphasis’), which is hard to explain – they probably did not know the collocation.

	CLIL	NON-CLIL	TOTAL
In twijfel trekken	55 (42%)	41 (51%)	96 (45%)
*In twijfel duwen	31 (24%)	17 (21%)	48 (23%)
*In twijfel schoppen	45 (34%)	23 (28%)	68 (32%)
TOTAL	131 (100%)	81 (100%)	212 (100%)

The Dutch target collocation does not have a direct French equivalent (literally **tirer en doute*) and the two pseudo-collocations could not possibly have misled the learners due to a potential similarity with their mother tongue (**pousser en doute*, **botter en doute*). Most of the CLIL and NON-CLIL learners chose the correct verb collocate, with a bit more correct answers in the NON-CLIL group.

	CLIL	NON-CLIL	TOTAL
*Een kijkje zien	17 (13%)	35 (44%)	52 (24%)
*Een kijkje doen	56 (42%)	22 (27%)	78 (37%)
Een kijkje nemen	60 (45%)	23 (29%)	83 (39%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

The Dutch collocation *een kijkje nemen* and the corresponding French collocation *jeter un œil* are not equivalent since ‘nemen’ would be translated as *prendre*. The two distractors do not share similarities with the learners’ L1 either, but ‘zien’ could have misled the learners because it is semantically compatible with the noun. The two learner groups went for different answers, with the CLIL learners being more successful.

	CLIL	NON-CLIL	TOTAL
Onder woorden brengen	82 (62%)	22 (27.5%)	104 (49%)
*Onder woorden zeggen	50 (37%)	50 (62.5%)	100 (47%)
*Onder woorden maken	1 (1%)	8 (10%)	9 (4%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

The target collocation *onder woorden brengen* would most probably be translated by a single verb in French: *formuler* or *exprimer*. A possible equivalent in French that would resemble the Dutch collocation is *mettre en mots* (*mettre* being one of the meanings of ‘brengen’), although it does not sound very natural. Many NON-CLIL learners chose ‘onder woorden zeggen’ (literally *dire en mots*, which probably does not exist, but has a similar meaning to the French single verb *formuler*). There is also semantic compatibility between the noun ‘woorden’ and the verb ‘zeggen’ in this case, and this may have deceived the learners.

	CLIL	NON-CLIL	TOTAL
*In scène maken	39 (29%)	16 (20%)	55 (26%)
*In scène doen	35 (26%)	28 (34%)	63 (29%)
In scène zetten	59 (45%)	37 (46%)	96 (45%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

The Dutch and French collocations are equivalent: *in scène zetten* and *mettre en scène* are direct translations of each other. The wrong answers might be due to a lack of understanding of the

context combined to the fact that the verbs ‘doen’ and ‘maken’ both refer to *faire* in French as in the existing *faire de la scène*. The way CLIL and NON-CLIL learners answered to the test did not seem to differ much.

	CLIL	NON-CLIL	TOTAL
Op iemands zenuwen werken	54 (41%)	29 (36%)	83 (39%)
*Op iemands zenuwen drukken	47 (36%)	31 (39%)	78 (37%)
*Op iemands zenuwen lopen	31 (23%)	20 (25%)	51 (24%)
TOTAL	132 (100%)	80 (100%)	213 (100%)

The collocations are not congruent, as the Dutch collocation would be translated by **travailler sur les nerfs* instead of *taper sur les nerfs*. The other two options sound more or less acceptable in French (*peser sur les nerfs* and *courir sur les nerfs*), so the mistakes might be induced by the L1. The proportions of answers given to each item are quite similar in the two groups of learners.

	CLIL	NON-CLIL	TOTAL
*Verband doen	13 (10%)	19 (23%)	32 (15%)
*Verband hangen	47 (36%)	32 (39%)	79 (37%)
Verband houden	72 (54%)	30 (38%)	102 (48%)
TOTAL	132 (100%)	81 (100%)	213 (100%)

The collocations in the two languages are not similar (*verband houden*, literally **tenir en lien* does not correspond to the French *être lié à*). ‘Verband hangen’ does not seem to correspond to something equivalent in French either (**pendre en lien/rapport avec*) and yet a number of CLIL and NON-CLIL learners chose that option.

We will now turn to the cases where the majority of the learners chose an incorrect verb collocate over the correct one (when looking at the total proportions).

	CLIL	NON-CLIL	TOTAL
*De telefoon ophalen	25 (19%)	14 (19%)	39 (18%)
De telefoon opnemen	66 (50%)	14 (17%)	80 (38%)
*De telefoon beantwoorden	42 (31%)	52 (64%)	94 (44%)
TOTAL	133 (100%)	80 (100%)	213 (100%)

The fact that most of the learners chose the distractor ‘de telefoon beantwoorden’ is not really surprising since it corresponds to the French *répondre au téléphone*. ‘Opnemen’ and ‘beantwoorden’ share the same French translation *répondre*, but ‘beantwoorden’ does not collocate with the noun ‘telefoon’ and learners did not seem to be aware of that restriction, especially the NON-CLIL ones.

	CLIL	NON-CLIL	TOTAL
*Aan tafel staan	46 (34.5%)	21 (26%)	67 (31%)
*Aan tafel zetten	46 (34.5%)	38 (47%)	84 (39%)
Aan tafel gaan	41 (31%)	22 (27%)	63 (30%)
TOTAL	133 (100%)	81 (100%)	214 (100%)

The Dutch target collocation *aan tafel gaan* does not have a direct French equivalent (*aller à table* does not sound very natural). Most learners chose the pseudo-collocation ‘aan tafel zetten’. It may be that they were misled by the meaning of ‘zetten’ (literally *mettre* in French as in *se mettre à table*, the French equivalent for *aan tafel gaan*). They may also have been confused by the existing Dutch collocation ‘de tafel zetten’ (literally ‘set the table’). The tendencies were quite similar in both groups.

	CLIL	NON-CLIL	TOTAL
Openbaar maken	59 (44%)	29 (37%)	88 (41.5%)
*Openbaar hangen	20 (15%)	15 (19%)	35 (16.5%)
*Openbaar leggen	54 (41%)	35 (44%)	89 (42%)
TOTAL	133 (100%)	79 (100%)	212 (100%)

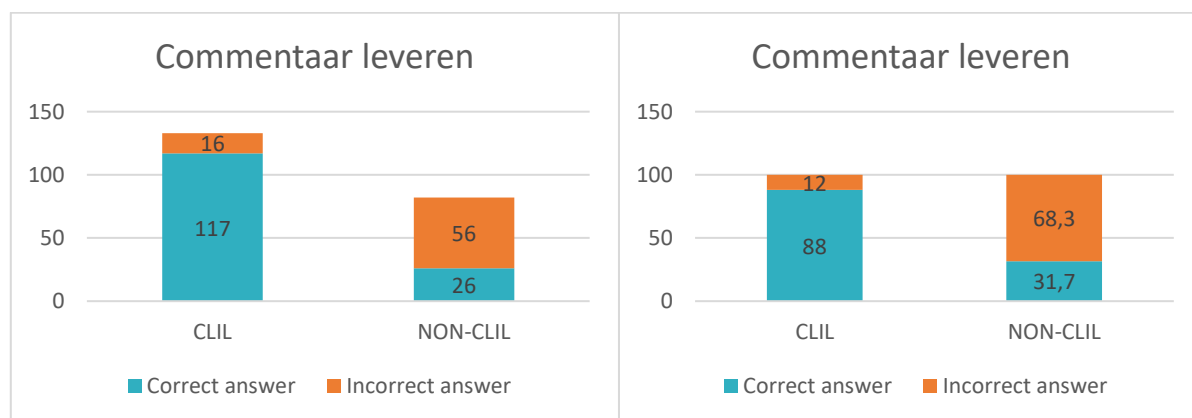
The correct collocation *openbaar maken* does not have a direct equivalent in French (literally **faire public* instead of *rendre public*) and neither do the other two options. The distractor ‘openbaar leggen’, literally *mettre public*, seems however more likely for a French native speaker and most of the NON-CLIL learners selected that item.

The items discussed in this section turned out to be difficult for the learners – they either correctly chose the target collocation but not in a prominent way, or failed to retrieve the correct item. The decrease in correct answers is undoubtedly due to the fact that the majority of the collocations are non-congruent, i.e. they do not have a direct equivalent in the learners’ mother tongue. Nevertheless, the case of *in scène zetten* showed that pupils did not always go for the

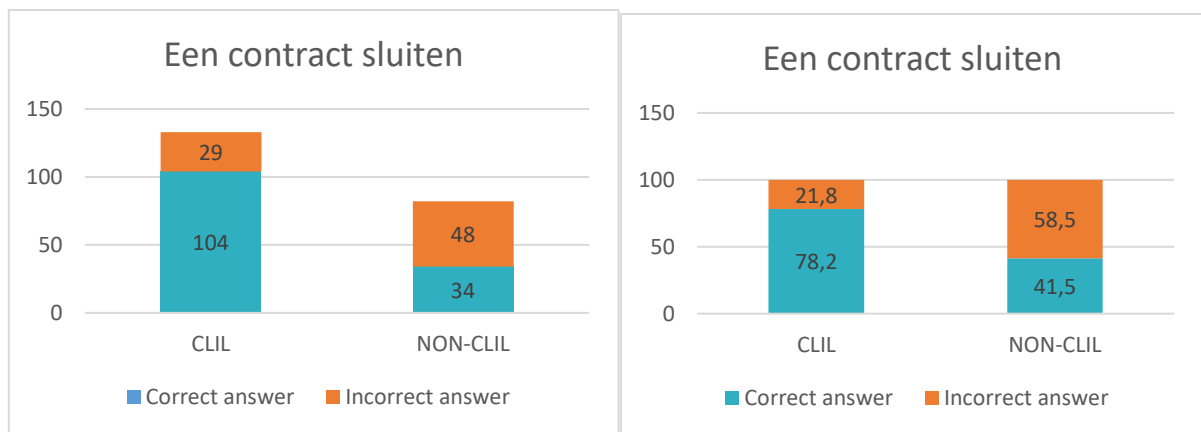
French equivalent when it was available – maybe because they did not know the Dutch collocation and thought it was too easy. The learners' L1 was clearly the cause of mistakes in the case of *de telefoon opnemen*. Regarding the items *op iemands zenuwen werken*, *onder woorden brengen*, and *aan tafel gaan*, one of the other options (or the other two) shared more similarities with French than the correct Dutch collocation. Synonymy between the Dutch verbs 'zetten' en 'leggen' probably confused the learners in the case of *de nadruk leggen (op)*. No clear explanation can be given for all the other difficult items – maybe pupils did not know the collocation and/or did not understand the context sentence and answered at random or chose what they thought sounded the most Dutch-like.

Most of the CLIL learners correctly retrieved the collocations labelled as 'difficult' in nine cases out of ten compared to two cases out of ten for the NON-CLIL group.

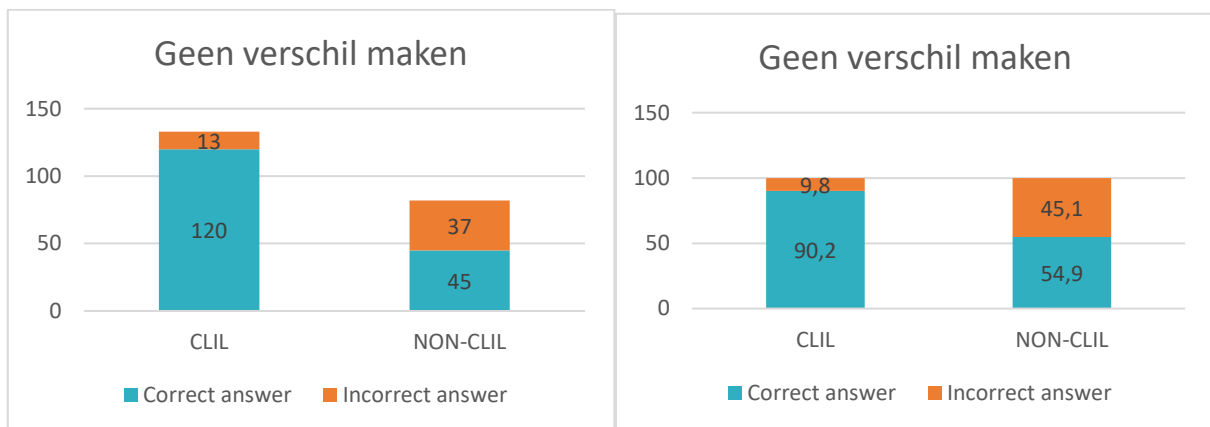
As previously mentioned in the quantitative analysis, the contrast between CLIL and NON-CLIL learners seemed more salient for the following collocations: *commentaar leveren*, *een contract sluiten*, *geen verschil maken* en *risico lopen*. In fact, CLIL learners gave a higher number of correct answers for these items compared to NON-CLIL learners. The figures below show the distribution of correct and incorrect answers (in absolute frequencies on the left and in proportions on the right) per learner group for each of these items.



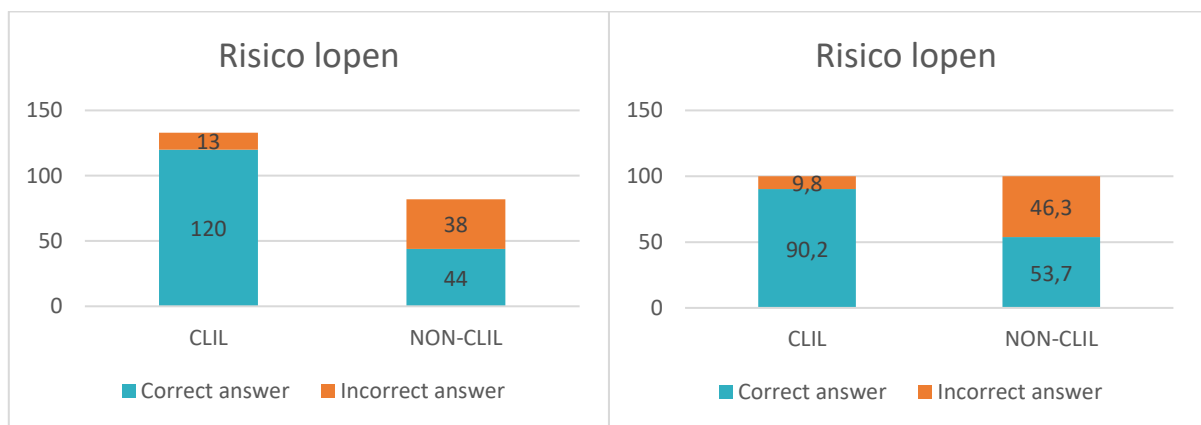
Easy item



Easy item



Very easy item

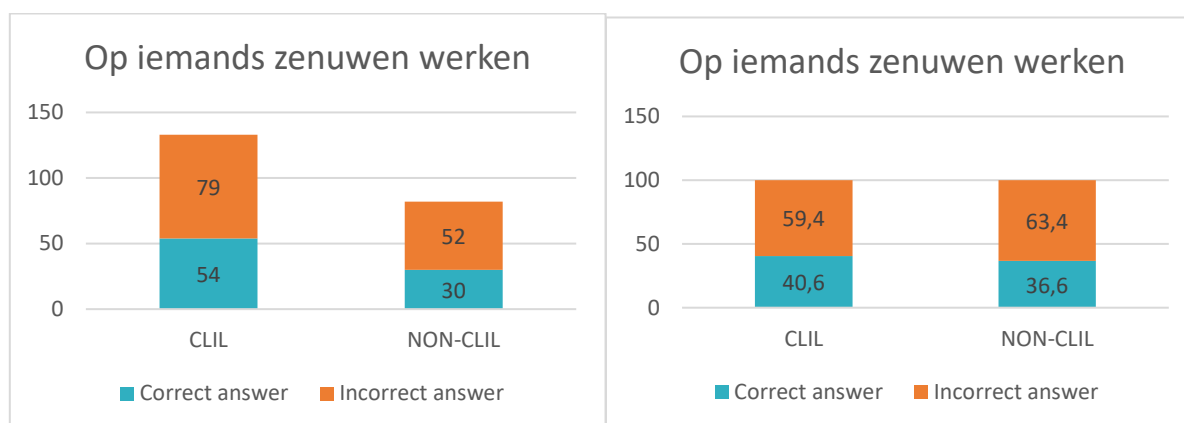


Very easy item

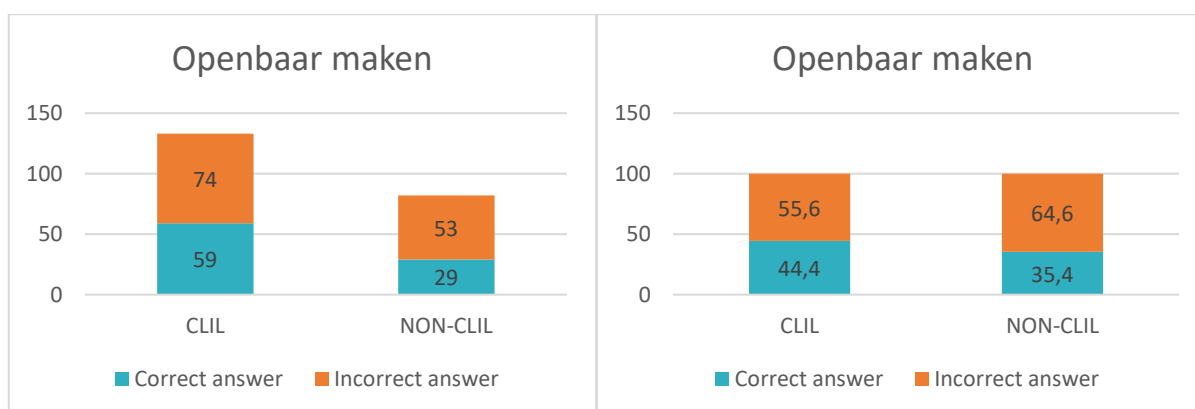
Among the items that seemed easy on the whole, some of them turned out to be very easy for the CLIL group, but actually difficult for the NON-CLIL group (e.g. *commentaar leveren* and *een contract sluiten*). These items do not have equivalent French collocations and might be easier for CLIL learners to retrieve thanks to the larger TL input they received. Some items were easy for all, but even more so for the CLIL learners (e.g. *risico lopen* and *een verschil*

maken). In this case, the Dutch collocations have a direct French equivalent; this seems to indicate that CLIL learners might be more aware of the congruence between Dutch and French.

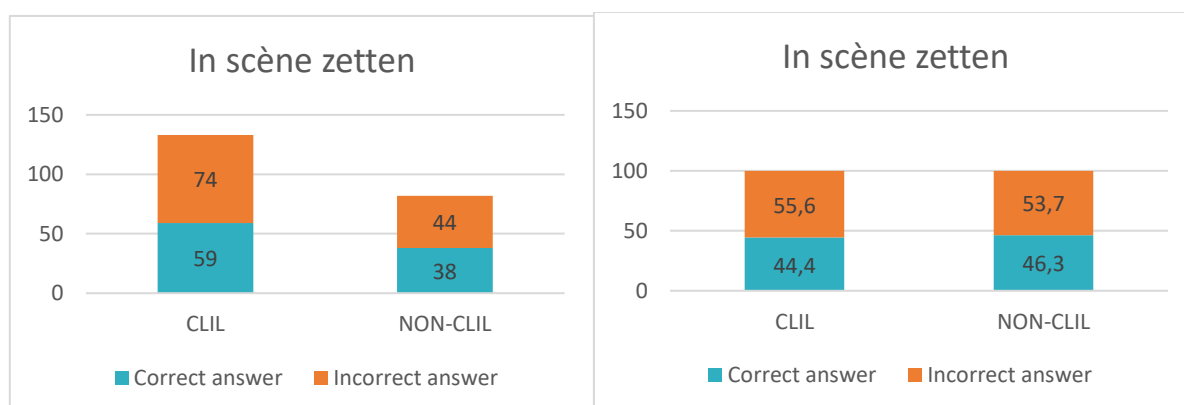
The cases below concern collocations that seem difficult for both CLIL and NON-CLIL learners: *op iemands zenuwen werken*, *openbaar maken*, *in scène zetten*, *aan tafel gaan* and *een kijkje nemen*. Frequencies and proportions are given in the figures below.



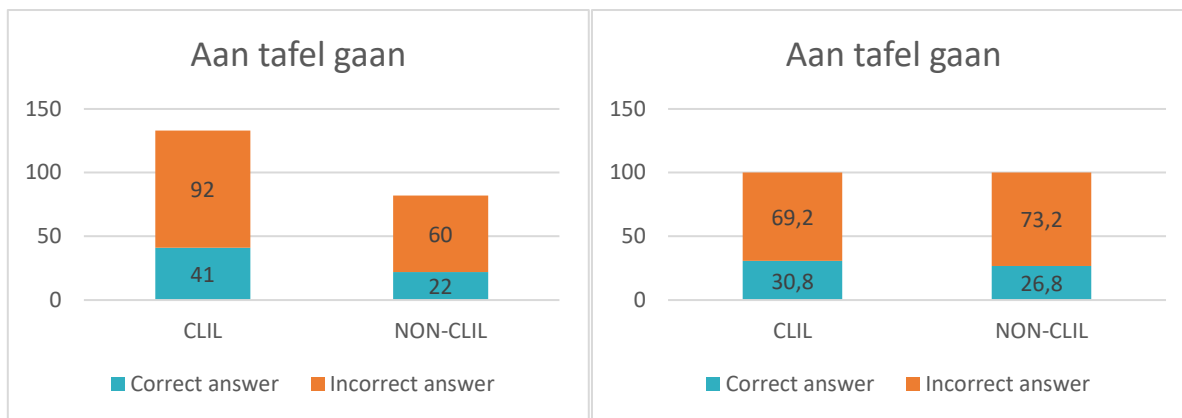
Difficult item



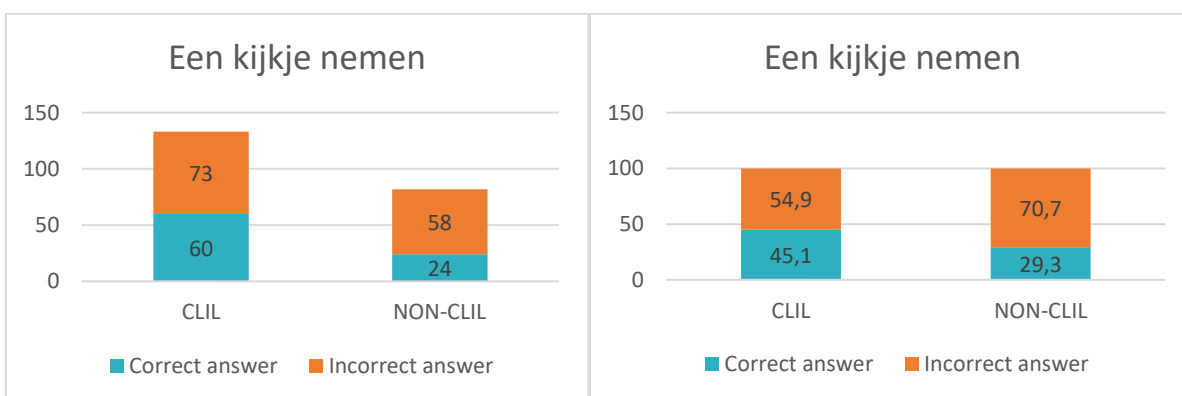
Difficult item



Difficult item



Difficult item



Difficult item

All these items seem to be difficult for both groups of learners. The case of *in scène zetten* is particularly surprising since it has a direct French translation. The other four collocations involve verbs that are undoubtedly familiar to the learners – ‘gaan’, ‘nemen’, ‘maken’ and ‘werken’ are frequent verbs – but the fact that they are here used in a more restricted sense with no clear meaning probably makes their retrieval more problematic.

6.5. Concluding remarks and discussion

This study’s aim was to further refine the comparison between CLIL and NON-CLIL learners of English and Dutch regarding their phraseological competence. The learners’ productive knowledge of phraseological units had already been analyzed in a previous chapter (Chapter 5), which revealed higher scores in the CLIL group.

The literature on learners’ receptive knowledge of phraseological units reveals that those who benefit from more exposure to TL input show a greater receptive collocation knowledge (e.g. Shehata, 2008; Yamashita & Jiang, 2010; Alsakran, 2011).

As hypothesized, the present study demonstrated that CLIL learners were significantly better at retrieving the correct verb collocate in verb-noun collocations. In fact, they systematically

scored better than NON-CLIL learners, except for two items (both in TL English and TL Dutch). This is in line with previous research and seems to confirm Gyllstad's (2007) statement that learners who benefit from more exposure in the TL demonstrate a higher ability to establish "a direct link between L2 lexical forms (single words or word combinations/collocations) and the conceptual store" (p. 249). In early stages of TL acquisition, words in the TL are often linked to L1 translations, which in turn are linked to conceptual representation. When TL proficiency increases, direct conceptual connections from TL words to semantics develops (see Kroll & Stewart's Revised Hierarchical Model, 1994). In other words, higher TL exposure helps learners make a direct link between, for instance, the form *in twijfel trekken* and its conceptual representation: to cast doubt on something. Interferences from the L1 are still present but are less strongly activated in processing. A qualitative look at the TL English collocations where CLIL and NON-CLIL learners differed the most revealed a higher ability in CLIL learners to make use of positive transfer and even more so when a less core meaning of the verb is involved (*draw a conclusion*) or when the verb is less frequent (*seize an opportunity*). Surprisingly, NON-CLIL learners scored better than CLIL learners for two 'difficult' items, but the differences were not radical. We can suppose that for these items, both groups of learners just guessed. As for TL Dutch, CLIL learners scored better at retrieving congruent collocations and even some of the non-congruent collocations. However, some items turned out to be as difficult for the NON-CLIL group as for the CLIL group, probably because learners did not encounter these enough to be able to recognize them as correct Dutch collocations. The high number of incorrect answers found for the congruent collocations *in scène zetten* is surprising since the verb is frequent and its meaning quite transparent.

Beyond that, this study also showed that the results strongly varied according to the collocation itself, and this for both target languages. In fact, some collocations were (very) easily retrieved by the learners while others seemed to cause more trouble. A first statistical analysis showed a positive correlation between frequency of the collocation in a reference corpus and its correct retrieval, but only for English as a TL. A more qualitative analysis of every selected item revealed that direct equivalence with the learners' mother tongue seemed to have a major impact. In fact, all TL English collocations and the majority of the TL Dutch collocations labelled as 'easy' (over 50% of correct answers) or 'very easy' (more than 75% of correct answers) were congruent collocations – they had a literal translation equivalent in the learners' L1. A qualitative look at the items labelled as 'difficult' (less than 50% of correct answers) helped to uncover different possible reasons for the learners' lower performance. As far as TL

English is concerned, a first explanation may be that some of the verb collocates were not used in their core meaning, but in a more restricted sense. This refers to what Cowie (1991, p. 102) calls ‘restricted collocations’, i.e. “word-combinations in which one element (usually the verb) has a technical sense, or a long-established figurative sense which has lost most of its analogical force”. These can refer to collocations where the verb is used in a delexicalized way (as in *do damage*, *do harm* and *pay a visit*), which means that the verb is semantically neutral and lacks distinct meaning (Fillmore, Johnson & Petruck, 2003). They can also refer to collocations where the verb is used in a figurative way (as in *draw a conclusion* with the verb ‘draw’ being mostly used in the literal sense of ‘sketching’ or ‘pulling’), or in a technical sense, which means that it usually occurs in a special register (for example *press charges*, usually found in legal texts or newspaper articles (Gyllstad, 2007)). All these specialized uses of the verb collocate reduce its level of transparency. The use of phrasal verbs in combination with nouns also seemed to raise the level of difficulty (as in *put out a fire*), corroborating the fact that their meaning remains hard to understand for learners (Siyanova & Schmitt, 2007; Schmitt & Redwood, 2011). Another possible explanation for TL English is the low frequency of the verb collocate (as in *seize an opportunity* and *shed tears*, for instance), which minimizes the chance of learners encountering it. Similarly, the TL Dutch items that turned out to be difficult involved a higher number of non-congruent collocations. Moreover, they involved a number of verb collocates that were used in a more restricted sense than their literal meaning, making them less transparent to the learners (e.g. *op iemands zenuwen werken*, *een kijkje nemen*, *in twijfel trekken*).

While congruence with the learners’ mother tongue appeared to be a facilitating factor in the case of ‘very easy’ and ‘easy’ collocations, it was not particularly helpful in the case of ‘difficult’ items. In fact, it seemed that when learners are not familiar with the meaning of the verb collocate, they cannot have access to the French equivalent or are reluctant to resort to it (cf. Kellerman, 1979). Moreover, resorting to French probably caused some incorrect answers in the case of the TL English *do damage* and *do harm* – the learners clearly hesitated between the verbs ‘do’ and ‘make’ because French only has one verb (*faire*) that encompasses the meanings of the two English verbs (Nehls, 1991). The learners of Dutch – and especially the NON-CLIL ones – also seemed to have been influenced by their L1 when incorrectly choosing the pseudo-collocations ‘openbaar leggen’, ‘de telefoon beantwoorden’, ‘zich zorgen doen’, ‘fouten doen’ and ‘onder woorden zeggen’. This had already been noticed in Gilquin’s study (2007) in which French-speaking learners of English had to evaluate the acceptability of correct and incorrect verb-noun collocations using the verb ‘make’. According to her, when learners

do not know a collocation, “they tend to rely on what they know best, namely the corresponding collocation in their mother tongues, hence the importance of (positive and negative) transfer” (p. 9).

It seems important to note that the analysis of the Dutch learners’ productive phraseological language carried out in Chapter 5 revealed a number of errors related to the verbs ‘doen’ and ‘maken’ in combination with nouns – learners often used ‘doen’ instead of ‘maken’ and the other way round (e.g. ‘een tocht *doen’, ‘een vuur *doen’, ‘dingen *maken’, ‘aan sport *maken’, etc.). The difficulty lies in the fact that French has only one verb (*faire*) that encompasses the meanings of the two Dutch verbs. The current study showed that this distinction between ‘doen’ and ‘maken’ tends to be less problematic at the receptive level for the learners of Dutch. In fact, the cases of *fouten maken* and *zich zorgen maken* seem to confirm this statement, since around 70% of correct answers were given for these items, even though ‘fouten doen’ and ‘zich zorgen doen’ were used as distractors.

Regarding the collocations that belonged to the C1-C2 levels of the CEFR (this only concerns TL English collocations), that is one or two levels above that expected by the learners, they turned out to be ‘easy’ or ‘difficult’ to retrieve. The ones labelled as ‘easy’ (*exercise rights*, *lose count* and *serve a purpose*) had direct French equivalents. This would suggest that congruence clearly facilitates the receptive acquisition of TL collocations, even the ones corresponding to higher levels of ability. The items labelled as ‘difficult’ (*shed tears* and *press charges*) did not have direct equivalents in the learners’ L1 and were apparently not familiar to the learners to the extent that they showed no clear preference for any of the three options. When we compared the number of answers given for these higher-level items in each learner group, we saw that CLIL learners gave a higher number of correct answers for all items but one (*press charges*).

One of this study’s limitations is the small number of items selected for the test. In fact, analyzing the learners’ receptive knowledge of 18 collocations for English and 22 collocations for Dutch is less than the 50 recommended by Nation (2001). This could lend less weight to our findings and may not be enough to draw well-founded conclusions. However, given the large number of tasks the learners had to perform in one day, we had to face time constraints and were concerned about concentration issues. Also, giving learners the opportunity to indicate whether they resorted to guessing or not could have supported some of our suppositions (for example in the strange case of *press charges*). However, Gyllstad (2007) decided not to include a control for guessing in the last version of his test because according to him, “guessing is not a major problem [...] some sort of knowledge is the driving factor behind learners’ responses” (p. 156).

Moreover, no first test was carried out in order to see whether the learners knew the individual nouns and verbs, which could have been useful since “learners’ claimed level of knowledge of a collocation may depend on their knowledge of the parts of the combination” (Gyllstad, 2007, p. 58). However, we made sure to select items that are supposed to be acquired by the end of Grade 12 (at the exception of the five discriminating items from C1 and C2 levels). Another possible limitation related to the Dutch test is the absence of information about its validity and the fact that its format differs from the English one. The Dutch verb collocates, because they were presented to the learners in context sentences, were sometimes found in their infinitive form, the present tense or the past participle. It may be that learners have not been exposed to the different forms of the verb in their input, making the retrieval more difficult. This relates to Gyllstad’s ‘structural aspect’, which actually refers to “a deviation from the patterns of language that a language user has experienced” (p. 40). However, we would expect the verbs to be used in their different forms in the (oral and written) input of the learners (except in vocabulary lists). On the other hand, the presence of the context in the Dutch test, however minimal, could also have been a facilitating factor in some cases, compared to the English test.

On a more positive note, this study hopefully contributes to the existing research on learners’ receptive collocational knowledge. As reported by Gyllstad (2007), the bulk of studies investigating learners’ knowledge of collocations analyzed their productive knowledge. The qualitative analysis of the selected items provided several interesting insights into the features that help or hinder the acquisition of English and Dutch collocations by French-speaking learners. The present study also showed supplementary evidence that CLIL – or at least greater TL exposure – appears to positively affect the receptive knowledge of collocations. A regression analysis including control variables such as the learners’ non-verbal intelligence and socio-economic status, as well as linguistic variables such as frequency of formal and informal TL input and general vocabulary knowledge is carried out in the next chapter in order to isolate the actual effect of CLIL on the learners’ receptive phraseological knowledge.

CHAPTER 7: Effects of target language exposure and vocabulary knowledge on the acquisition of verb-noun collocations

7.1. Introduction

This study follows on from Chapter 6 in which the learners' receptive knowledge of verb-noun collocations was analyzed. The results revealed that the CLIL group had a higher receptive knowledge of this type of phraseological units than the NON-CLIL group. It also showed that the target collocation itself (due to its inherent characteristics) had a moderate impact on the answer given to the test, both in English and Dutch as target languages. The frequency of occurrence of the collocation in a reference corpus had also an influence, but for English only.

The present chapter will assess the influence of factors such as formal and informal target language exposure as well as general vocabulary knowledge on the learners' receptive knowledge of verb-noun collocations. Literature shows that time of exposure is an important predictor of language proficiency outcomes (Lambelet & Berthele, 2015) and receptive vocabulary knowledge in particular (Golberg, Paradis & Crago, 2008; Terrazas Gallego & Agustín Llach, 2009). Informal contact with the TL also appears to have a positive impact on the acquisition of the TL (Kinsella, 2009; Moyer, 2009), and specifically on the acquisition of phraseological units (Fernández & Schmitt, 2015). Besides, a previous study carried out in the framework of this project and using the same sample (Van Mensel, Bulon, Hendrikx, Meunier & Van Goethem, submitted) has found that:

1. CLIL and informal TL contact positively correlated with the learners' written lexical diversity in Dutch;
2. formal TL instruction and CLIL significantly predicted mean word length and lexical diversity in English.

This encourages us to predict that these variables will also influence the acquisition of TL collocations. In addition, it has been argued that the knowledge of single words and receptive collocational knowledge are positively related (Gyllstad, 2007), even though Bahns and Eldaw (1993) suggested that the reason for the difficult acquisition of collocations is that it does not develop according to general vocabulary knowledge. Looking at the influence of (productive and receptive) vocabulary size on the receptive knowledge of phraseological units will disentangle the effect of CLIL from the effect of the learners' general vocabulary knowledge.

This chapter first introduces the various variables whose influence is analyzed in this study, namely: target language exposure variables (Section 7.2.), productive and receptive vocabulary knowledge variables (Section 7.3.), and two background variables (socio-economic status and non-verbal intelligence) (Section 7.4.). The research questions and hypotheses are presented in Section 7.5. Section 7.6. discusses the statistical analyses before proceeding to the results in Section 7.7., the discussion in Section 7.8. and the conclusion in Section 7.9.

7.2. Measures of target language exposure⁵⁰

Target language exposure is usually divided into two main categories: formal and informal. The former type of exposure refers to the amount of formal instruction the learners receive at school (often measured in number of years). The latter type of exposure refers to the frequency of informal contact with the TL (e.g. informal conversations with native speakers, reading, watching television or surfing the internet in the TL). The computation of these two different input measures is based on Muñoz's work (2011, 2014) and the data are derived from extensive questionnaires.

It is important to isolate the CLIL factor from these two types of input in order to assess its real impact on the learners' higher scores as found in the previous chapter.

7.2.1. Duration of target language instruction

As previously discussed in Chapter 2, Section 2.4.1., research on age showed that early starters reached higher levels of language proficiency than late starters in second language learning contexts (e.g. Patkowski, 1980; Johnson & Newport, 1989). However, this does not seem to apply to foreign language learning contexts (Muñoz, 2011), where late starters eventually catch up on younger learners (e.g. Cenoz, 2002, 2003; García Lecumberri & Gallardo, 2003; Miralpeix, 2006; Muñoz, 2006, 2008, 2011). Intensity of TL exposure appears to be a more important factor in these contexts (Torrás & Celaya, 2001).

This finding also holds for receptive vocabulary knowledge (Golberg et al., 2008; Terrazas Gallego & Agustín Llach, 2009). In terms of receptive phraseological knowledge, Gyllstad (2007) investigated the receptive collocational knowledge of Swedish learners of English with one year difference in terms of learning level (upper-secondary school and university level students) – which means that the latter group had been exposed to a larger amount of TL input – and he found higher scores in the university level group. However, the differences were not

⁵⁰ This section is partly based on Van Mensel et al. (submitted).

always significant when the groups were only one term – 4-6 months – apart.

The measure of formal TL instruction in the present study refers to the learners’ length of TL instruction calculated in years. The learners were asked to mention from which year onwards – in kindergarten, primary school or secondary school – they started to learn their TL and the number of years of TL learning was calculated on the basis of that information. In doing so, we obtained a mean number of years of TL instruction per learner group (Table 7.1.).

Table 7.1. Number of years of TL instruction per learner group

	English CLIL	English NON-CLIL	Dutch CLIL	Dutch NON-CLIL
<i>N</i>	82	69	128	72
Mean	6.28	5.38	9.22	7.82
SD	3.056	2.177	2.638	2.542

Table 7.1. shows that a) (CLIL and NON-CLIL) learners of Dutch as a TL benefited from more years of TL instruction than learners of English as a TL and that b) CLIL learners (of both English and Dutch) spent more years learning the TL than their peers in traditional foreign language instruction. This difference between the groups is significant for Dutch ($U=3198.5$, $z=-3.650$, $p=.000$, $r=-.26$).

7.2.2. Informal target language input outside school⁵¹

It seems from previous research (Kinsella, 2009; Moyer, 2009) that contact with native speakers clearly has a positive impact on language learning, and even more so in informal than in formal contexts. With regard to phraseological competence, various studies showed a positive influence of extracurricular activities on the acquisition of phraseological units. Schmitt & Redwood (2011) looked at factors affecting learners’ productive and receptive knowledge of TL English phrasal verbs and found a positive relationship between extensive reading and watching English language movies and TV programs and the acquisition of phrasal verbs. Lin (2014) confirmed the suggestion that television can facilitate the acquisition of TL formulaic sequences. Fernández & Schmitt (2015) reported that the productive collocational knowledge of Spanish learners of English correlated with everyday engagement with the TL outside school, including activities such as extensive reading, watching movies/TV, internet use or social networking.

⁵¹ Also referred to as “extramural” by some researchers (e.g. Sundqvist & Sylvén, 2016).

The measure of informal TL input used in this study refers to current informal contact with the TL outside school, calculated as a composite measure consisting of: frequency of internet use in the TL, frequency of TL (productive and receptive) use outside school and frequency of contact with native speakers outside school. The measure (Cronbach's alpha .78)⁵² is based on the following items from the questionnaires administered to the learners; they replied on a scale from 1 to 5 (never – rarely – sometimes – often – very often):

[Original in French]

1. Sur internet, à quelle fréquence utilises-tu l'anglais / le néerlandais?
2. A quelle fréquence parles-tu l'anglais / le néerlandais en dehors de l'école?
3. A quelle fréquence entends-tu l'anglais / le néerlandais en dehors de l'école?
4. As-tu des contacts avec des anglophones / des néerlandophones en dehors de l'école ou de la maison?

[English translation]

1. How often do you use English / Dutch on the Internet?
2. How often do you speak English / Dutch outside school?
3. How often do you hear English / Dutch outside school?
4. Do you have any contact with English-speakers / Dutch-speakers outside school or home?

Table 7.2. describes the results for each sub-group.

Table 7.2. Current informal target language input outside school

	English CLIL	English NON-CLIL	Dutch CLIL	Dutch NON-CLIL
<i>N</i>	85	71	120	68
Mean	3.36	2.91	2.52	2.15
SD	0.840	0.793	0.783	0.720

Table 7.2. shows that (CLIL and NON-CLIL) learners of Dutch are less in contact with the TL outside school than the learners of English, which could be surprising since Dutch is one of the national languages in Belgium. However, as mentioned in Chapter 2, Section 2.4.2., Dutch cannot be considered as a second language for the French-speaking learners because they do

⁵² Cronbach's alpha is a measure of internal consistency and it is also used as a measure of scale reliability. In other words, it indicates how closely answers to a set of items are related.

not encounter the language frequently in their everyday life. By contrast, English being an international language, it is certainly more present in the learners' daily life through music, TV series, etc. As for the comparison between the CLIL and the NON-CLIL groups, CLIL learners appear to be more engaged with the TL outside school than their NON-CLIL peers, and this in both target languages. This difference was significant for English ($U=2147.5$, $z=-3.109$, $p=.002$, $r=-.25$) and Dutch ($U=3076$, $z=-2.817$, $p=.005$, $r=-.20$).

7.3. Measures of productive and receptive vocabulary knowledge

Some studies (e.g. Bahns & Eldaw, 1993; Arnaud & Savignon, 1997) found that advanced learners had more difficulties acquiring collocations than single word items, hence arguing that TL general vocabulary knowledge and TL collocation knowledge did not develop simultaneously. However, more recent research suggested that vocabulary size is an indicator of collocational knowledge. Gitsaki (1999) studied the development of collocational knowledge and reported a significant increase in collocational knowledge as the learners' vocabulary knowledge developed. Koya (2005) found the same findings looking at Japanese learners of English at university level. Gyllstad (2007) investigated the influence of vocabulary size on the knowledge of collocations and found a close relationship between advanced learners' scores on the receptive vocabulary size test and scores on two receptive tests of collocational knowledge. Mutlu & Kaşlıoğlu (2016) analyzed whether the receptive vocabulary size and the productive and receptive verb-noun collocational knowledge of high school Turkish learners of English were related. The analysis revealed a positive correlation between the learners' average vocabulary size and their collocational knowledge. Bergström (2008) examined the vocabulary and receptive collocation knowledge of Swedish upper secondary school learners of English and found a clear correlation between the two. Kadlecová's study (2014) involving Slovak secondary school learners also suggested that bigger TL vocabulary size and better TL collocational knowledge were linked.

The measures of productive and receptive vocabulary knowledge used in the present study were developed as follows. The learners' written productions (see Chapter 4, Section 4.5.1.) were analyzed in terms of global written competence (pilot study described in Chapter 4 Section 4.2.). One of the measures used in the study as a proxy for productive TL vocabulary knowledge is MTLTD (Measure of Textual Lexical Diversity)⁵³. It is calculated as the "mean length of

⁵³ The MTLTD scores used in the present study were calculated on the basis of the written productions collected at Time 2, since the learners performed the receptive collocational test at Time 2.

sequential word strings in a text that maintain a given TTR value” (McCarthy & Jarvis 2010, p. 385) and is less sensitive to text length compared to other measures of lexical diversity (see Bulon et al., 2017 for a more detailed discussion of these measures). As for the measure of the learners’ receptive vocabulary knowledge, two tests were carried out according to the TL: the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-IV) for English (Dunn & Dunn, 2007) and the Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL) for Dutch (Dunn, Dunn & Schlichting, 2005). They actually work the same way; the learners wore headphones and listened to recorded TL single words. For each word, the learners had the choice between four pictures that appeared on the computer screen, including one that depicted the word they had just heard. They could listen to the word as many times as they needed.

Table 7.3. shows the descriptive statistics for the learners’ productive and receptive vocabulary knowledge according to the TL.

Table 7.3. Productive and receptive vocabulary knowledge

		English		Dutch	
		CLIL	NON-CLIL	CLIL	NON-CLIL
MTLD based on written TL productions	<i>n</i>	82	77	129	79
	Mean	78.54	65.68	83.40	64.26
	SD	17.880	18.418	19.348	19.000
Receptive vocabulary knowledge based on PPVT-IV/ PPVT-III-NL	<i>n</i>	84	71	127	75
	Mean	149.92	107.76	129.47	92.93
	SD	30.149	32.352	19.579	28.741

We can see from Table 7.3. that CLIL learners achieved better scores of productive vocabulary knowledge (MTLD) than NON-CLIL learners, and this was significant for English ($U=1953$, $z=-4.150$, $p=.000$, $r=-.33$) and for Dutch ($U=2273$, $z=-6.700$, $p=.000$, $r=-.46$). The same holds for receptive vocabulary knowledge (PPVT-IV/ PPVT-III-NL): we found higher scores in the CLIL group compared to the NON-CLIL group. Again, the difference was significant for English ($U=1032.5$, $z=-7.002$, $p=.000$, $r=-.56$) and for Dutch ($U=1199$, $z=-8.879$, $p=.000$, $r=-.62$).

7.4. Non-verbal intelligence and socio-economic status

The Raven test (Raven, Court & Raven, 1998) was used to measure the learners' non-verbal intelligence (IQ). It consisted in solving progressively more complex puzzles. As for the learners' economic status (SES), we used the highest degree of the mother and distinguished between three different categories: 1 = low (secondary degree or lower), 2 = medium (short-type higher education), and 3 = high (long-type higher education).

Table 7.4. below provides the results regarding the learners' non-verbal intelligence per subgroup.

Table 7.4. Non-verbal intelligence (Raven score)

		English		Dutch	
		CLIL	NON-CLIL	CLIL	NON-CLIL
Non-verbal IQ (Raven score)	<i>n</i>	84	69	126	76
	Mean	43.74	42.06	46.40	42.64
	SD	6.262	7.812	6.495	8.128

Higher scores of non-verbal intelligence were found in the CLIL group and the difference between the groups was significant for Dutch ($U=3515.5$, $z=-3.165$, $p=.002$, $r=-.22$).

The descriptive results for the learners' socio-economic status are shown in Table 7.5.

Table 7.5. The level of education of the learners' mother (used as a proxy for the learners' SES)

		English		Dutch	
		CLIL	NON-CLIL	CLIL	NON-CLIL
Socio-economic status (SES)	<i>n</i>	64	56	116	65
	Mean	2.19	1.93	2.49	2.05
	SD	0.732	0.735	0.666	0.799

A higher socio-economic status was observed among the CLIL learners. The difference was significant for Dutch ($U=2616.5$, $z=-3.735$, $p=.000$, $r=-.28$)⁵⁴.

Pearson correlations were carried out to see whether the two control variables (non-verbal intelligence and socio-economic status) correlated with our dependent variable (receptive verb-noun collocation knowledge) (see Table 7.6. below).

⁵⁴ The difference was almost significant for English ($p=.056$).

Table 7.6. Pearson correlation between control variables (non-verbal intelligence and SES) and dependent variable (receptive collocational knowledge)

	English		Dutch	
	CLIL	NON-CLIL	CLIL	NON-CLIL
Receptive collocational knowledge				
Non-verbal IQ	<i>n</i> =84	<i>n</i> =69	<i>n</i> =126	<i>n</i> =76
	0.005 (<i>p</i> >0.05)	-0.017 (<i>p</i> >0.05)	0.056 (<i>p</i> >0.05)	0.109 (<i>p</i> >0.05)
SES	<i>n</i> =64	<i>n</i> =56	<i>n</i> =116	<i>n</i> =65
	0.233 (<i>p</i> >0.05)	-0.136 (<i>p</i> >0.05)	0.340** (<i>p</i> =.000)	0.311* (<i>p</i> =.012)

We found no correlations between the control variables and our dependent variable for English. However, the socio-economic status of the CLIL and NON-CLIL learners of Dutch appears to correlate with their receptive collocational knowledge.

7.5. Research questions and hypotheses

The receptive knowledge of verb-noun collocations was investigated in the previous chapter and revealed a better collocational knowledge among the CLIL learners in both English and Dutch as target languages. In the present study, we intend to gauge the impact of other input-related variables (besides CLIL) and general vocabulary knowledge on receptive collocational skills.

The present study addresses the following research questions:

1. How does CLIL relate to receptive knowledge of verb-noun collocations (RQ1)?

Since CLIL learners benefit from more TL exposure than NON-CLIL learners and are believed to acquire the new language in a more meaningful context, they tend to display higher levels of proficiency in the TL, including in the receptive knowledge of verb-noun collocations. However, it is important to analyze the relative contribution of CLIL in comparison with other input-related variables in order to shed light on its real impact, hence the present study.

2. What is the relation between duration of target language instruction and receptive knowledge of verb-noun collocations (RQ2)?

Duration of formal target language instruction has been recognized as an important factor influencing the acquisition of a TL in general and of vocabulary more specifically. We expect the number of years the learners spent learning the TL at school to have an impact on their receptive collocational knowledge.

3. How is informal target language input outside school associated with receptive knowledge of verb-noun collocations (RQ3)?

Among all the different input-related variables, informal contact with the target language outside school also appeared to positively affect the acquisition of phraseological language.

4. Are receptive and productive vocabulary knowledge linked to receptive knowledge of verb-noun collocations (RQ4)?

Although studies looking at the relation between vocabulary size and the receptive knowledge of collocations are not unanimous, its potential impact on our learners' receptive knowledge of verb-noun collocations should be analyzed.

7.6. Methodology

This section describes the statistical tests that were carried out to answer the research questions outlined above. The learner's receptive collocational knowledge was calculated on the basis of his/her total number of correct answers given to the test. The number of correct answers was divided by the number of items (18 for English; 22 for Dutch), resulting in one-decimal scores between 0 and 1. Spearman's rho correlation tests are first performed in order to measure the strength of association between the predictor variables⁵⁵ presented in Sections 7.2., 7.3. and 7.4., and the outcome variable (the scores of receptive knowledge of verb-noun collocations). The variables that are significantly related to the learners' receptive collocational knowledge are then included in linear regression models (hierarchical regression analysis – forced entry) to analyze to what extent they contribute to the learners' receptive collocational skills. The analyses are done for English and Dutch separately.

⁵⁵ It is important to note that, even though we use the term 'predictor', the present analysis is cross-sectional; all the variables were collected (almost) at the same time.

7.7. Results

The relation between the predictors and the outcome variable is analyzed in Section 7.7.1. (correlations), while the effect of the predictor on the outcome variable is scrutinized in Section 7.7.2. (linear regression models).

7.7.1. Correlations between the predictors and the outcome variable

We ran Spearman correlations to establish a potential relationship between the predictor variables (CLIL, measures of formal and informal target language exposure, receptive and productive measures of vocabulary knowledge) and the outcome variable (receptive knowledge of verb-noun collocations).

English

Table 7.7. presents the results of the Spearman correlations for English as a TL.

Table 7.7. Correlations predictor and outcome variables for English⁵⁶

	Receptive knowledge of verb-noun collocations
CLIL	$n=165$ $0.410^{**} (p=.000)$
Duration of target language instruction	$n=151$ $0.187^* (p=.021)$
Informal target language input outside school	$n=156$ $0.304^{**} (p=.000)$
Receptive vocabulary knowledge (PPVT-IV)	$n=155$ $0.530^{**} (p=.000)$
Productive vocabulary knowledge (MTLD)	$n=159$ $0.396^{**} (p=.000)$

* $p<.05$ (two-tailed) ** $p<.01$ (two-tailed)

As can be seen from Table 7.7., CLIL and the two other input-related variables (duration of target language instruction and informal target language input outside school) significantly correlate with the learners' receptive knowledge of verb-noun collocations. The learners' receptive and productive vocabulary knowledge are also significantly related to the number of

⁵⁶ The sample sizes (n) in the table are not the same because some data were missing.

correct answers given to the receptive collocational test. The correlations presented here lead to us to think that besides the (expected) effect of CLIL, the two other input measures and general vocabulary size affect the receptive knowledge of TL English verb-noun collocations.

Dutch

Table 7.8. presents the results of the Spearman correlations for Dutch as a TL.

Table 7.8. Correlations predictor and outcome variables for Dutch⁵⁷

	Receptive knowledge of verb-noun collocations
CLIL	<i>n</i> =214 0.602** (<i>p</i> =.000)
Duration of target language instruction	<i>n</i> =200 0.331** (<i>p</i> =.000)
Informal target language input outside school	<i>n</i> =188 0.254** (<i>p</i> =.000)
Receptive vocabulary knowledge (PPVT-IV)	<i>n</i> =201 0.717** (<i>p</i> =.000)
Productive vocabulary knowledge (MTLD)	<i>n</i> =208 0.580** (<i>p</i> =.000)

* *p*<.05 (two-tailed) ** *p*<.01 (two-tailed)

The same conclusions can be drawn for Dutch, as shown in Table 7.8. In fact, all input and vocabulary measures significantly correlate with the learners' receptive knowledge of TL Dutch verb-noun collocations. The strength of the correlations (as shown by the correlation coefficients) appear to be slightly higher than for English.

7.7.2. Regression analyses

We conducted multiple hierarchical regression analyses (forced entry) to analyze the relative contribution of the different predictors on the learners' receptive knowledge of verb-noun collocations. The same regression models were used for English and for Dutch so that comparisons between the target languages could be drawn. The learners' SES is included in the

⁵⁷ The sample sizes in the table (*n*) are not the same because some data were missing.

first block⁵⁸. The input-related variables (duration of target language instruction in years and informal target language input outside school) are included in the second block. The vocabulary-related variables (receptive and productive vocabulary knowledge) are included in the third block. CLIL was included in the fourth block in order to investigate its effect in isolation from all the other variables. All assumptions required for this type of statistical analysis were met⁵⁹.

English

Table 7.9. shows the results of the regression model for English as a TL.

Table 7.9. Regression model of the effect of different types of TL exposure and vocabulary knowledge on the percentage of correct answers to the multiple-choice test of English verb-noun collocations

English (n= 109)	<i>B</i>	<i>SE B</i>	β	VIF	<i>P</i>
Block 1					
(Constant)	0.454	0.047			0.000
SES*	0.047	0.022	0.205	1.000	0.033
Block 2					
(Constant)	0.252	0.073			0.001
SES	0.032	0.021	0.141	1.035	0.125
Duration of TL instruction**	0.017	0.006	0.260	1.036	0.005

⁵⁸ SES did not significantly correlate with the outcome variable when looking at English (see Section 7.4., Table 7.6.), but the p-value was almost significant ($p=.064$) for the CLIL group. In addition, the partial correlations of the regression model showed a significant correlation when taking the other variables into account.

⁵⁹ We checked the following assumptions before carrying out the multiple hierarchical regression analysis. First, the predictors should have some variation in value (i.e. they should not have variances of 0). This was indeed the case, hence the assumption of non-zero variance is met. Second, there should be no perfect linear relationship between two or more of the predictors. This was the case (VIF scores are $>1<10$). Hence, the assumption of no perfect multicollinearity was met. Third, the assumption of independent errors was respected. The residuals are only slightly positively correlated in the regression models for TL English (Durbin-Watson=1.601) and TL Dutch (Durbin-Watson=2.033). Fourth, at each level of the predictor variables, the variance of the residual terms should be constant. In other words, the residuals at each level of the predictors should have the same variance. We plotted the regression standardized residual against the regression standardized predicted value in a scatterplot. The points are randomly and evenly dispersed throughout the plot. This pattern is indicative of a situation in which the assumptions of linearity and homoscedasticity are met. Fifth, to test the normality of residuals, we looked at the histogram and normal probability plot. The histogram looked like a normal distribution (a bell-shaped curve). Hence, this assumption was also met. Sixth, the assumption of independence is also met, as all values of the outcome variable are independent – each value of the outcome variable comes from a different learner.

Informal TL input outside school*	0.042	0.019	0.202	1.040	0.028

Block 3

(Constant)	0.040	0.077			0.600

SES	0.022	0.019	0.095	1.066	0.245

Duration of TL instruction	0.010	0.005	0.150	1.096	0.071

Informal TL input outside school	0.024	0.017	0.114	1.079	0.164

PPVT-IV**	0.002	0.000	0.436	1.258	0.000

MTLD	0.001	0.001	0.093	1.230	0.288

Block 4

(Constant)	0.028	0.077			0.716

SES	0.018	0.019	0.077	1.090	0.345

Duration of TL instruction	0.010	0.005	0.153	1.097	0.064

Informal TL input outside school	0.022	0.017	0.105	1.086	0.201

PPVT-IV**	0.002	0.000	0.374	1.564	0.000

MTLD	0.001	0.001	0.085	1.235	0.331

CLIL	0.045	0.031	0.135	1.415	0.149

Note: $R^2=0.042$, $F(1,108)=4.684$, $\Delta R^2=0.042$, $\Delta F=4.684$ ($p<.033$)* $p=0.033$ for Block 1.
 $R^2=0.162$, $F(3,108)=6.762$, $\Delta R^2=0.120$, $\Delta F=7.516$ ($p<.001$)** $p=0.000$ for Block 2.
 $R^2=0.365$, $F(5,108)=11.863$, $\Delta R^2=0.203$, $\Delta F=16.515$ ($p<.000$)** $p=0.000$ for Block 3.
 $R^2=0.378$, $F(6, 108)=10.344$, $\Delta R^2=0.013$, $\Delta F=2.111$ ($p>.149$) $p=0.000$ for Block 4.

In the first block, we see that the learners' SES positively correlates with their receptive knowledge of verb-noun collocations when no other variable is considered.

When the two input-related variables are included (Block 2), SES stops contributing to the model, while the number of years of TL instruction and the current informal TL input outside school are positively correlated with the learners' receptive collocational knowledge in a significant way.

In the third block, the vocabulary measures are included in the regression model alongside the input-related variables and SES. Only the learners' receptive vocabulary (PPVT-IV) is significantly associated with their receptive knowledge of verb-noun collocations. The input-related variables do not contribute significantly to the model anymore, once the vocabulary measures are included in the regression model.

The results in the fourth block indicate that, isolated from the effects of the other predictor variables (SES, the other input-related variables and the measures of vocabulary knowledge), the CLIL variable does not appear to have a significant effect on the learners' receptive knowledge of verb-noun collocations anymore ($p=0.149$). The only remaining variable being significantly related to the learners' receptive knowledge of verb-noun collocations is the learners' receptive vocabulary knowledge (PPVT-IV) ($\beta=0.374$).

In other words, the variable of receptive vocabulary knowledge correlates more strongly with the learners' receptive knowledge of verb-noun collocations, compared to their SES, their productive vocabulary knowledge, the amount of formal and informal TL input they received, and CLIL. These are relatively less associated to the receptive knowledge of verb-noun collocations. The finding that CLIL does not significantly contribute to the model when other input-related variables and vocabulary measures are taken into account may result from the fact that there is a strong correlation between the learners' receptive vocabulary knowledge and CLIL ($\rho=0.550^{**}$, see Appendix 7). Although there is no violation of the assumption of perfect multicollinearity ($VIF >1 < 10$), it makes the model harder to interpret. We saw from Table 7.7. (simple correlations) that both variables significantly correlated with the learners' collocational knowledge and the present results would indicate that receptive vocabulary knowledge explains the variance that CLIL could explain if it were the only predictor included in the model. It is not really surprising to find that receptive vocabulary size is the variable that explains most of the variance since it logically includes receptive collocational knowledge. Knowing that CLIL, overall receptive vocabulary knowledge and receptive collocational knowledge highly correlate, we can only hypothesize that CLIL learners display a larger receptive vocabulary thanks to the specificities of the CLIL approach, and that this larger receptive vocabulary positively impacts their better performance on the receptive collocational test.

The final model explains 37.8% of the variance ($R^2=0.378$, adjusted $R^2=0.342$),⁶⁰ which can be considered as a relatively high predictive power. However, it also means that other factors may

⁶⁰ The adjusted R^2 is the modified R^2 based on the sample size and the number of predictor variables in the model.

contribute to the learners' receptive knowledge of verb-noun collocations as well, such as the presence of verb-noun collocations found in the teaching materials, or the teacher itself, for instance.

Dutch

Table 7.10. shows the results of the regression model for Dutch as a TL.

Table 7.10. Regression model of the effect of different types of TL exposure and vocabulary knowledge on the percentage of correct answers to the multiple-choice test of Dutch verb-noun collocations

Dutch (n= 158)	B	SE B	β	VIF	P
Block 1					
(Constant)	0.304	0.045			0.000
SES**	0.111	0.018	0.438	1.000	0.000
Block 2					
(Constant)	0.101	0.060			0.096
SES**	0.094	0.018	0.371	1.044	0.000
Duration of TL instruction**	0.018	0.005	0.261	1.084	0.000
Informal TL input outside school*	0.036	0.017	0.148	1.081	0.036
Block 3					
(Constant)	-0.164	0.055			0.003
SES*	0.037	0.015	0.147	1.230	0.014
Duration of TL instruction*	0.010	0.004	0.143	1.135	0.013
Informal TL input outside school	0.008	0.014	0.032	1.138	0.572
PPVT-III-NL**	0.003	0.000	0.422	1.437	0.000
MTLD**	0.003	0.001	0.291	1.349	0.000
Block 4					

(Constant)	-0.177	0.054			0.799
SES*	0.034	0.015	0.136	1.237	0.001
Duration of TL instruction*	0.009	0.004	0.131	1.144	0.021
Informal TL outside school	0.007	0.013	0.028	1.139	0.622
PPVT-III-NL**	0.002	0.000	0.346	1.761	0.000
MTLD**	0.002	0.001	0.261	1.397	0.000
CLIL*	0.067	0.026	0.170	1.608	0.011

Note: $R^2=0.192$, $F(1,157)=36.995$, $\Delta R^2=0.192$, $\Delta F=36.995$ ($p < .001$)** $p=0.000$ for Block 1.
 $R^2=0.296$, $F(3, 157)=21.631$, $\Delta R^2=0.105$, $\Delta F=11.467$ ($p < .001$)** $p=0.000$ for Block 2.
 $R^2=0.569$, $F(5, 157)=40.092$, $\Delta R^2=0.272$, $\Delta F=47.984$ ($p < .001$)** $p=0.000$ for Block 3.
 $R^2=0.587$, $F(6, 157)=35.741$, $\Delta R^2=0.018$, $\Delta F=6.600$ ($p < .001$)** $p=0.000$ for Block 4.

In the regression model for Dutch, we observe that the socio-economic status (SES) of the learners significantly contributes to the model when it is the sole variable considered (Block 1).

SES continues to have a significant impact when the two input-related variables are added to the model (Block 2), and both the duration of TL instruction in years and the current informal input the learners received outside school significantly predict the learners' receptive collocational knowledge.

The inclusion of the vocabulary measures in the model (Block 3) shows that these two variables are significant predictors of the learners' receptive knowledge of verb-noun collocations, alongside SES and number of years of TL instruction. The informal type of input does not contribute significantly anymore.

Adding CLIL to the regression model (Block 4) indicates that, isolated from the effects of the other predictor variables, CLIL still seems to have a significant influence on the learners' receptive knowledge of verb-noun collocations ($\beta=0.170$). Alongside CLIL, both the learners' receptive (PPVT-III-NL) and productive (MTLD) vocabulary knowledge significantly relate to the learners' receptive knowledge of collocations – more strongly in the case of their receptive vocabulary knowledge ($\beta=0.346$) than their productive vocabulary knowledge ($\beta=0.261$). Both the learners' SES and the number of years they spent learning the TL at school significantly impact their receptive collocational knowledge, but to a lesser extent ($\beta=0.136$ and $\beta=0.131$).

All variables but the frequency of informal contact with the TL outside school end up significantly influencing the learners' receptive knowledge of verb-noun collocations. The two vocabulary measures seem to be stronger predictors than SES, duration of TL instruction or CLIL, especially the receptive vocabulary measure. The final model explains 58.7% of the variance ($R^2=0.587$, adjusted $R^2=0.570$), which corresponds to a relatively high predictive power.

7.8. Discussion

A number of observations can be made from the analyses presented above. First, it is important to note that the significant correlations found between the predictor variables and the outcome measure were overall rather modest, indicating the importance of other factors that contribute to explaining variability in the learners' receptive knowledge of verb-noun collocations. Factors such as the frequency with which learners are exposed to verb-noun collocations in the TL input (at and outside school), the classroom practices (the fact that this type of collocations is focused on in teaching materials or not, the teachers' teaching style, the type of feedback, etc.), as well as individual variables such as motivation, may also play a role. Interestingly, the results from the regression analyses showed that the percentage of variance which is unaccounted for differs according to the TL. The present model appeared to be better at predicting receptive collocational knowledge for the learners of Dutch than for the learners of English, thereby suggesting that the variables unaccounted for play a larger role in explaining learners' variability for the latter⁶¹.

Another study carried out in the framework of the ARC project looking at the effect of different types of input on the learners' general writing skills in the target languages English and Dutch (Van Mensel et al., submitted) showed that CLIL was an important predictor (if not the strongest) of the learners' written productions in both target languages – CLIL significantly correlated with all four measures used as proxies for the learner's writing skills (text length, word length, sentence length and lexical diversity). The present study revealed that CLIL was one of the variables significantly impacting the learners' receptive collocational knowledge in Dutch, but not in English. This argues in favor of the CLIL approach as enhancing the receptive knowledge of verb-noun collocations, at least for the learners of Dutch as a TL. As hypothesized above, the fact that CLIL did not significantly correlate with receptive collocational knowledge

⁶¹ Plonsky and Ghanbar (2018) conducted a meta-analysis on multiple regression analysis in TL research and reported that the median of the R^2 values in the studies they assessed was 0.32. The R^2 values of the models in our study were higher (0.378 for English and 0.587 for Dutch).

in English may be due to the strong correlations found between CLIL, receptive vocabulary knowledge and receptive collocational knowledge.

Van Mensel et al.'s study (submitted) also showed that besides CLIL, the duration of formal TL instruction was significantly related to text length, word length and lexical diversity in English. As for Dutch, it was rather the frequency of informal TL input outside school (alongside CLIL) that positively influenced text length and lexical diversity. The present study looked at these two input-related variables, in addition to two measures of general vocabulary knowledge and the results showed that the effect of the two input measures tended to disappear in favor of the impact of the vocabulary measures. The receptive vocabulary knowledge of the learners of English was clearly the only remaining and strongest predictor of their receptive collocational knowledge. Similarly, the receptive and productive vocabulary knowledge of the learners of Dutch were the strongest predictors of their receptive collocational knowledge, the number of years of formal TL instruction being the weakest significant one. This corroborates the findings of other research showing the close and positive relationship of overall vocabulary size and receptive knowledge of verb-noun collocations (e.g. Gyllstad, 2007; Mutlu & Kaşlıoğlu, 2016).

The results were different according to the TL. We would have expected a somewhat similar picture for both TL learning conditions given the fact that significant correlations were observed between the predictor variables and the outcome variable both in Dutch and in English. Whereas only receptive vocabulary knowledge significantly predicted higher scores at the receptive collocational test for English, all variables but the current informal TL input outside school were significant predictors for Dutch. The effect of the learners' SES in Dutch can be related to findings of a background analysis carried out with the sample (Van Mensel, Hiligsmann, Mettwie & Galand, 2019) showing that the Dutch CLIL track was the most 'selective', with possible (self-)selection processes mostly associated with socio-economic factors. The most unexpected result would be the fact that CLIL only significantly impacted the receptive knowledge of verb-noun collocations in Dutch and not in English.

All in all, we can say that the better performance showed by the CLIL learners over their NON-CLIL peers at the receptive collocational test (cf. Chapter 6) is once more emphasized in the present study, although it only actually holds for the learners of Dutch in this specific case. Besides, this study has demonstrated that a large receptive vocabulary knowledge is predictive of the learners' receptive knowledge of verb-noun collocations for English and Dutch as target languages.

7.9. Conclusion

The goal of the present study was to explore the impact of two input-related variables – besides CLIL – as well as two vocabulary measures on learners' receptive knowledge of verb-noun collocations in two different target languages (English and Dutch). The study follows on from Chapter 6 which analyzed the receptive collocational knowledge of CLIL and NON-CLIL learners of English or Dutch and which revealed higher scores in the CLIL group. We expected CLIL to be positively related to the receptive knowledge of collocations, even if we disentangle its effect from the effect of other types of input such as duration of formal TL instruction at school and frequency of current informal contact with the TL outside school. We also expected the two vocabulary measures (receptive and productive TL vocabulary knowledge) to be positively associated with receptive collocational knowledge. All these predictor variables significantly correlated with the learners' receptive knowledge of verb-noun collocations (both in English and Dutch).

Regarding English, results from the regression model showed that none of the input-related variables were significant predictors of receptive collocational skills, not even CLIL. Only receptive vocabulary knowledge remained the sole and strongest predictor of the learners' receptive knowledge of verb-noun collocations.

As for Dutch, findings from the regression model pointed out that CLIL remained one of the significant predictors, even after including the two other input-related variables and the two vocabulary measures. SES, duration of TL instruction and the two vocabulary measures also appeared to significantly relate to the receptive acquisition of collocations. Only the informal type of extracurricular input did not significantly contribute to the model, which is not in line with the results reported in Van Mensel et al.'s study (submitted) looking at the effect of different types of input on the learners' overall writing skills.

Overall, we can conclude that the learners' receptive knowledge of verb-noun collocations is significantly associated with overall vocabulary knowledge. For learners of English, receptive vocabulary knowledge is in fact the only variable explaining the variance in the learners' receptive collocational knowledge. For learners of Dutch, CLIL is a significant predictor as well, alongside SES and the number of years the learners spent acquiring the TL at school, meaning that these variables explain part of the variance that remains unexplained by the vocabulary measures only. This indicates that proficiency in recognizing correct TL collocations is not entirely an effect of overall vocabulary knowledge; other factors should be

taken into consideration and this study helped to identify some of them, at least for Dutch. Moreover, the study confirmed once again that the investigation of learners with the same L1 acquiring different target languages may be beneficial for S/FLA research and research on CLIL since different findings were found according to the TL.

CONCLUSION

The present doctoral dissertation was part of a multidisciplinary project on CLIL in French-speaking Belgium and contributed to the project's linguistic research line by investigating the acquisition of phraseological units in English and Dutch by French-speaking learners in CLIL and NON-CLIL contexts.

Previous research on the linguistic outcomes of CLIL has largely focused on vocabulary at the single word level. The lack of focus on phraseological competence constituted a gap in the literature on CLIL as both L1 and S/FLA studies have demonstrated that authentic language is in essence phraseological: it is made of words that combine together and constitute one lexical choice rather than of separate lexical items that are put one next to another. Besides, phraseological competence is closely linked to linguistic competence in first and second/foreign language acquisition as it helps achieve fluency and facilitates the development of receptive and productive skills. Prior studies looking at the phraseological competence of learners acquiring a TL have shown that phraseology remains a stumbling block, even for advanced learners. The limited amount of input provided to learners in instructed settings does not allow them to become aware of the frequency with which some words co-occur. Interferences from the mother tongue also throw a spanner in the works of the learning process. The CLIL approach being considered as a more implicit and input-rich type of approach to learning, it seemed interesting to examine its effect on phraseology.

More specifically, **three main research hypotheses** guided the present doctoral research. These were **formed according to prior work done in the field of phraseology (Chapter 1), insights provided by the usage-based and emergentist theories of language acquisition (Chapter 2), existing research carried out on CLIL (Chapter 3), and results from a pilot study into the learners' global written competence (Chapter 4, Section 4.2.)**. The first hypothesis predicted that CLIL learners would display a higher phraseological competence in terms of frequency of use and accuracy when writing in the target languages English and Dutch. The second hypothesis was concerned with the learners' phraseological development over time. We expected both CLIL and NON-CLIL learners to increase their phraseological production and accuracy over the course of two consecutive school years, with CLIL learners showing more substantial gains than their NON-CLIL peers. The third hypothesis concerned the learners' receptive knowledge of verb-noun collocations and foretold that CLIL learners would score better at the receptive collocational test. These research hypotheses were tested out on the

basis of corpus-based and experimental data, as explained in Chapter 4. The results of the analyses were presented in Chapters 5 to 7.

Chapter 5 presented the findings of two related studies: 1) a cross-sectional analysis of the learners' phraseological productions, and 2) a longitudinal analysis of their phraseological development over time. The cross-sectional study showed that CLIL learners significantly produced a higher number (and a wider range) of phraseological units and that they significantly made fewer phraseological errors than NON-CLIL learners. The difference between the groups was larger in Dutch. The longitudinal study revealed slightly different patterns according to the TL. The CLIL and NON-CLIL learners of English significantly increased their phraseological production in a number of categories, but larger gains were found in the NON-CLIL group. As for Dutch, the NON-CLIL learners significantly increased their production of phraseological units in a large number of categories, while the CLIL learners significantly decreased in a small number of categories. No clear trend was found in terms of accuracy.

Chapter 6 reported the outcomes of a study investigating the learners' receptive knowledge of one specific type of phraseological units, namely verb-noun collocations. The results of the multiple-choice test demonstrated that the CLIL learners of English and Dutch have acquired a higher receptive knowledge of verb-noun collocations than the NON-CLIL learners. The study also showed that the collocation itself had an influence on the type of answer given to the test. Further analyses revealed that, for English, the frequency of the collocation (measured in a reference corpus) was significantly related to the learners' receptive collocational knowledge. A qualitative analysis showed that congruence also appeared to play a role. On the whole, English and Dutch congruent collocations, i.e. collocations that have a direct translation in the learners' L1, were more easily retrieved by the learners. However, collocations with a more restricted meaning or collocations composed of low frequency items seemed to cause more trouble. In relation to that, CLIL learners appeared to make a more correct use of positive transfer, even when the collocation had a less core meaning.

Chapter 7 presented the results of a follow-up study to the analysis of the learners' receptive collocational knowledge (Chapter 6), which examined the actual effect of CLIL on the receptive knowledge of verb-noun collocations when considering the influence of general vocabulary knowledge (receptive and productive), two other input-related variables (duration of TL instruction and current informal contact with the TL outside school), and two control variables (learners' non-verbal intelligence and socio-economic status). The results of the regression analyses showed that CLIL remained a significant predictor of the learners' receptive

knowledge of verb-noun collocations in Dutch after considering the impact of the other variables. The learners' SES, the number of years they spent learning the TL, and their receptive and productive vocabulary knowledge were also significant predictors. As for English, the only remaining predictive variable of the learners' receptive collocational knowledge was their global receptive vocabulary knowledge.

The results derived from these studies largely corroborate our first and third hypotheses. In fact, the **CLIL learners of English and Dutch showed a higher phraseological competence at the productive and receptive level.** This was shown by a greater use of phraseological units in their written productions and by a higher level of accuracy of their phraseological language. Regarding their receptive knowledge, the CLIL learners significantly scored better at the multiple-choice test, although we have seen that the educational setting was not the only significant predictor of the learners' knowledge of verb-noun collocations. **However, the results do not seem to support our second hypothesis since the phraseological language of the CLIL learners did not develop substantially after two consecutive years of TL learning.** Both CLIL and NON-CLIL learners of English actually increased their phraseological production, but the latter more than the former. The NON-CLIL learners of Dutch increased their production of phraseological units, while a decrease was found in the CLIL group. There was no clear improvement in the accuracy of their phraseological language.

A first observation that can be made on the basis of the different studies carried out in the present doctoral dissertation is that **the difference between the CLIL and the NON-CLIL learners is recursively greater in Dutch than in English.** In fact, our preliminary study on the learners' general written proficiency (Chapter 4, Section 4.2.) had already pointed out a greater disparity between the learners of Dutch than between the learners of English, and this appeared to be confirmed at the phraseological level. Moreover, Hendrikx (2019) observed the same tendency while carrying out her doctoral research on the acquisition of intensifying constructions by the same learners. One possible explanation could lie in the socio-economic background of the learners. Van Mensel et al.'s study (2019) into the socio-economic profile of the participants revealed that the differences between the groups were even more substantial in Dutch, with the Dutch CLIL track being the most 'selective' (as already mentioned in Chapter 7). This may thus explain the larger differences found at the linguistic level between the two Dutch learner groups. Another possible reason might be related to the learners' motivation and attitudes towards the TL – Chapter 2, Section 2.4.2. discussed the paramount role of such

learner-related variable in S/FLA. Besides, De Smet's (forthcoming) doctoral research on socio-affective factors carried out in the framework of the project revealed different findings according to the TL. In fact, the learners of English reported significantly more enjoyment and less anxiety (De Smet, Mettwie, Galand, Hiligsmann & Van Mensel, 2018) as well as more positive attitudes and a higher motivation (De Smet, Mettwie, Galand, Hiligsmann & Van Mensel, submitted) than the learners of Dutch. In sum, the fact that the Dutch CLIL track seems to appeal to learners coming from more privileged families, combined with the fact that English is associated with positive emotions and a higher motivation (no matter the educational approach), probably partly explains the greater differences found between the learners of Dutch at the linguistic level.

Second, although a series of positive linguistic outcomes were found in favor of the CLIL groups throughout our different studies, the longitudinal results did not live up to our expectations. In fact, **the CLIL setting did not seem to speed up the progress in phraseological competence of the learners, as we would have expected.** The additional hours of TL input the CLIL learners received during two consecutive school years did not seem to be sufficient to widen the gap between them and their NON-CLIL peers. Second language settings – where learners are immersed in a L2 environment for an extended period of time – probably remain the best language learning contexts in order to develop a native-like command of phraseological language since they provide learners with maximum opportunities for repeated exposure to multi-word units (Groom, 2009; Reppen, 2009). Nevertheless, introducing variability in the task (e.g. changing the text type or genre) could have brought about larger differences or have uncovered different types of development. As suggested in Gentil and Meunier's (2018) longitudinal study on nominalizations in L2 academic writing, whilst controlling for genre allows to compare the evolution of use of a specific linguistic feature, collecting writing sampled from a different genre can provide interesting insights related to varying genre demands.

While the cross-sectional study revealed a clear **advantage in the CLIL groups** in terms of frequency of production of phraseological units, this was **less obvious regarding the accuracy of their phraseological language.** Despite the fact that CLIL learners made fewer phraseological errors on the whole, a look at the different types of phraseological errors produced by the learners showed that some mistakes were equally found in the CLIL and NON-CLIL learners' productions (e.g. wrong selection of (non-)existing words in English, creation of deviant verbs, or prepositions missing in Dutch) (see Chapter 5). Also, no substantial

development in the quality of the learners' phraseological productions was noticed. This is in line with prior research on the (development of) phraseological competence of TL learners in general; accuracy very often tends to lag behind and to be an almost inevitable negative consequence of an increased phraseological productivity. In our case, this is possibly related to the informal nature of the tasks; the learners were asked to write as spontaneously as possible to a friend about a casual topic. Besides, they were repeatedly reminded of the non-evaluative character of the exercises. It may also be due to the CLIL approach itself, that is characterized by a more natural and communicative way of learning and that would therefore favor fluency over accuracy and an explicit treatment of errors. As mentioned in the theoretical part of this thesis (see Chapter 2, Section 2.3.3.), research into the CLIL classroom discourse has actually pointed to a lack of form-focused feedback. Significantly more language errors are found in the productions of CLIL learners (Hampl, 2011) and errors appear to be much less treated in CLIL than in traditional TL classrooms (Lochtmann, 2007; Schuitemaker-King, 2012). When attention is paid to accuracy in CLIL lessons, this is mostly done implicitly (Dalton-Puffer, 2007; Hampl, 2011); very few instances of metalinguistic feedback were reported in CLIL compared to NON-CLIL settings (Schuitemaker-Ring, 2012). An analysis of the data collected from the classroom observations made in some of the schools participating in the project (which goes beyond the scope of the present thesis) could probably support or reject this hypothesis.

Third, the follow-up study carried out in Chapter 7 demonstrated how **crucial** it is **to consider a number of other variables when researching CLIL**. The regression analyses actually showed that, at least for the learners' receptive knowledge of verb-noun collocations, it is their general (receptive) vocabulary knowledge that turned out to be the best and strongest predictor. That study also showed that part of the variability between the learners remained unexplained (especially for those learning English), which means that other variables still need to be taken into account. Actually, we cannot state with certainty that the better results found in CLIL in the present thesis are necessarily due to CLIL per se. It may be that the CLIL approach crystallizes a number of factors that are favorable to language acquisition and the present research design does not allow us to rule them out. One of the most obvious factors is the **greater exposure to TL input**. As we have seen in Chapter 2, *usage-based approaches* consider input as key to a target-like acquisition of the language. The extra hours of exposure provided in CLIL may have reinforced the *entrenchment* of multi-word units in the learners' mental lexicon. Another possible factor would be related to the fact that CLIL appears to attract learners with a **'better' profile**. Even though no official selection procedures are permitted in

French-speaking Belgium, Van Mensel et al.'s study (2019) into the socio-economic profile of the learners involved in the project (and therefore in the present thesis) seemed to demonstrate the existence of (self-)selection processes, with CLIL schools attracting learners from socially privileged families. Their study showed that the two learner populations (CLIL/NON-CLIL) significantly differed in terms of socio-economic background and school trajectory; the CLIL learners are coming from families with a higher socio-economic background and tend to be more successful at school (less repetitions of school years). The learners' **motivation** can arguably be taken into account too. What we know from the project's research line on socio-affective aspects is that the CLIL learners who took part in the study experienced less anxiety, reported more positive attitudes towards the TL and displayed a higher motivation to learn the language than their NON-CLIL peers (De Smet et al., 2018; submitted). Causal relationships between socio-affective factors and the acquisition of phraseological units have not been clearly established, but a study carried out by Dörnyei, Durow and Khawla (2004) recognized motivation as one of the factors influencing the learning and use of phraseological language. Other explanations may relate to educational aspects, such as **teaching practices**. It may be that the CLIL learners' attention has been more explicitly drawn to phraseology as part of explicit vocabulary teaching. In fact, observations made in some classrooms showed a CLIL teacher (native speaker of Dutch) explicitly drawing his learners' attention to the phraseological behavior of words in the feedback he gave to them after an oral exercise and expanding his learners' vocabulary by encouraging them to find synonyms. This teaching practice undoubtedly fosters language processing and acquisition, as we have seen that *explicit learning* or *focus-on-form* activities are as important for the acquisition of phraseological units as *implicit learning* or a more *focus-on-meaning* type of instruction for word combinations that are not very *frequent, salient, or prototypical*. In sum, while the amount of exposure to the TL clearly plays a role, findings observed in the other research lines of the project and related to socio-linguistic, motivational and educational aspects show how important it is to take a more holistic approach to CLIL, as a large number of factors seem to be at work.

A series of **limitations** have to be noted. A first limitation relates to the fact that only one methodology was used to analyze the learners' productive phraseological language (see Chapter 4, Section 4.6. for a reminder of the extraction and classification procedure). The approach adopted in this doctoral research was mostly based on frequency and the classification made according to the type of phraseological unit (cf. Granger and Paquot's 2008 taxonomy). It gave us interesting insights into the frequency, variety and accuracy of the learners' phraseological

language. However, literature shows that research carried out in the field of phraseology may yield different findings according to the criteria selected by the researcher for distinguishing between phraseological units and non-phraseological ones (Granger & Paquot, 2008). A study carried out by a Master's student (Dupal, 2018) made it possible to apply a different methodology (and therefore different selection criteria) to the same learner written data as the ones used in the present thesis. The *IdiomSearch* tool⁶² was used to extract and classify the phraseological units produced by the learners. Based on a 200-million token corpus of reference, the tool retrieves and classifies the learners' phraseological units according to two different criteria: frequency (frequent/not frequent) and degree of fixation (fixed/partly fixed/very fixed), the latter criterion giving an idea of the complexity of the phraseological units produced by the learners. The study allowed the extraction of about 30% more phraseological units, but the general findings remained the same for frequency of use and variety (the phraseological errors were not analyzed). A follow-up study (Meunier & Bulon, in preparation) on additional data also allowed us to show that using two different methodologies when analyzing the same learner data yielded similar results in this case, which reinforces the validity of the results found in the present thesis. A second limitation concerns our longitudinal study, which consisted of only two data collection points (Chapter 5). Actually, our research might not be considered as 'longitudinal' for some authors. Ployhart and Vandenberg (2010, p. 97) define longitudinal studies as: "emphasizing the study of change and containing at minimum three repeated observations on at least one of the substantive constructs of interest". Given the amount and variety of data that had to be collected for the different research lines of the project, it would have been quite challenging – both for the participating schools and the research team – to collect more longitudinal data from the learners. Moreover, a period of 18 months has passed between the two data collection points, which leaves room for potential changes to take place. In addition, Hendrikx' doctoral research (2019) focusing on the acquisition of intensifying constructions by the same learners also included a longitudinal study that revealed the same tendencies as those presented in the present manuscript, which also reinforces our findings. A last limitation has to do with the receptive tests (Chapter 6). The test used for English has been the object of seven empirical studies aiming at its design, its use and the evaluation of its effectiveness and has eventually been reported to yield valid and reliable scores of receptive knowledge of English collocations with upper-secondary school Swedish learners

⁶² Colson, J. P. (2017). The *IdiomSearch* Experiment: Extracting Phraseology from a Probabilistic Network of Constructions. In R. Mitkov (Ed.), *Computational and Corpus-Based Phraseology. EUROPHRAS 2017. Lecture Notes in Computer Science* (pp.16-28). Springer, Cham.

(Gyllstad, 2007). No tried and tested receptive test could be found for Dutch, though, and even if a pre-test was carried out with first year university students, the reliability of the results for Dutch may be reduced. Having said that, both tests seem to have been able to distinguish between CLIL and NON-CLIL learners' collocational knowledge, which is what we wanted to achieve.

Current (post-thesis) work is being carried out to investigate the learners' **spoken phraseological language** in the two language learning conditions (CLIL and NON-CLIL) and to make a comparison with their written phraseological language. Studies comparing CLIL and NON-CLIL learners' phraseological written competence are scarce, but those looking at their spoken phraseological productions and comparing these with their written productions are, to our knowledge, inexistent. According to Paquot and Granger (2012, pp. 142-143), "in spite of the flurry of learner-corpus-based studies on the L2 phrasicon, some important areas are still largely underexplored [...] there is a need for more investigations of the L2 spoken phrasicon which is interesting in its own right but will also allow for useful comparisons with the L2 written phrasicon". As more quantitative and qualitative analyses need to be carried out (see Bulon, forthcoming), the study has not been the object of an empirical chapter in the present thesis. It is however interesting to mention that initial results show that, when we compare the learners' written and spoken productions, we find significantly more phraseological units in the written corpus (with a significantly higher number of *referential phrasemes* and *textual phrasemes*). However, we find significantly more *communicative phrasemes* in the spoken corpus, which is consistent with this type of communication mode. When we compare the spoken productions of CLIL and NON-CLIL learners, we observe significant differences in favor of CLIL. The CLIL learners of English significantly produce more *lexical collocations* and *attitudinal formulae* than their NON-CLIL counterparts and the CLIL learners of Dutch significantly use more phraseological units in total, more *referential phrasemes* and *textual phrasemes*, and more *lexical collocations* and *grammatical collocations* than their NON-CLIL peers. Examples of *attitudinal formulae* that were exclusively used by CLIL learners of English are personal stance expressions such as *I guess (that)*, *I would prefer* or *I have to admit that*, discourse markers such as *you see what I mean* or *the thing is that*, hedges such as *kind of* or *sort of*, indirect forms such as *would you prefer* or *I thought you could*, or special conversational expressions such as *well then* or *okay well*.

Other **avenues for future research** are, among others, the potential influence of a number of other variables, as previously discussed. One of these is the teacher effect. A qualitative analysis

of the classroom observations made in some of the secondary schools is in order as we have noticed interesting practices (e.g. drawing the learners' attention to phraseology, as described above). In addition, an analysis of the data collected via the teacher questionnaires and related to the teachers' professional background (proficiency in the TL, teaching experience and training) should reveal itself interesting. In fact, we know that some CLIL teachers were native speakers of the TL, while others were not, despite their very good level in the TL. It would therefore be interesting to see whether this difference leads to differences in phraseological knowledge within the CLIL group of learners. Another important aspect related to the teachers is the methodological CLIL training. According to Paran (2013), one of the factors contributing to the success of CLIL is the teachers' education in CLIL and their understanding of the links between language and content. A recent study conducted by Alcaraz-Mármol (2018) with trained and non-trained CLIL teachers in Spain showed that methodological training made significant differences in terms of teachers' attitudes about CLIL efficiency for learning, the type of activities and materials used in class and the use of the L1. In French-speaking Belgium, no specific training is required to teach in a CLIL program but schools and teachers have collaborated with academics since 2005 to develop their own pedagogical toolkits and an 'immersion teaching certificate' can be obtained since 2014 in some universities (see Chapter 3). It is therefore possible that some teachers had received some training and others not. Closely related to this issue is the degree of commitment of the schools in the implementation of their CLIL programs. The frequency of contact and collaboration between the content teachers and the language teachers may vary from one school to another. Some schools also seem to provide for a person in charge of the coordination of the CLIL programs, but this is not always the case. The schools may also differ in the availability of material resources that are appropriate for the learning of target languages, such as language labs, computer rooms, data projectors, interactive boards, etc. A thorough analysis of the data collected from the teacher and school principals questionnaires should provide interesting insights into these matters. A last potentially relevant variable is the content of the teaching materials. In a study of the acquisition of the English passive in German CLIL and NON-CLIL learners, Möller (2017) analyzed the textbooks used in the two educational settings and found that the passive was more prevalent in CLIL materials. Another possible avenue to explore would involve a detailed analysis of transfer effects on the learners' phraseological language by means of a CIA (Granger, 1996; see Chapter 1, Section 1.5.1.2.). Traces of potential interferences from the learners' L1 (French) have been observed several times in this doctoral research and the data collected in the framework of the project lend themselves to this type of analysis. In fact, written data from native speakers of English

and Dutch have also been collected using the exact same task. Comparisons between L1 productions from learners (viz. French) and from native speakers of the target languages (viz. English and Dutch) can be made in order to identify cross-linguistic differences and possible difficulties for the learners. The TL productions from the learners (English and Dutch) can then be compared to the native speakers' L1 productions to see to what extent and how their interlanguage differs from the native speakers' L1. Finally, a comparison between the learners' L1 and TL productions can be drawn to find possible transfer effects from their L1. This would help establish transfer with a higher degree of confidence. Also, having investigated the learners' phraseological language based on production data, we only had access to 'learner-external' sequences, that is the linguistic patterns produced. The use of devices such as keystroke logging (Strömqvist, Holmqvist, Johansson, Karlsson & Wengelin, 2006), which records keyboard activities during the learners' writing process, would give us access to 'learner-internal' patterns. Previous studies' results seem to point to differences between what appears to be a phraseological unit in the learner's end product and how it is internally stored in his/her holistic lexicon (Meunier, forthcoming). Similarly, an analysis of the pauses in the learners' oral interactions would be valuable for psycholinguistic research. In fact, another accepted characteristic of phraseological units is that they are articulated in a fluent manner with a 'normal' intonation contour (e.g. van Lancker, Canter & Terbeek, 1981); in other words, they should be produced as a single intonation unit (Lin, 2010). Hesitations found between words within a phraseological unit (for example: *as a matter* (1 second pause) *of fact*) would be evidence that phraseological units are not holistically stored in the learners' mind. Another interesting avenue for research would be to replicate the study described in Chapter 5 using written data from different text genres in order to reveal different types of development in the learners' phraseological competence. Finally, it could be interesting to scrutinize the data for grammatical aspects, since a lack of focus on learner grammar development in LCR has been pointed out (Rankin, 2015).

Co-authored publications

Bulon, A. (forthcoming). The phraseological lexicon of CLIL and NON-CLIL learners in written and oral productions: A comparative analysis.

Meunier, F., & Bulon, A. (in preparation). Using two different taxonomies of MWUs to assess learners' productive EFL phrasicon in CLIL and traditional EFL settings: On the benefits of comparing apples and oranges.

Bulon, A., & Meunier, F. (submitted). Comparing CLIL and NON-CLIL pupils' learning curves in Dutch L2 phraseological competence: the winner does not take it all. *International Journal of Bilingual Education and Bilingualism*.

Hendrikx, I., Bulon, A., Meunier, F., Van Goethem, K., Van Mensel, L., & Naets, H. (submitted). Presenting MulTINCo: Multilingual Traditional Immersion and Native Corpus. *International Journal of Bilingual Education and Bilingualism*.

Van Mensel, L., Bulon, A., Hendrikx, I., Meunier, F., & Van Goethem, K. (submitted). Effects of input on L2 writing skills in English and Dutch CLIL and non-CLIL learners in French-speaking Belgium. *Journal of Immersion and Content-Based Language Education*.

Bulon, A. (accepted). Comparing the 'phrasicon' of teenagers in immersive and non-immersive settings: does input quantity impact range and accuracy? *Journal of Immersion and Content-Based Language Education*.

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.

https://sites.uclouvain.be/bkl-cbl/wp-content/uploads/2017/04/Bulon_et_al_2017.pdf

Hilgsmann, P., Van Mensel, L., Galand, B., Mettwie, 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*, 17(109), 1-25.

https://cdn.uclouvain.be/groups/cms-editors-girsef/demey/Cahier_109_final.pdf

References

- Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation*, 12, 75-93. DOI: 10.1080/13803610500392160
- Agustín Llach, M. P. (2009). The role of Spanish L1 in the vocabulary use of content and non-content EFL learners. In Y. R. de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and Language Integrated Learning: Evidence from Research in Europe* (pp. 112-129). Bristol: Multilingual Matters.
- Agustín Llach, M. P. (2010). An overview of variables affecting lexical transfer in writing: A review study. *International Journal of Linguistics*, 2(1), 1-17. DOI: 10.5296/ijl.v2i1.445
- Allerton, D. J. (1984). Three (or four) levels of cooccurrence restriction. *Lingua*, 63(1), 17-40. DOI: 10.1016/0024-3841(84)90030-5
- Alsakran, R. A. (2011). *The productive and receptive knowledge of collocations by advanced Arabic-speaking ESL/EFL learners*. (Unpublished doctoral dissertation). Colorado State University, Colorado.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. Cowie (Ed.), *Phraseology: Theory, Analysis and Applications* (pp. 101-122). Oxford: Oxford University Press.
- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195. DOI: 10.1093/applin/22.2.173
- Álvarez, E. (2006). Rate and route in EFL narrative development at different ages. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 127-155). Clevedon: Multilingual Matters.
- Amosova, N. N. (1963). *Fundamentals of English phraseology*. Leningrad: Publishing house of the Leningrad University.
- Anthony, L. (2014). AntConc (Version 3.4.3) [Computer Software]. Tokyo: Waseda University.
- Arnaud, P. J. L., & Savignon, S. J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 157-173). Cambridge: Cambridge University Press.
- Aronin, L., & Singleton, D. (2012). *Multilingualism*. Amsterdam/Philadelphia: John Benjamins.
- Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, 19(4), 537-552. DOI: 10.1017/S0142716400010353
- Babault, S., & Markey, M. (2016). Articulation langue 1 - langue 2 dans le répertoire langagier des élèves inscrits en programme immersif: quelles ressources lexicales pour les cours de sciences ? *The Canadian Modern Language Review/La revue canadienne des langues vivantes*, 72(4), 454-479. DOI: 10.3138/cmlr.3364

- Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.
- Baetens Beardsmore, H. (2009). Language promotion by European supra-national institutions. In O. García (Ed.), *Bilingual Education in the 21st Century. A Global Perspective* (pp. 197-217). Chichester: Wiley-Blackwell.
- Baetens Beardsmore, H. (Ed.) (1993). *European models of bilingual education*. Clevedon: Multilingual Matters.
- Bahns, J. (1997). *Kollokationen und Wortschatzarbeit im Englischunterricht*. Tübingen: Narr.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114. DOI: 10.1016/0346-251X(93)90010-E
- Baker, C. (2011). *Foundations of bilingual education and bilingualism*. Buffalo, N.Y.: Multilingual matters.
- Barfield, A. (2003). *Collocation recognition and production: Research insights*. Tokyo: Chuo University.
- Barfield, A. (2007). *An exploration of second language collocation knowledge and development*. (Unpublished doctoral dissertation). University of Swansea, Wales.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 629-654. Retrieved from http://barsaloulab.org/Online_Articles/1985-Barsalou-JEPLMC-determinants_graded_structure.pdf
- Bartning, I. & Forsberg, F. (2006). Stades de développement et séquences préfabriquées. In M. Olsen & E. H. Swiatek (Eds.), *Actes du XVI^e congrès des romanistes scandinaves* (pp. 1-22). Department of Language and Culture, Roskilde University.
- Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, 142(3), 402-409. DOI: 10.1016/j.actpsy.2013.01.010
- Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics*, 47(2), 383-411. DOI: 10.1515/LING.2009.014
- Berendse, E. P. H. (2014). *Acquiring L2 English prepositions in an L1 Dutch environment: The effect of immersion through CLIL teaching*. (Unpublished master's thesis). Utrecht University, Utrecht.
- Bergström, K. (2008). *Vocabulary and receptive knowledge of English collocations among Swedish upper secondary school students*. (Unpublished bachelor's thesis). Stockholm university, Stockholm.
- Berns, M., Claes, M. T., de Bot, K., Evers, R., Hasebrink, U., Huibregtse, I., Truchot, C., & van der Wijst, P. (2007). English in Europe. In M. Berns, K. de Bot & U. Hasebrink (Eds.), *In the presence of English: Media and European Youth* (pp. 15-42). New-York: Springer.
- Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, 26, 28-41. DOI: 10.1016/j.jslw.2014.09.004

- Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, 122(1), 67-73. DOI: 10.1016/j.cognition.2011.08.003
- Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, 24(1), 40-49. DOI: 10.1016/j.neuroimage.2004.09.044
- Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at...: Lexical bundles in university teaching and textbooks*. *Applied Linguistics*, 25(3), 371-405. DOI: 10.1093/applin/25.3.371
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of written and spoken English*. Harlow: Longman.
- Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24, 199-208. DOI: 10.1017/S0272263102002048
- Biskup, D. (1992). L1 influence on learners' renderings of English collocations. A Polish/German empirical study. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 85-93). London: Macmillan.
- Blondin, C. (2003). L'immersion linguistique dans l'enseignement fondamental en communauté française de Belgique: l'état de la question. *Journal de l'immersion*, 25(2), 19-31. Retrieved from <https://orbi.uliege.be/bitstream/2268/11055/1/BLONDIN-2003-25-JDI.pdf>
- Blondin, C., Candelier, M., Edelenbos, P., Johnstone, R., Kubaneck, A., & Taeschner, T. (1998). *Foreign languages in primary and pre-school education: a review of recent research within the European Union*. London: Centre for Information on Language Teaching.
- Blondin, C., & Straeten, M.-H. (2002). A propos d'immersion linguistique dans l'enseignement fondamental: un état de la question. *Le point sur la Recherche en Education*, 25, 3-21. Retrieved from <https://orbi.uliege.be/bitstream/2268/8869/1/BLONDIN-STRAETEN-2002-PRE.pdf>
- Boers, F., Deconinck, J., & Lindstromberg, S. (2010). Choosing motivated chunks for teaching. In S. De Knop, F. Boers & T. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 239-256). Berlin: Mouton de Gruyter.
- Bolinger, D. (1977). Idioms have relations. *Forum Linguisticum*, 2, 157-169.
- Bollen, K., & Baten, K. (2010). Bilingual education in Flanders: Policy and press debate (1999–2006). *The Modern Language Journal*, 94(3), 412-433. DOI:10.1111/j.1540-4781.2010.01089.x.
- Bonk, W. J. (2000). *Testing ESL learners' knowledge of collocations* (Report No. FL 801 384). (ERIC Document Reproduction Service No. ED442309). Retrieved from <https://files.eric.ed.gov/fulltext/ED442309.pdf>
- Bouillon, H., & Descamps, M. (2011). L'immersion linguistique par la pratique. Le cas de l'Ecole primaire de Court-Saint- Etienne. In L. Beheydt & P. Hiligsmann (Eds.), *Met immersie*

aan de slag / Au travail, en immersion (pp. 9-31). Louvain-la-Neuve: Presses universitaires de Louvain.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development*, 41(1), 1-104. DOI: 10.2307/1165959

Braun, A., De Man-De Vriendt, M.-J., & De Vriendt, S. (2006). Sept ans après ... *Éducation-Formation*, 275, 4-16.

Braun, A., & Vergallo, E. (2010). Influences de l'immersion linguistique sur la maîtrise du français. *Education et Formation*, e292, 155-165. Retrieved from https://s3.amazonaws.com/academia.edu.documents/25026835/manuels_scolaires_et_materiel_didactique.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1545324757&Signature=AU0uzVM27s%2B4rD2uB9MjeGixqVY%3D&response-content-disposition=inline%3B%20filename%3DStereotypes_et_manuels_scolaires.pdf#page=155

Brown, H., & Bradford, A. (2017). EMI, CLIL, & CBI: Differing approaches and goals. In P. Clements, A. Krause & H. Brown (Eds.), *Transformation in language education* (pp. 328-334). Tokyo: JALT.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39, 523-532. DOI: 10.1016/j.system.2011.08.002

Bueraheng, N., & Laohawiriyanon, C. (2014). Does learners' degree of exposure to English language influence their collocational knowledge? *International Journal of English and Literature*, 4(3), 1-10. Retrieved from <http://fs.libarts.psu.ac.th/research/sar/%E0%B8%9C%E0%B8%A5%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B8%95%E0%B8%B5%E0%B8%9E%E0%B8%B4%E0%B8%A1%E0%B8%9E%E0%B9%8C%E0%B9%83%E0%B8%99%E0%B8%A7%E0%B8%B2%E0%B8%A3%E0%B8%AA%E0%B8%B2%E0%B8%A3/%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2/6-DOES%20LEARNERS%E2%80%99%20DEGREE%20OF%20EXPOSURE%20TO%20ENGLISH%20LANGUAGE%20INFLUENCE.pdf>

Bulon, A. (accepted). Comparing the 'phrasicon' of teenagers in immersive and non-immersive settings: does input quantity impact range and accuracy? *Journal of Immersion and Content-Based Language Education*.

Bulon, A. (forthcoming). The phraseological lexicon of CLIL and NON-CLIL learners in written and oral productions: A comparative analysis.

Bulon, A., Hendriks, 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, *Travaux du Cercle Belge de Linguistique*, 11(1), 1-25. Retrieved from https://sites.uclouvain.be/bkl-cbl/wp-content/uploads/2017/04/Bulon_et_al_2017.pdf

Bulon, A., & Meunier, F. (submitted). Comparing CLIL and NON-CLIL pupils' learning curves in Dutch L2 phraseological competence: the winner does not take it all. *International Journal of Bilingual Education and Bilingualism*.

- Burger, H. (1998). *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin: Erich Schmidt.
- Burns, A. (2001). Analyzing spoken discourse: Implications for TESOL. In A. Burns & C. Coffin (Eds.), *Analyzing English in a global context: A reader* (pp. 123-148). London: Routledge.
- Buyl, A., & Housen, A. (2014). Factors, processes and outcomes of early immersion education in the Francophone Community in Belgium. *International Journal of Bilingual Education and Bilingualism*, 17(2), 178-196. DOI: 10.1080/13670050.2013.866627
- Bya, N., & Chohey-Paquet, M. (Eds.). (2005). *CLIL/EMILE: Digest préparé par le groupe d'accompagnement de l'immersion linguistique*. Brussels: FESeC.
- Bybee, J. (1998). The emergent lexicon. *Chicago Linguistic Society*, 34, 421-435. Retrieved from <https://www.unm.edu/~jbybee/downloads/Bybee1998EmergentLexicon.pdf>
- Bybee, J. (2003). Mechanisms in change in grammaticalization: The role of repetition. In R. Janda & B. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602-623). Oxford: Blackwell.
- Bybee, J. (2006). *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216-236). New York: Routledge.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Bybee, J., & Moder, C. (1983). Morphological classes as natural categories. *Language*, 59, 251-270. DOI: 10.2307/413574
- Callies, M. (2015). Learner corpus methodology. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 35-56). Cambridge: Cambridge University Press.
- Carroll, J. B. (1969). What does the Pennsylvania foreign language research project tell us? *Foreign Language Annals*, 3(2), 214-236. DOI: 10.1111/j.1944-9720.1969.tb01281.x
- Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign-language teaching. *Linguistics*, 11(112), 5-14. DOI: 10.1515/ling.1973.11.112.5
- Carter, R. (1998). *Vocabulary. Applied vocabulary perspectives*. London: Routledge.
- Castel, N., Hiligsmann, P., & Leloup, A. L. (2015). Verzoeken in het Nederlands van Franstalige CLIL- en niet-CLIL-leerlingen: een verkennend onderzoek. In S. Lestrade, P. de Swart & L. Hogeweg (Eds.), *Addenda. Artikelen voor Ad Foolen*, (pp. 49-68). Nijmegen: Radboud Universiteit.

- Celaya, M. L. (2008). 'I study *natus* in English': lexical transfer in CLIL and regular learners. In R. Monroy & A. Sánchez (Eds.), *25 Years of Applied Linguistics in Spain: Milestones and Challenges* (pp. 43-49). Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Celaya, M. L. (2012). "I wish I were three!": Learning EFL at an early age. In M. G. Davies & A. Taronna (Eds.), *New trends in early foreign language learning: The age factor, CLIL and languages in contact. Bridging research and good practices* (pp. 2-11). Newcastle: Cambridge Scholars Publishing.
- Celaya, M. L., & Torras, M. R. (2001). L1 influence and EFL vocabulary: do children rely more on L1 than adult learners? In M. F. Sierra, M. D. Dueñas & J. M. Pérez Fernández (Eds.), *Proceedings of the 25th International Conference of the Spanish Association for Anglo-American Studies (AEDEAN)* (pp. 1-14). Granada: University of Granada.
- Cenoz, J. (2002). Age differences in foreign language learning. *I.T.L. Review of Applied Linguistics*, 135(1), 125-142. DOI: 10.1075/itl.135-136.06cen
- Cenoz, J. (2003). Facteurs déterminant l'acquisition d'une L3: âge, développement cognitif et milieu. *Acquisition et Interaction en Langue Étrangère*, 18, 37-51. Retrieved from <http://journals.openedition.org/aile/1151>
- Chan, T., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231-251. DOI: 10.1080/09588220500185769
- Chen, M. (2013). Phrasal verbs in a longitudinal learner corpus: Quantitative findings. In S. Granger, G. Gilquin & F. Meunier (Eds.), *Twenty years of learner corpus research: Looking back, moving ahead* (pp. 89-101). Louvain-la-Neuve: Presses universitaires de Louvain.
- Chen, M., & Hin-tat, C. (2015). The use of recurrent multi-word expressions involving prepositions in Hong Kong L2 written English: A pseudo-longitudinal study. [Abstract]. In *Third International Learner Corpus Research Conference. Book of Abstracts* (pp. 38-40). Radboud University.
- Chen, P. C. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan*. (Unpublished master's thesis). National Taiwan Normal University, Taipei.
- Cheng, L. (2012). English immersion schools in china: Evidence from students and teachers. *Journal of Multilingual and Multicultural Development*, 33(4), 379-391. DOI: 10.1080/01434632.2012.661436
- Chomsky, N. (1980). Discussion of Putnam's comments. In M. Piatelli-Palmarini (Ed.), *Language and learning: The debate between Jean Piaget and Noam Chomsky* (pp. 310-324). MA: Harvard University Press.
- Chomsky, N. (1986). *Knowledge of language*. Berlin: Praeger.
- Chohey-Paquet, M. (2007). Belgium (French-Speaking). In D. Marsh & D. Wolff (Eds.), *Diverse contexts – Converging goals: CLIL in Europe* (pp. 25-32). Frankfurt: Peter Lang.
- Chohey-Paquet, M. (2008). CLIL in French-Speaking Belgium: Transforming paradox into potential. In C. M. Coonan (Ed.), *CLIL e l'Apprendimento delle Lingue. Le Sfide del Nuovo*

Ambiented di Apprendimento. Venezia: Libreria Editrice Cafoscarina (pp. 239-250). Venezia: Libreria Editrice Cafoscarina.

Chopey-Paquet, M. (2015). *Investigating teacher partnerships for CLIL: Developing a model for subject-content and language teacher pedagogic collaboration towards integration*. (Unpublished doctoral dissertation). Université de Namur & University of Aberdeen.

Cloud, N., Genesee, F., & Hamayan, E. (2000). *Dual language instruction: A handbook for enriched education*. Boston: Heinle and Heinle.

Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review*, 59(3), 393-424. DOI: 10.3138/cmlr.59.3.393

Cobb, T. (2007). The compleat lexical tutor (v.4.5 03/06). For data-driven language learning on the web [Electronic Version] from <http://132.208.224.131/>.

Colson, J. P. (2017). The *IdiomSearch* experiment: Extracting phraseology from a probabilistic network of constructions. In R. Mitkov (Ed.), *Computational and corpus-based phraseology. EUROPHRAS 2017. Lecture Notes in Computer Science* (pp. 16-28). Springer, Cham.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34(2), 302-312. DOI: 10.1037/0278-7393.34.2.302

Comblain, A., & Rondal, J. A. (2001). *Apprendre les langues. Où, quand, comment ?* Sprimont, Belgique: Mardaga.

Conklin, K., & Carrol, G. (2018). First language influence on a second language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 62-77). New-York: Routledge.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72-89. DOI: 10.1093/applin/amm022

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45-61. DOI: 10.1017/S0267190512000074

Cook, V. J. (2002). Language teaching methodology and the L2 user perspective. In V. J. Cook (Ed.), *Portraits of the L2 user* (pp. 325-346). Clevedon: Multilingual Matters.

Corder, S. P. (1973). *Introducing applied linguistics*. New-York: Penguin.

Cordier, C. (2013). *The presence, nature and role of formulaic sequences in English advanced learners of French: A longitudinal study* (Unpublished doctoral dissertation). Newcastle University, Newcastle.

Coseriu, E. (1954). Forma y sustancia en los sonidos del lenguaje. In E. Coseriu (Ed.), *Teoría del lenguaje y lingüística general* (pp. 115-234). Montevideo: Universidad de la República.

Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, 2(3), 223-235. DOI: 10.1093/applin/II.3.223

- Cowie, A. P. (1988). Stable and creative aspects of vocabulary use. In R. Carter & M. J. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 126-137). London: Longman.
- Cowie, A. P. (1991). Multiword Units in Newspaper Language. In S. Granger (Ed.), *Perspectives on the English lexicon: A tribute to Jacques van Roey* (pp. 101-116). Louvain-la-Neuve: Cahiers de l'Institut de Linguistique de Louvain.
- Cowie, A. P. (1994). Phraseology. In R. E. Asher (Ed.), *The Encyclopedia of language and linguistics* (pp. 3168-3171). Oxford: Oxford University Press.
- Cowie, A. P. (1998). Introduction. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 1-20). Oxford: Oxford University Press.
- Coxhead, A. (2008). Phraseology and English for academic purposes. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 149-161). Amsterdam/Philadelphia: John Benjamins.
- Coxhead, A. (in prep.). Exploring vocabulary use in an EAP writing task.
- Coyle, D. (2005). Developing CLIL: Towards a theory of practice. In N. Figueras (Ed.), *CLIL in Catalonia, from theory to practice* (pp. 5-29). Barcelona: APAC.
- Coyle, D. (2007). Content and Language Integrated Learning: Towards a connected research agenda for CLIL pedagogies. *The International Journal of Bilingual Education and Bilingualism*, 10(5), 543-562. DOI: 10.2167/beb459.0
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Language Teaching*, 49(1), 1-26. DOI: 10.1515/iral.2011.001
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles: California State University, National Evaluation, Dissemination and Assessment Center.
- Cummins, J., & Hornberger, N. (Eds.). (2008). *Encyclopedia of language and education. Volume 5: Bilingual education*. New York: Springer Science.
- Dabrowska, E. (2000). Form formula to schema: The acquisition of English questions. *Cognitive Linguistics*, 11(1/2), 83-102. DOI: 10.1515/cogl.2001.013
- Dabrowska, E., & Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 13(3), 437-474. DOI: 10.1515/cogl.2005.16.3.437
- Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. Amsterdam/Philadelphia: John Benjamins.

Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future perspectives for English language teaching* (pp. 139-157). Heidelberg: Carl Winter.

Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of applied linguistics*, 31, 182-204. DOI: 10.1017/S0267190511000092

Dalton-Puffer, C., Nikula, T., & Smit, U. (Eds.). (2010). *Language use and language learning in CLIL classrooms*. Amsterdam/Philadelphia: John Benjamins.

De Cock, S. (2003). *Recurrent sequences of words in native speaker and advanced learner spoken and written English: A corpus-driven approach*. (Unpublished doctoral dissertation). Université catholique de Louvain, Louvain-la-Neuve.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL)*, 2(1), 225-246. Retrieved from <http://dial.uclouvain.be/pr/boreal/object/boreal:75157>

de Graaff, R. (1997). THE *EXPERANTO* EXPERIMENT: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19(2), 249-276. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/A0844F17A48DF07333C2BDC63AB5B0C0/S0272263197002064a.pdf/experanto_experiment.pdf

de Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in Content and Language Integrated Learning (CLIL). *The International Journal of Bilingual Education and Bilingualism*, 10(5), 603-624. DOI: 10.2167/beb462.0

De Haan, P., & van der Haagen, M. (2013). The search for sophisticated language in advanced EFL writing: A longitudinal study. In S. Granger, G. Gilquin & F. Meunier (Eds.), *Twenty years of learner corpus research: Looking back, moving ahead. Corpora and language in use* (pp. 103-116). Louvain-la-Neuve: Presses universitaires de Louvain.

De Kleijn, P. (2003). *Combinatiewoordenboek van Nederlandse substantieven (zelfstandige naamwoorden) met hun vaste verba (werkwoorden)*. Amsterdam: Rozenberg Publishers.

De le Vingne, C. (2015). *Immersie-onderwijs met het Nederlands als doeltaal. Onderzoek naar de attitudes en motivatie van de ouders van immersie- en niet-immersie-leerlingen uit Luikse secundaire scholen*. Presented at the Anéla Applied Linguistics Conference. Egmond aan Zee, The Netherlands. May 22-23. Retrieved from <http://hdl.handle.net/2268/182924>

De Smet, A., Mettwie, L., Galand, B., Hiligsmann, P., & Van Mensel, L. (2018). Classroom anxiety and enjoyment in CLIL and non-CLIL: Does the target language matter? *Studies in Second Language Learning and Teaching*, 8(1), 47-71. DOI: 10.14746/ssllt.2018.8.1.3

De Smet, A., Mettwie, L., Galand, B., Hiligsmann, P., & Van Mensel, L. (submitted). Does CLIL shape language attitudes and motivation? Interactions with target languages and instruction levels. *International Journal of Bilingual Education and Bilingualism*.

de Villiers, J. G. (1980). The process of rule learning in child speech: A new look. In K. E. Nelson (Ed.), *Children's language* (pp. 1-44). New York: Gardner.

- de Zarobe, Y. R. (2008). CLIL and foreign language learning: A longitudinal study in the Basque country. *International CLIL Research Journal*, 1, 60-73.
- de Zarobe, Y. R. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 191-212). Amsterdam/Philadelphia: John Benjamins.
- de Zarobe, Y. R. (2011). Which language competencies benefit from CLIL? An insight into applied linguistics research. In Y. R. de Zarobe, J. M. Sierra & F. G. del Puerto (Eds.), *Content and Foreign Language Integrated Learning. Contributions to multilingualism in European contexts* (pp. 129-153). Bern: Peter Lang.
- de Zarobe, Y. R., Sierra, J. M., & del Puerto, F. G. (Eds.). (2011). *Content and Foreign Language Integrated Learning. Contributions to multilingualism in European contexts*. Bern: Peter Lang.
- Diessel, H. (2016). Frequency and lexical specificity in grammar: A critical review. In H. Behrens & S. Pfänder (Eds.), *Experience counts: Frequency effects in language* (pp. 209-238). Berlin/Boston: De Gruyter.
- Dörnyei, Z., Durow, V., & Khawla, Z. (2004). Individual differences and their effects on formulaic sequence acquisition. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 87-106). Amsterdam/Philadelphia: John Benjamins.
- Doughty, C., & Williams, J. (1998). Issues and terminology. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 1-11). Cambridge: Cambridge University Press.
- Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.
- Dunn, L. M., Dunn, L. M., & Schlichting, J. E. P. T. (2005). *Peabody picture vocabulary test-III-NL*. Harcourt Test Publishers.
- Dupal, J. (2018). *Investigating the phrasicon of CLIL and NON-CLIL students: A corpus-based comparative analysis using IdiomSearch*. (Unpublished master's dissertation). Université catholique de Louvain, Louvain-la-Neuve.
- Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26(2), 163-188. DOI: 10.1177/0267658309349431
- Ebeling, S. O., & Hasselgård, H. (2015). Learner corpora and phraseology. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 207-229). Cambridge: Cambridge University Press.
- Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, 79(1), 67-89. DOI: 10.1111/j.1540-4781.1995.tb05417.x
- Ellis, N. C. (Ed.). (1994). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press.

- Ellis, N. C. (1995). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, 4, 123-146. DOI: 10.1080/09658416.1995.9959876
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18(1), 91-126. DOI: 10.1017/S0272263100014698
- Ellis, N. C. (2002a). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188. DOI: 10.1017/S0272263102002024
- Ellis, N. C. (2002b). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, 24(2), 297-339. DOI: 10.1017/S0272263102002140
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 63-103). Oxford: Blackwell.
- Ellis, N. C. (2004). The processes of second language acquisition. In B. VanPatten, J. Williams, S. Rott & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 49-77). Mahwah, NJ: Erlbaum.
- Ellis, N. C. (2006a). Cognitive perspectives on SLA: The Associative-Cognitive CREED. *AILA Review*, 19, 100-121. DOI: 10.1075/aila.19.08ell
- Ellis, N. C. (2006b). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 1-31. DOI: 10.1093/applin/aml015
- Ellis, N. C. (2008a). Optimizing the input: Frequency and sampling in usage-based and form-focussed learning. In M. H. Long & C. Doughty (Eds.), *Handbook of second and foreign language teaching* (pp. 139-158). Oxford: Blackwell
- Ellis, N. C. (2008b). Phraseology: The periphery and the heart of the language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 1-13). Amsterdam/Philadelphia: John Benjamins.
- Ellis, N. C. (2008c). The dynamics of language use, language change, and first and second language acquisition. *The Modern Language Journal*, 41(3), 232-249. DOI: 10.1111/j.1540-4781.2008.00716.x
- Ellis, N. C. (2008d). Usage-based and form-focused SLA: The implicit and explicit learning of constructions. In A. Tyler, Y. Kim & M. Takada (Eds.), *Language in the context of use: Discourse and cognitive approaches to language* (pp. 93-121). Berlin/Boston: de Gruyter.
- Ellis, N. C. (2008e). Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention and the limited L2 end state. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372-406). London: Routledge.
- Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics*, 32, 17-44. DOI: 10.1017/S0267190512000025

- Ellis, N. C. (2013). Frequency-based accounts of second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 193-210). New York: Routledge.
- Ellis, N. C. (2016). Frequency in language learning and language change. In H. Behrens & S. Pfänder (Eds.), *Experience counts: Frequency effects in language* (pp. 239-256). Berlin/Boston: de Gruyter.
- Ellis, N. C., & Cadierno, T. (2009). Constructing a second language: Introduction to the special section. *Annual Review of Cognitive Linguistics*, 7, 111-139. DOI: 10.1075/arcl.7.05ell
- Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93, 370-386. DOI: 10.1111/j.1540-4781.2009.00896.x
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second-language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42, 375-396. DOI: 10.1002/j.1545-7249.2008.tb00137.x
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Elman, J. L. (2003). Generalization from sparse input. In M. Andronis, E. Debenport, A. Pycha & K. Yoshimura (Eds.), *Proceedings of the 38th Meeting of the Chicago Linguistics Society* (pp. 175-200). Chicago: Chicago Linguistics Society.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development*, 65, 1521-1540. DOI: 10.1111/j.1467-8624.1994.tb00833.x
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse*, 20(1), 29-62. DOI: 10.1515/text.1.2000.20.1.29
- Eskildsen, S. W. (2008). Constructing another language – Usage-based linguistics in second language acquisition. *Applied Linguistics*, 30(3), 335-357. DOI: 10.1093/applin/amn037
- Eskildsen, S. W., & Cadierno, T. (2007). Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics. In M. Nenonen & S. Niemi (Eds.), *Collocations and idioms 1: Papers from the first Nordic conference on syntactic freezes* (pp. 86-99). Joensuu: Joensuu University Press.
- Eurydice (2006). *Content and language integrated learning (CLIL) at school in Europe*. Eurydice European Unit. Retrieved from <http://bookshop.europa.eu/en/content-and-languageintegrated-learning-clil-at-school-ineurope-pbNCX106001/>
- Evert, S. (2005). *The statistics of word cooccurrences: word pairs and collocations*. (Unpublished doctoral dissertation). Universität Stuttgart, Stuttgart.

Evert, S. (2008). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp. 1212-1248). Berlin/Boston: de Gruyter.

Evert, S., & Krenn, B. (2001). Methods for the qualitative evaluation of lexical association measures. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics* (pp. 188-195). Toulouse, France: Association for Computational Linguistics. Retrieved from http://delivery.acm.org/10.1145/1080000/1073037/p188-evert.pdf?ip=130.104.201.142&id=1073037&acc=OPEN&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E6D218144511F3437&acm_=1559572417_a5cfddc25ddaf42e778d60cbc83c7a22

Fernández, B. G., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? *ITL-International Journal of Applied Linguistics*, 166(1), 94-126. DOI: 10.1075/itl.166.1.03fer

Ferris D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1-10. DOI: 10.1016/S1060-3743(99)80110-6

Fillmore, C. (1988). The mechanisms of construction grammar. *Berkeley Linguistics Society*, 14, 35-55. Retrieved from <http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/viewFile/1794/1566>

Fillmore, C., Johnson, C., & Petruck, M. (2003). Background to Framenet. *International Journal of Lexicography*, 16(3), 235-250. DOI: 10.1093/ijl/16.3.235

Fillmore, C., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language*, 64, 501-538. DOI: 10.2307/414531

Firth, J. R. (1957). *A synopsis of linguistic theory, 1930-1955*. Oxford: Basil Blackwell.

Forlot, G., & Beaucamp, J. (2008). Heurs et malheurs de la proximité linguistique dans l'enseignement de l'anglais au primaire. *Etudes de Linguistique Appliquée*, 149(1), 77-92. Retrieved from <https://www.cairn.info/revue-ela-2008-1.htm-page-77.htm>

Fox, H. J. (2002). Phrasal cohesion and statistical machine translation. In *Proceedings of the ACL-02 conference on empirical methods in natural language processing – Volume 10* (pp. 304-311). Philadelphia: Association for Computational Linguistics. Retrieved from http://delivery.acm.org/10.1145/1120000/1118732/p304-fox.pdf?ip=130.104.201.142&id=1118732&acc=OPEN&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E6D218144511F3437&acm_=1556893311_49bf3cf9ec3532c8b35277a76d2210ce

Francis, B., & Poole, R. (Eds.). (2009). *Oxford collocations dictionary*. Oxford: Oxford University Press.

François, T. (2011). *Les apports du traitement automatique du langage à la lisibilité du français langue étrangère*. (Unpublished doctoral dissertation). Université catholique de Louvain, Louvain-la-Neuve.

- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67, 155-179. DOI: 10.1111/lang.12225
- García Lecumberri, M. L., & Gallardo, F. (2003). English FL pronunciation in school students of different ages. In M. Pilar García Mayo & M. Luisa García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 115-135). Clevedon: Multilingual Matters.
- García Mayo, M. P. (2003). Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign language. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 94-114). Clevedon: Multilingual Matters.
- García, O. (2009). *Bilingual education in the 21st Century. A global perspective*. Chichester: Wiley-Blackwell.
- Gardner, R. C. (1979). Social psychological aspects of second language acquisition. In H. Giles & R. St. Clair (Eds.), *Language and social psychology* (pp. 193-220). Oxford: Blackwell.
- Gardner, R. C. (1985). *Social psychology and second language learning. The role of attitudes and motivation*. London: Edward Arnold.
- Gass, S. M., & Selinker, L. E. (1983). *Language transfer in language learning. Issues in second language research*. Rowley, MA: Newbury House Publishers.
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., Roberts, E. J., Hugues, E. K., & Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5, 1-14. DOI: 10.3389/fpsyg.2014.00011
- Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Writing development under CLIL provision. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 139-161). Berlin: Springer International Publishing.
- Genesee, F. (1988). Neuropsychology and second language acquisition. In L. M. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 138-150). Cambridge, Mass.: Newbury House.
- Genesee, F. (1991). Second language learning in school settings: Lessons from immersion. In A. G. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning: The McGill conference in honour of Wallace E. Lambert* (pp. 183-201). New York: Lawrence Erlbaum Associates.
- Gentil, G., & Meunier, F. (2018). A systemic functional linguistic approach to usage-based research and instruction: The case of nominalization in L2 academic writing. In A. E. Tyler, L. Ortega, M. Uno & H. I. Park (Eds.), *Usage-inspired L2 instruction: Researched pedagogy* (pp. 267-289). Amsterdam/Philadelphia: John Benjamins.
- Gernsbacher, M. A. (1994). *A handbook of psycholinguistics*. San Diego, CA: Academic Press.
- Giddens, A. (1984). *The constitution of society*. Cambridge: Polity Press.

- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363-374. DOI: 10.1093/elt/58.4.363
- Gilquin, G. (2007). To err is not all. What corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift für Anglistik und Amerikanistik*, 55(3), 273-291. DOI: 10.1515/zaa.2007.55.3.273
- Gilquin, G. (2009). *The (non-)use of phrasal verbs in L2 varieties of English*. Presented at the 30th International Computer Archive of Modern and Medieval English Conference (ICAME 30). Lancaster University. May 27-31. Retrieved from <http://hdl.handle.net/2078.1/112517>
- Gilquin, G. (2015). The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory*, 11(1), 51-88. DOI: 10.1515/cllt-2014-0005
- Gilquin, G., & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, 1(1), 41-61. DOI: 10.1075/etc.1.1.05gil
- Ginsburgh V., & Weber, S. (2007). La connaissance des langues en Belgique. *Reflets et perspectives de la vie économique*, 1, 31-43. DOI: 10.3917/rpve.461.0031
- Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of collocational knowledge*. Maryland: International Scholar Publications.
- Gläser, R. (1986). *Phraseologie der englischen Sprache*. Tübingen: Max Niemeyer.
- Gläser, R. (1998). The stylistic potential of phraseological units in the light of genre analysis. In A. P. Cowie (Ed.), *Phraseology. Theory, analysis, and applications* (pp. 125-143). Oxford: Oxford University Press.
- Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, 29(1), 41-65. DOI: 10.1017/S014271640808003X
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: The University of Chicago Press.
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, 7, 219-224. DOI: 10.1016/S1364-6613(03)00080-9
- Goldschneider, J. M., & DeKeyser, R. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51, 1-50. DOI: 10.1111/1467-9922.00147
- González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? : The effects of frequency and amount of exposure. *International Journal of Applied Linguistics*, 166, 94-126. DOI: 10.1075/itl.166.1.03fer
- Goodall, G. (2008). *Poverty/richness of the stimulus in Spanish instructed second language acquisition*. Presented at the Hispanic Linguistics Symposium. Université de Laval, Québec. October 23-26.
- Gouverneur, C. (2008). The phraseological patterns of high-frequency verbs in advanced English for general purposes: A corpus-driven approach to EFL textbooks analysis. In F.

Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 223-243). Amsterdam/Philadelphia: John Benjamins.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments & computers*, 36(2), 193-202. DOI: 10.3758/BF03195564

Granena, G. (2012). *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment*. (Unpublished doctoral dissertation). University of Maryland, College Park.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Languages in contrast* (pp. 37-51). Lund: Lund University Press.

Granger, S. (1998a). *Learner English on computer*. London: Longman.

Granger, S. (1998b). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 145-160). Oxford: Oxford University Press.

Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Amsterdam/Philadelphia: John Benjamins.

Granger, S. (2005). Pushing back the limits of phraseology: How far can we go. In C. Cosme, C. Gouverneur, F. Meunier & M. Paquot (Eds.), *Proceedings of Phraseology 2005. The many faces of Phraseology. An interdisciplinary conference* (pp. 165-168). Louvain-la-Neuve : Université catholique de Louvain.

Granger, S. (2008). Learner corpora. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp. 259-276). Berlin/New-York: Walter de Gruyter.

Granger, S., Gilquin, G., & Meunier, F. (2015). Introduction: learner corpus research – past, present and future. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 1-5). Cambridge: Cambridge University Press.

Granger, S., & Meunier, F. (2008). Phraseology in language learning and teaching: Where to from here? In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 247-252). Amsterdam/Philadelphia: John Benjamins.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: an interdisciplinary perspective* (pp. 27-49). Amsterdam/Philadelphia: John Benjamins.

Granger, S., & Wynne, M. (1999). Optimizing measures of lexical variation in EFL learner corpora. In J. Kirk (Ed.), *Corpora Galore* (pp. 249-257). Amsterdam: Rodopi.

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics*, 25(1), 38-61. DOI: 10.1093/applin/25.1.38

Gries, S. T. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 3-26). Amsterdam/Philadelphia: John Benjamins.

Gries, S. T. (2012). Frequencies, probabilities, and association measures in usage-/exemplar-based linguistics: Some necessary clarifications. *Studies in Language*, 36(3), 477-510. DOI: 10.1075/sl.36.3.02gri

Gries, S. T., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, 9(1), 97-129. DOI: 10.1075/ijcl.9.1.06gri

Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 21-33) London: Palgrave Macmillan.

Gyllstad, H. (2005). Words that go together well: Developing test formats for measuring learner knowledge of English collocations. *The Department of English in Lund: Working papers in linguistics*, 5, 1-31. Retrieved from <https://lucris.lub.lu.se/ws/files/5346153/625034.pdf>

Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners*. (Unpublished doctoral dissertation). Språk- och litteraturcentrum, Lunds universitet, Lund.

Gyllstad, H. (2009). Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 153-170). London: Palgrave Macmillan.

Hagglund, M. (2001). Do Swedish advanced learners use spoken language when they write English? A quantitative study of some common phrasal verbs in the written language of Swedish advanced learners, native students and professional writers. *Moderna Språk*, 95(1), 2-8.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning*, 24, 287-298. DOI: 10.1111/j.1467-1770.1974.tb00509.x

Hallin, A. E., & Van Lancker-Sidtis, D. (2017). A closer look at formulaic language: Prosodic characteristics of Swedish proverbs. *Applied Linguistics*, 38(1), 68-89. DOI: 10.1093/applin/amu078

Hampl, M. (2011). *Error and error correction in classroom conversation - A comparative study of CLIL and traditional EFL lessons in Austria*. (Unpublished doctoral dissertation). University of Vienna, Austria.

Handl, S. (2008). Essential collocations for learners of English: The role of collocational direction and weight. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 43-66). Amsterdam/Philadelphia: John Benjamins.

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(3), 379-400. Retrieved from <http://www.jstor.org/stable/44487973>

- Harrop, E. (2012). Content and language integrated learning (CLIL): Limitations and possibilities. *Encuentro*, 21, 57-70. Retrieved from <https://files.eric.ed.gov/fulltext/ED539731.pdf>
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), 237-258. DOI: 10.1111/j.1473-4192.1994.tb00065.x
- Heid, U. (2008). Computational phraseology: An overview. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 337-360). Amsterdam/Philadelphia: John Benjamins.
- Hendriks, I. (submitted). Receptive knowledge of intensifying compounds: Belgian French-speaking CLIL and non-CLIL learners of Dutch and English. *International Journal of Bilingual Education and Bilingualism*.
- Hendriks, I. (2019). The acquisition of intensifying constructions in Dutch and English by French-speaking CLIL and non-CLIL students: Cross-linguistic influence and exposure effects. (Unpublished doctoral dissertation). Université catholique de Louvain, Louvain-la-Neuve.
- Hendriks, I., Van Goethem, K., & Wulff, S. (2019). Intensifying constructions in French-speaking L2 learners of English and Dutch: Cross-linguistic influence and exposure effects. *International Journal of Learner Corpus Research*, 5(1), 63-103. DOI: 10.1075/ijlcr.18002.hen
- Hilgsmann, P., Van Mensel, L., Galand, B., Mettwie, L., Meunier, F., Szmalec, A., Van Goethem, K., Bulon, A., De Smet, A., Hendriks, 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*, 17(109), 1-25. Retrieved from [file:///C:/Users/bulon/Downloads/Hilgsmann%20et%20al%20\(2017\)%20Assessing%20CLIL%20in%20French-speaking%20Belgium%20\(1\).pdf](file:///C:/Users/bulon/Downloads/Hilgsmann%20et%20al%20(2017)%20Assessing%20CLIL%20in%20French-speaking%20Belgium%20(1).pdf)
- Hoffman, T., & Trousdale, G. (2013). *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.
- Hong, A. L., Rahim, H. A., Hua, T. K., & Salehuddin, K. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis. 3L; Language, Linguistics and Literature. *The Southeast Asian Journal of English Language Studies*, 17(special issue), 31-44. Retrieved from <http://ejournals.ukm.my/3l/article/view/986/900>
- Howarth, P. A. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tübingen: Niemeyer.
- Howarth, P. A. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161-186). Oxford: Oxford University Press.
- Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *Studies in Second Language Acquisition*, 24(3), 347-386. DOI: 10.1017/S0272263102003017

- Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13-23. DOI: 10.1016/j.system.2015.06.011
- Huat, C. M. (2012). Learner corpora and second language acquisition. In K. Hyland, C. M. Huat & M. Handford (Eds.), *Corpus applications in applied linguistics* (pp. 191-207). London/New York: Continuum.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Ishikawa, S., Uemura, T., Kaneda, M., Shimizu, S., Sugimori, N., & Tono, Y. (2003). *JACET 8000: JACET List of 8000 Basic Words*. Tokyo: JACET.
- Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning*, 67, 631-664. DOI: 10.1111/lang.12242
- Jaén, M. M. (2007). A Corpus-driven design of a test for assessing the ESL collocational competence of university students. *International Journal of English Studies*, 7(2), 127-147. Retrieved from <http://revistas.um.es/ijes/article/view/49031>
- James, C. (2013). *Errors in language learning and use: Exploring error analysis*. London/New York: Routledge.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, 50(2), 245-309. DOI: 10.1111/0023-8333.00118
- Jexenflcker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 169-190). Amsterdam/Philadelphia: John Benjamins.
- Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, 24(1), 131-161. DOI: 10.1017/S0142716403000079
- Jiménez-Catalán, R. M., & de Zarobe, Y. R. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In Y. R. de Zarobe & R. M. Jiménez-Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 81-92). Bristol: Multilingual Matters.
- Jiménez-Catalán, R. M., de Zarobe, Y. R., & Cenoz, J. (2006). Vocabulary profiles of English foreign language learners in English as a subject and as a vehicular language. *Vienna English Working Papers*, 15(3), 23-27. Retrieved from https://s3.amazonaws.com/academia.edu.documents/34290312/views15_3_clil_special.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1539076857&Signature=4mnlIfN%2FkCxyUVHUH2Dg5ot%2Bhvo%3D&response-content-disposition=inline%3B%20filename%3DThe_Acquisition_of_Knowledge_in_Bilingu.pdf#page=23
- Johnson, K., & Johnson, H. (Eds.). (1999). *Encyclopedic dictionary of applied linguistics*. Oxford: Blackwell.

- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99. DOI: 10.1016/0010-0285(89)90003-0
- Johnstone, R. (2007). Nationally-sponsored innovations at school in Scotland: Issues of evidence, generalizability and sustainability. *Innovation in Language Learning and Teaching*, 1(1), 111-128. DOI: 10.2167/illt046.0
- Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 269-300). Amsterdam/Philadelphia: John Benjamins.
- Juan-Garau, M., Salazar-Noguera, J., & Gené-Gil, M. (2013). Some insights into longitudinal EFL written development under CLIL provision: Holistic assessment of language gains and attrition. In R. Arias, M. López Rodríguez, A. Moreno Ortiz & C. Pérez Hernández (Eds.), *Hopes and fears: English and American studies in Spain. Proceedings of the 36th AEDEAN Conference* (pp. 324-330). Universidad de Málaga. Retrieved from <http://tecnolengua.uma.es/misc/AEDEAN36-Proceedings.pdf#page=326>
- Kadlecová, D. (2014). *L2 collocations and the L2 mental lexicon* (Unpublished master's thesis). Utrecht University, Utrecht.
- Kalberer, U. (2007). *Rate of L2 acquisition and the influence of instruction time on achievement*. (Unpublished master's thesis). University of Manchester, Manchester.
- Källkvist, M. (1999). *Form-class and task-type effects in learner English: A study of advanced English learners*. Lund: Lund University Press.
- Kaszubski, P. (2000). *Selected aspects of lexicon, phraseology and style in the writing of Polish advanced learners of English: A contrastive, corpus-based approach*. (Unpublished doctoral dissertation). Adam Mickiewicz University, Poznań.
- Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism Toronto*, 15, 59-92.
- Kellerman, E. (1979). Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition*, 2(1), 37-57. DOI: 10.1017/S0272263100000942
- Kellerman, E. (2000). What fruit can tell us about lexicosemantic transfer: A non-structural dimension to learners' perceptions of linguistic relations. In C. Muñoz (Ed.), *Segundas lenguas. Adquisición en la aula* (pp. 21-37). Barcelona: Ariel.
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 7-28), Stanford, California: CSLI Publications.
- Kennedy, G. (2008). Phraseology and language pedagogy. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 21-41). Amsterdam/Philadelphia: John Benjamins.
- Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17(1), 81-92. DOI: 10.1111/j.1473-4192.2007.00134.x

- Kinsella, C. (2009). *An investigation into the proficiency of successful late learners of French*. (Unpublished doctoral dissertation). Trinity College, Dublin.
- Kinsella, C., & Singleton, D. (2014). Much more than age. *Applied Linguistics*, 35(4), 441-462. DOI: 10.1093/applin/amu032
- Kizil, A. Ş., & Kilimci, A. (2014). Recurrent phrases in Turkish EFL learners' spoken interlanguage: A corpus-driven structural and functional analysis. *Journal of Language and Linguistic Studies*, 10(1), 195-210. Retrieved from <http://jlls.org/index.php/jlls/article/view/176/178>
- Klampfl, A. (2010). *A comparative study of writing proficiency between an Austrian CLIL and mainstream EFL class with regard to vocabulary*. (Unpublished doctoral dissertation). Universität Wien, Vienna.
- Klégr, A. (1997). English complex prepositions of the prepositional phrase type. *Prague Studies in English XXII, AUC, Philologica*, 5, 51-78. Retrieved from http://emsa.ff.cuni.cz/system/files/English%20Complex%20Prepositions%20of%20the%20Prepositional%20Phrase%20Type_Kl%C3%A9gr.pdf
- Knell, E., Haiyan, Q., Miao, P., Yanping, C., Siegel, L. S., Lin, Z., & Wei, Z. (2007). Early English immersion and literacy in X'ian, China. *The Modern Language Journal*, 91(3), 395-417. DOI: 10.1111/j.1540-4781.2007.00586.x
- Koizumi, R. (2012). Relationships between text length and lexical diversity measures: Can we use short texts of less than 100 tokens? *Vocabulary Learning and Instruction*, 1(1), 60-69. DOI: 10.7820/vli.v01.1.koizumi
- Kormi-Nouri, R., Shojaei, R. S., Moniri, S., Gholami, A. R., Moradi, A. R., Akbari-Zardkhaneh, S., & Nilsson, L. G. (2008). The effect of childhood bilingualism on episodic and semantic memory tasks. *Scandinavian Journal of Psychology*, 49(2), 93-109. DOI: 10.1111/j.1467-9450.2008.00633.x
- Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 131-152). Amsterdam/Philadelphia: John Benjamins.
- Koya, T. (2005). *The acquisition of basic collocations by Japanese learners of English*. (Unpublished doctoral dissertation). Waseda University, Tokyo.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149-174. DOI: 10.1006/jmla.1994.1008
- Kuo, C. L. (2009). An analysis of the use of collocation by intermediate EFL college students in Taiwan. *Arecls*, 6, 141-155. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.697.4814&rep=rep1&type=pdf>
- Kuperman, V., & Bertram, R. (2013). Moving spaces: Spelling alternation in English noun-noun compounds. *Language and Cognitive Processes*, 28(7), 939-966. DOI: 10.1080/01690965.2012.701757

- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lambelet, A., & Berthele, R. (2015). *Age and foreign language learning in school*. New York: Palgrave Macmillan.
- Langacker, R. (1987). *Foundations of cognitive grammar, Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141-165. DOI: 10.1093/applin/18.2.141
- Larsen-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research*, 24(1), 35-63. DOI: 10.1177/0267658307082981
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1(1), 30-41. DOI: 10.2174/1874913500801010030
- Lasagabaster, D. (2009). The implementation of CLIL and attitudes towards trilingualism. *ITL - International Journal of Applied Linguistics*, 157(0), 23-43. DOI: 10.2143/ITL.157.0.2042586
- Lasagabaster, D., & Doiz, A. (2003). Maturational constraints on foreign-language written production. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 136-160). Clevedon: Multilingual Matters.
- Lasagabaster D., & Sierra, J. M. (2009a). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64(4), 367-375. DOI: 10.1093/elt/ccp082
- Lasagabaster, D., & Sierra, J. M. (2009b). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal*, 1(2), 4-17. Retrieved from <https://pdfs.semanticscholar.org/e29a/ecfad1deacdc32caf59a0d23188018d8458a.pdf>
- Laufer, B. (2010). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, 24(1), 29-49. DOI: 10.1093/ijl/ecq039
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694-716. DOI: 10.1093/applin/amn018
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), 365-391. DOI: 10.1111/0023-8333.00046
- Laufer, B. & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647-672. DOI: 10.1111/j.1467-9922.2010.00621.x
- Lecocq, K., Kolinsky, R., Goetry, V., Morais, J., Alegria, J., & Mousty, P. (2009). Reading development in two alphabetic systems differing in orthographic consistency: A longitudinal

study of French-speaking children enrolled in a Dutch immersion program. *Psychologica Belgica*, 49(2&3), 111-156. Retrieved from https://www.researchgate.net/profile/Regine_Kolinsky/publication/233636207_Reading_Development_in_two_Alphabetic_Systems_Differing_in_Orthographic_Consistency_A_longitudinal_study_of_French-speaking_children_enrolled_in_a_Dutch_immersion_program/links/55437b040cf234bdb21b99f8.pdf

Lecocq, K., Mousty, P., Kolinsky, R., Goetry, V., Morais, J., & Alegria, J. (2007a). Evaluation de programmes d'immersion en Communauté française: Etude longitudinale du développement des habiletés écrites en L1 et L2. In L. Puren & S. Babault (Eds.), *L'éducation au-delà des frontières, apprentissage précoce en néerlandais dans la zone frontalière francobelge* (pp. 259-293). Paris: L'Harmattan.

Lecocq, K., Mousty, P., Kolinsky, R., Goetry, V., Morais, J., & Alegria, J. (2007b). Évaluation de programmes d'immersion en Communauté française: Une étude longitudinale comparative du développement de compétences linguistiques d'enfants francophones immergés en néerlandais. *Recherches en éducation*. Retrieved from http://www.enseignement.be/download.php?do_id=5113&do_check=

Leech, L. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning*, 50, 675-724. DOI: 10.1111/0023-8333.00143

Lemmens, M. (2002). The semantic network of posture verbs. In J. Newman (Ed.), *The linguistics of sitting, standing, and lying* (pp. 103-140). Amsterdam/Philadelphia: John Benjamins.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12), 59-67. DOI: 10.1080/21548331.1967.11707799

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove: Language Teaching Publications.

Lewis, M. (Ed.). (2000). *Teaching collocation. Further developments in the lexical approach*. Boston: Thomson Heinle.

Lewis, M. (2002). *Implementing the lexical approach*. Boston: Thomson Heinle.

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85-102. DOI: 10.1016/j.jslw.2009.02.001

Li, J., & Schmitt, N. (2010). The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 22-46). New York: Continuum.

Lin, P. M. S. (2010). The phonology of formulaic sequences: A review. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 174-193). New York: Continuum.

Lin, P. M. S. (2014). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System*, 42, 164-176. DOI: 10.1016/j.system.2013.11.010

- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861-883. DOI: 10.3758/s13423-013-0565-2
- Lindstromberg, S., & Boers, F. (2008a). Phonemic repetition and the learning of lexical chunks: The power of assonance. *System*, 36(3), 423-436. DOI: 10.1016/j.system.2008.01.002
- Lindstromberg, S., & Boers, F. (2008b). The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics*, 29(2), 200-222. DOI: 10.1093/applin/amn007
- Liu, C. P. (1999). An analysis of collocational errors in EFL writings. In *The Proceedings of the Eighth International Symposium on English Teaching* (pp. 483-494). Taipei: Crane.
- Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, 13(4), 491-518. DOI: 10.1075/ijcl.13.4.05liu
- Llinares, A., Morton, T., & Whittaker, R. (2012). *The role of language in CLIL*. Cambridge: Cambridge University Press.
- Llinares, A., & Whittaker, R. (2010). Writing and speaking in the history class: A comparative analysis of CLIL and first language contexts. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 125-144). Amsterdam/Philadelphia: John Benjamins.
- Lo, Y. Y., & Murphy, V. A. (2010). Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong-Kong. *Language and Education*, 24(3), 215-238. DOI: 10.1080/09500780903576125
- Lochtmann, K. (2007). Die mündliche Fehlerkorrektur in CLIL und im traditionellen Fremdsprachenunterricht: Ein Vergleich. [Oral corrective feedback in CLIL and traditional foreign language teaching: A comparison]. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical perspectives on CLIL classroom discourse* (pp. 119-138). Frankfurt, Germany: Peter Lang.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition* (pp. 413-468). New York: Academic.
- Long, M. H. (2000). Focus on form in task-based language teaching. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy. Essays in honor of A. Ronald Walton* (pp. 179-192). Philadelphia/Amsterdam: John Benjamins.
- Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford: Blackwell.
- Lorenzo, F., Casal, S., & Moore, P. (2005). *Orientaciones para la elaboración del currículo integrado de las lenguas en los centros bilingües*. Sevilla: Consejería de Educación (Junta de Andalucía).
- Lorette, P., & Mettewie, L. (2013). CLIL in Wallonië: Franstaligen leren graag Nederlands. *Levende Talen Magazine*, 100(7), 26-29.

- Lowie, W. (2017). Emergentism: Wide ranging theoretical framework or just one more meta-theory? *Recherches en didactique des langues et des cultures*, 14(1), 1-9. DOI: 10.4000/rdlc.1101
- Lowie, W., & Verspoor, M. (2015). Variability and variation in second language acquisition orders: A dynamic reevaluation. *Language Learning*, 65(1), 63-88. DOI: 10.1111/lang.12093
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432. DOI: 10.1017/S0272263104263021
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam/Philadelphia: John Benjamins.
- MacWhinney, B. (Ed.). (1987). *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. (Ed.). (1999). *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341-371). New York: Routledge.
- Maljers, A. (2007). The Netherlands. In A. Maljers, D. Marsh & D. Wolff (Eds), *Windows on CLIL. Content and language integrated learning in the European spotlight* (pp. 130-137). Graz: European Centre for Modern Languages.
- Manning, C., & Schütze, H. (1999). *Collocations. Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Marsh, D. (2002). *CLIL/EMILE – The European dimension: Actions, trends and foresight potential*. University of Jyväskylä, Finland: Continuing Education Centre. Retrieved from <https://jyx.jyu.fi/dspace/handle/123456789/47616>
- Marsh, D., & Wolff, D. (Eds.). (2007). *Diverging contexts – diverging goals. CLIL in Europe*. Frankfurt: Peter Lang.
- Martínez Adrián, N., & Gutiérrez Mandago, M. J. (2015). L1 use, lexical richness, accuracy and syntactic complexity in the oral production of CLIL and NON-CLIL learners of English. *Atlantis* (0210-6124), 37(2), 175-197.
- Mattheoudakis, M., Alexiou, T., & Laskaridou, C. (2014). To CLIL or not to CLIL? The case of the 3rd experimental primary school in Evosmos. In N. Lavidas, T. Alexiou & A. M. Sougari (Eds.), *Major trends in theoretical and applied linguistics 3* (pp. 215-234). London: Versita.
- McCarthy, M., & Carter, R. (2002). This, that and the other: Multi-word clusters in spoken English as visible patterns of interaction. *Teanga: The Irish Yearbook of Applied Linguistics*, 21, 30-52.
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392. DOI: 10.3758/BRM.42.2.381
- McLaughlin, B. (1995). *Fostering second language development in young children: Principles and practices*. UC Berkely: Center for Research on Education, Diversity and Excellence. Retrieved from <https://escholarship.org/uc/item/23s607sr>

- McLeay, H. (2003). The relationship between bilingualism and the performance of spatial tasks. *International Journal of Bilingual Education and Bilingualism*, 6(6), 423-438. DOI: 10.1080/13670050308667795
- Meara, P., Lightbown, P., & Halter, R. (1997). Classrooms as lexical environments. *Language Teaching Research*, 1, 28-47. DOI: 10.1177/136216889700100103
- Mehisto, P., & Marsh, D. (2011). Approaching the economic, cognitive and health benefits of bilingualism: Fuel for CLIL. In Y. R. de Zarobe, J. M. Sierra & F. G. del Puerto (Eds.), *Content and foreign language integrated learning. Contributions to multilingualism in European contexts* (pp. 21-47). Bern: Peter Lang.
- Mel'čuk, I. (1995). Phrasemes in language and phraseology in linguistics. In M. Everaert, E. J. Van der Linden & A. Schenk (Eds.), *Idioms: Structural and psychological perspectives* (pp. 167-232). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mel'čuk, I. (1998). Collocations and lexical functions. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 24-53). Oxford: Oxford University Press.
- Mellow, J. D. (2006). The emergence of second language syntax: A case study of the acquisition of relative clauses. *Applied Linguistics*, 27, 645-670. DOI: 10.1093/applin/aml031
- Merino, J. A., & Lasagabaster, D. (2015). CLIL as a way to multilingualism. *International Journal of Bilingual Education and Bilingualism*, 21(1), 79-92. DOI: 10.1080/13670050.2015.1128386
- Merisuo-Storm, T. (2002). *Development of L1 reading and writing in bilingual education*. University of Turku, Turku.
- Merisuo-Storm, T., & Soininen, M. (2014). Students' first language skills after six years in bilingual education. *Mediterranean Journal of Social Sciences*, 5(22), 72-81. DOI: 10.5901/mjss.2014.v5n22p72
- Met, M. (1999). *Content-based instruction: Defining terms, making decisions*. NFLC Reports. Washington, DC: The National Foreign Language Center.
- Mettewie, L. (2004). *Attitudes en motivatie van taalleerders in België. Een sociaal-psychologisch onderzoek naar het verwerven van de eerste en tweede taal door Nederlandstalige, Franstalige en tweetalige leerlingen in het secundair onderwijs in België*. (Unpublished doctoral dissertation). Vrije Universiteit Brussel, Brussel.
- Mettewie, L. (2015). Apprendre la langue de "l'Autre" en Belgique: La dimension affective comme frein à l'apprentissage. *Le Langage et l'Homme*, 50(2), 23-42.
- Mettewie, L., & Lorette, P. (2014). Tweetalig onderwijs als Belgische «philtre d'amour». Attitudes van Franstalige CLIL-leerders van het Nederlands. In L. Degand, P. Hilgsmann, L. Rasier, M. Sergier, S. Vanasten & K. Van Goethem (Eds.), *In het teken van identiteit. Taal en cultuur van de Nederlanden* (pp. 195-201). Louvain-la-Neuve: Presses universitaires de Louvain.
- Mettewie, L., & Van Mensel, L. (2009). Multilingualism at all costs: Language use and language needs in business in Brussels. *Sociolinguistica*, 23, 131-149. Retrieved from https://www.researchgate.net/profile/Luk_Van_Mensel/publication/284592404_Multilinguali

[sm at all Costs Language Use and Language Needs in Business in Brussels/links/5940ff4aa6fdcccda9b5e260/Multilingualism-at-all-Costs-Language-Use-and-Language-Needs-in-Business-in-Brussels.pdf](https://www.researchgate.net/publication/314444444/Multilingualism-at-all-Costs-Language-Use-and-Language-Needs-in-Business-in-Brussels.pdf)

Meunier, F. (forthcoming). Introduction to learner corpus research. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge handbook of SLA and corpora*.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111-129. DOI: 10.1017/S0267190512000128

Meunier, F. (2015). Developmental patterns in learner corpora. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 379-400). Cambridge: Cambridge University Press.

Meunier, F., & Bulon, A. (in preparation). Using two different taxonomies of MWUs to assess learners' productive EFL phrasicon in CLIL and traditional EFL settings: On the benefits of comparing apples and oranges.

Meunier, F., & Granger, S. (2008). *Phraseology in foreign language learning and teaching*. Amsterdam/Philadelphia: John Benjamins.

Meunier, F., & Littré, D. (2013). Tracking learners' progress: Adopting a dual 'Corpus cum experimental data' approach. *The Modern Language Journal*, 97(S1), 61-76. DOI: 10.1111/j.1540-4781.2012.01424.x

Mewald, C. (2007). A comparison of oral language performance of learners in CLIL and in mainstream classes at lower secondary level in lower Austria. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical perspectives on CLIL classroom discourse* (pp. 139-178). Frankfurt: Peter Lang.

Miralpeix, I. (2006). Age and vocabulary acquisition in English as a foreign language. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 89-106). Clevedon: Multilingual Matters.

Möller, V. (2017). *Language acquisition in CLIL and NON-CLIL settings: Learner corpus and experimental evidence on passive constructions*. Amsterdam/Philadelphia: John Benjamins.

Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology. Theory, analysis, and applications* (pp. 79-100). Oxford: Oxford University Press.

Mora, J. C. (2006). Age effects on oral fluency development. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 65-88). Clevedon: Multilingual Matters.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933-947. DOI: 10.1162/jocn_a_00119

Morris, W. C., Cottrell, G. W., & Elman, J. L. (2000). A connectionist simulation of the empirical acquisition of grammatical relations. In S. Wermter & R. Sun (Eds.), *Hybrid neural symbolic integration* (pp. 175-193). Berlin: Springer.

Moyer, A. (2009). Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In T. Piske & M. Young-Sholten (Eds.), *Input matters in SLA* (pp. 159-174). Bristol: Multilingual Matters.

Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge: Cambridge University Press.

Moyer, A. (2014). What's age got to do with it? Accounting for individual factors in second language accent. *Studies in Second Language Learning and Teaching*, 3(4), 443-464. Retrieved from <https://www.cceol.com/search/article-detail?id=115831>

Mukherjee, J. (2009). The grammar of conversation in advanced spoken learner English: Learner corpus data and language-pedagogical implications. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 203-230). Amsterdam/Philadelphia: John Benjamins.

Muñoz, C. (2003). Variation in oral skills and age of onset. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 101-181). Clevedon: Multilingual Matters.

Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning*. Clevedon: Multilingual Matters.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 24(4), 578-596. DOI: 10.1093/applin/amm056

Muñoz, C. (2011). Input and long-term effects of starting age in foreign language learning. *International Review of Applied Linguistics in Language Teaching*, 49(2), 113-133. DOI: 10.1515/iral.2011.006

Muñoz, C. (Ed.). (2012). *Intensive exposure experiences in second language learning*. Bristol: Multilingual Matters.

Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, 35(4), 463-482. DOI: 10.1093/applin/amu024

Mutlu, G., & Kaşlıoğlu, Ö. (2016). Vocabulary size and collocational knowledge of Turkish EFL learners. *Journal of Theory and Practice in Education*, 12(6), 1231-1252. Retrieved from http://acikerisim.lib.comu.edu.tr:8080/xmlui/bitstream/handle/COMU/1651/Gizem_Mutlu_Makale.pdf?sequence=1

Myles, F. (2004). From data to theory: The over-representation of linguistic knowledge in SLA. *Transactions of the Philological Society*, 102(2), 139-168. DOI: 10.1111/j.0079-1636.2004.00133.x

Myles, F. (2017). Learning foreign languages in primary schools: Is younger better? *Journal of Languages, Society and Policy*, 1(1), 1-8. Retrieved from <https://www.repository.cam.ac.uk/bitstream/handle/1810/264340/Myles%20JLSP%202017%20.pdf?sequence=6>

Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323-364. DOI: 10.1111/0023-8333.00045

- Myles, F., Mitchell, R., & Hooper, J. (1999). Interrogative chunks in French L2: A basis for creative construction? *Studies in Second Language Acquisition*, 21(1), 49-80. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/C089B158322237452E3FEF43ED4F1214/S0272263199001023a.pdf/interrogative_chunks_in_french_l2.pdf
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82. DOI: 10.3138/cmlr.63.1.59
- Nattinger, J. R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, 14, 337-344. DOI: 10.2307/3586598
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Navés, T., Miralpeix, I., & Celaya, M. L. (2005). Who transfers more ... and what? Cross-linguistic influence in relation to school grade and language dominance in EFL. *International Journal of Multilingualism*, 2(2), 113-134. DOI: 10.1080/14790710508668380
- Navés, T., Torras, M. R., & Celaya, M. L. (2003). Long-term effects of an earlier start. An analysis of EFL written production. *EUROSLA Yearbook*, 3(1), 103-130. DOI: 10.1075/eurosla.3.08nav
- Neff van Aertselaer, J. (2008). Contrasting English-Spanish interpersonal discourse phrases: A corpus study. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 85-99). Amsterdam/Philadelphia: John Benjamins.
- Nehls, D. (1991). English DO/MAKE compared with German TUN/MACHEN and Dutch DOEN/MAKEN: A synchronic-diachronic approach. *International Review of Applied Linguistics in Language Teaching*, 29(4), 303-316. DOI: 10.1515/iral.1991.29.4.303
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242. DOI: 10.1093/applin/24.2.223
- Nesselhauf, N. (2004). What are collocations? In D. J. Allerton, N. Nesselhauf & P. Skandera (Eds.), *Phraseological units: Basic concepts and their application* (pp. 1-21). Basel: Schwabe.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam/Philadelphia: John Benjamins.
- Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, 21(3), 298-320. DOI: 10.1177/1362168816639619
- Nieto-Moreno-de-Diezmas, E. (2018). Acquisition of reading comprehension in L1 in bilingual programmes of primary education. A comparative study. *Ocnos*, 17(1), 43-54. DOI: 10.18239/ocnos_2018.17.1.1471

- Nikula, T. (2010). Effects of CLIL on a teacher's classroom language use. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 105-124). Amsterdam/Philadelphia: John Benjamins.
- Oakes, M. P. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erlbaum.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232-258. DOI: 10.1016/j.cogpsych.2012.12.002
- Pallotti, G. (2007). An operational definition of the emergence criterion. *Applied Linguistics*, 28(3), 361-382. DOI: 10.1093/applin/amm018
- Pander Maat, H., Kraf, R., Dekker, N., Sloot, K. van der, Bosch, A. van den, Gompel, M. van, & Klein, S. (2014). *Handleiding T-Scan*. Retrieved from <http://webservices-lst.science.ru.nl/>
- Paquot, M. (2005). *Towards a productively-oriented academic word list*. Presented at Practical Applications in Language and Computers. Łódź, Poland. April 7-9. Retrieved from <https://dial.uclouvain.be/pr/boreal/fr/object/boreal%3A76045>
- Paquot, M. (2008). Exemplification in learner writing: A cross-linguistic perspective. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 101-119). Amsterdam/Philadelphia: John Benjamins.
- Paquot, M. (2019). The phraseological dimension in interlanguage complexity research. *Second Language Research*, 35(1), 121-145. DOI: 10.1177/0267658317694221
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149. DOI: 10.1017/S0267190512000098
- Paquot, M., Naets, H., & Gries, S. T. (forthcoming). Using syntactic co-occurrences to trace phraseological complexity development in learner writing: verb + object structures in LONGDALE. In B. Le Bruyn & M. Paquot (Eds.), *Second language acquisition and learner corpora*. Cambridge: Cambridge University Press
- Pasquarella, A., Chen, X., Lam, K., Luo, Y. C., & Ramirez, G. (2011). Cross-language transfer of morphological awareness in Chinese-English bilinguals. *Journal of Research in Reading*, 34(1), 23-42. DOI: 10.1111/j.1467-9817.2010.01484.x
- Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30(2), 449-468. DOI: 10.1111/j.1467-1770.1980.tb00328.x
- Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics*, 7(5), 551-579. DOI: 10.1016/0378-2166(83)90081-4

- Pecman, M. (2008). Compilation, formalization and presentation of bilingual phraseology: Problems and possible solutions. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 203-222). Amsterdam/Philadelphia: John Benjamins.
- Penfield, W. (1963). *The second career*. Boston: Little, Brown.
- Perales, S., García Mayo, M. P., & Licerias, J. M. (2004). *The acquisition of English sentential negation by bilingual (Basque/Spanish) children: The age factor in an institutional setting*. Presented at the Annual Meeting of the American Association of Applied Linguistics. Portland, OR. May 1-4.
- Pérez Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction*, 57, 18-33. DOI: 10.1016/j.learninstruc.2017.12.002
- Pérez Cañado, M. L., & Ráez Padilla, J. (2015). Introduction and overview. In D. Marsh, M. L. Pérez Cañado & J. Ráez Padilla (Eds.), *CLIL in action: Voices from the classroom* (pp. 1-12). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Pérez-Vidal, C. (2007). The need for focus on form (FoF) in content and language integrated approaches: An exploratory study. *Revista española de lingüística aplicada*, (1), 39-54. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=2575491>
- Pfenninger, S. E., & Singleton, D. (2017). *Beyond age effects in instructional L2 learning: Revisiting the age factor*. Bristol: Multilingual Matters.
- Pierce, A. E. (2012). *Language acquisition and syntactic theory: A comparative analysis of French and English child grammars*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Pine, J. M., & Lieven, E. V. M. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics*, 18(2), 123-138. DOI: 10.1017/S0142716400009930
- Pinker, S. (1999). *Words and rules*. New York: Morrow.
- Pirskanen, H. (2009). *Attitudes towards the English language: A comparison between a CLIL class and a regular class*. (Unpublished bachelor's dissertation). University of Jyväskylä, Finland.
- Piske, T., & Young-Scholten, M. (Eds.). (2008). *Input matters in SLA*. Bristol: Multilingual Matters.
- Pladevall-Ballester, E. (2015). Exploring primary school CLIL perceptions in Catalonia: Students', teachers' and parents' opinions and expectations. *International journal of bilingual education and bilingualism*, 18(1), 45-59. DOI: 10.1080/13670050.2013.874972
- Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R² values. *The Modern Language Journal*, 102(4), 713-731. DOI: 10.1111/modl.12509
- Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, 36(1), 94-120. DOI: 10.1177/0149206309352110

- Porto, M. (1998). Lexical phrases and language teaching. *Forum*, 36(3), 22-26. Retrieved from <http://exchanges.state.gov/forum/vols/vol36/no3/p22.htm>
- Pritchett, L., Vaid, J., & Tosun, S. (2016). Of black sheep and white crows: Extending the bilingual dual coding theory to memory for idioms. *Cogent Psychology*, 3(1), 1-18. DOI: 10.1080/23311908.2015.1135512
- Rankin, T. (2015). Learner corpora and grammar. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 231-254). Cambridge: Cambridge University Press.
- Rasier, L., Bui, A. V., Jouniaux, A., & Hiligsmann, P. (2014). Klemtoon in het Nederlands van Franstalige immersie- en niet-immersieerlingen. In L. Degand, P. Hiligsmann, L. Rasier, M. Sergier, S. Vanasten & K. Van Goethem (Eds.), *In het teken van identiteit. Taal en cultuur van de Nederlanden* (pp. 203-220). Louvain-la-Neuve: Presses universitaires de Louvain.
- Raven, J. C., Court, J. H., & Raven, J. (1998). *Progressive coloured matrices*. Oxford: Oxford Psychologists Press.
- Reppen, R. (2009). Exploring L1 and L2 writing development through collocations: A corpus-based look. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 49-59). London: Palgrave Macmillan.
- Rica-Peromingo, J. P. (2009). The use of lexical bundles in the written production of Spanish EFL university students. *Applied Linguistics for Specialized Discourse*, 1, 1-7. Retrieved from https://www.researchgate.net/profile/Juan_Pedro_Rica-Peromingo/publication/303818351_The_use_of_lexical_bundles_in_the_written_production_of_Spanish_EFL_university_students/links/57557ea708ae0405a575369f.pdf
- Richards, J. C. (1974). A non-contrastive approach to error analysis. In J. C. Richards (Ed.), *Error analysis: Perspectives on second language acquisition* (pp. 172-188). New York: Routledge.
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47(1), 45-99. DOI: 10.1111/0023-8333.21997002
- Robinson, P., & Ellis, N. C. (2008). Conclusion: Cognitive linguistics, second language acquisition and L2 instruction – Issues for research. In P. Robinson & N. C. Ellis (Eds.), *A handbook of cognitive linguistics and SLA* (pp. 489-545). Mahwah, NJ: Lawrence Erlbaum.
- Roever, C. (2011). What learners get for free: Learning of routine formulae in ESL and EFL environments. *ELT journal*, 66(1), 10-21. DOI: 10.1093/elt/ccq090
- Roquet i Pugès, H. (2011). A study of the acquisition of English as a foreign language: Integrating content and language in mainstream education in Barcelona. (Unpublished doctoral dissertation). Universitat Pompeu Fabra, Barcelona.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605. DOI: 10.1016/0010-0285(75)90024-9
- Rumlich, D. (2016). *Evaluating bilingual education in Germany: CLIL students' general English proficiency, EFL self-concept and interest*. Frankfurt am Main: Lang.

- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 1-31). Oxford: Oxford University Press.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action. An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 1-22). Amsterdam/Philadelphia: John Benjamins.
- Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 127-151). Amsterdam/Philadelphia: John Benjamins.
- Schmitt, N., & Redwood, S. (2011). Learner knowledge of phrasal verbs: A corpus-informed study. In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds.), *A taste for corpora. In honour of Sylviane Granger* (pp. 173-208). Amsterdam/Philadelphia: John Benjamins.
- Schmitt, N., & Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 173-189). Amsterdam/Philadelphia: John Benjamins.
- Schuitmaker-King, J. (2012). *Teachers' strategies in providing opportunities for second language development*. (Unpublished doctoral dissertation). University of Groningen, The Netherlands.
- Scott, M. (2012). *WordSmith Tools*. Liverpool: Lexical Analysis Software.
- Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, 21(4), 328-341. DOI: 10.2167/le635.0
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-231. DOI: 10.1515/iral.1972.10.1-4.209
- Selinker, L., & Lakshmanan, U. (1992). Language transfer and fossilization: The multiple effects principle. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 197-216). Amsterdam/Philadelphia: John Benjamins.
- Shehata, A. K. (2008). L1 Influence on the reception and production of collocations by advanced ESL/EFL Arabic learners of English. (Unpublished doctoral dissertation). Ohio University, Athens.
- Shin, D. (2007). The high frequency collocations of spoken and written English. *English Teaching*, 62(1), 199-218. Retrieved from http://kate.bada.cc/wp-content/uploads/2015/02/kate_62_1_9.pdf
- Shirato, J., & Stapleton, P. (2007). Comparing English vocabulary in a spoken learner corpus with a native speaker corpus: Pedagogical implications arising from an empirical study in Japan. *Language Teaching Research*, 11(4), 393-412. DOI: 10.1177/1362168807080960
- Siepmann, D. (2005a). Collocation, colligation and encoding dictionaries. Part I: lexicological aspects. *International Journal of Lexicography*, 18(4), 409-443. DOI: 10.1093/ijl/eci042
- Siepmann, D. (2005b). *Discourse markers across languages: A contrastive study of second-level discourse markers in native and non-native text with implications for general and pedagogic lexicography*. London: Routledge.

- Siepmann, D. (2008). Phraseology in learners' dictionaries: What, where and how. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 185-202). Amsterdam/Philadelphia: John Benjamins.
- Simonis, M. (2019). The cognitive profile of children enrolled in Content and Language Integrated Learning in French-speaking Belgium. (Unpublished doctoral dissertation). Université catholique de Louvain, Louvain-la-Neuve.
- Sinclair, J. (Ed.). (1987). *Looking up. An account of the COBUILD project in lexical computing*. London: Collins Cobuild.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2008). Preface. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. xv-xix). Amsterdam/Philadelphia: John Benjamins.
- Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.
- Siyanova-Chanturia, A. (2015). Collocation in beginner learner writing: A longitudinal study. *System*, 53, 148-160. DOI: 10.1016/j.system.2015.07.003
- Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multi-word sequences. *Journal of Experimental Psychology: Language, Memory and Cognition*, 37, 776-784. DOI: 10.1037/a0022531
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2018). Formulaic language: Setting the scene. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 1-15). New-York/London: Routledge.
- Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics in Language Teaching*, 45(2), 119-139. DOI: 10.1515/IRAL.2007.005
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking”. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge: Cambridge University Press.
- Snow, C. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of language* (pp. 257-276). Mahwah, NJ: Erlbaum
- Snow, C., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 49(4), 1114-1128. DOI: 10.2307/1128751
- Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in the development of accuracy and complexity: A longitudinal case study on the acquisition of Finnish. *Applied Linguistics*, 31(4), 532-553. DOI: 10.1093/applin/amq001
- Stefanowitsch, A., & Gries, S. T. (2003). Collocations: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209-243. DOI: 10.1075/ijcl.8.2.03ste

- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *IRAL - International Review of Applied Linguistics in Language Teaching*, 49(4), 321-343. DOI: 10.1515/iral.2011.017
- Strömqvist, S., Holmqvist, K., Johansson, V., Karlsson, H., & Wengelin, Å. (2006). What keystroke-logging can reveal about writing. In K. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging and writing: Methods and applications* (pp. 45-72). Amsterdam: Elsevier.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
- Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, 7(2), 215-244. DOI: 10.1075/ijcl.7.2.04stu
- Sundkvist, P., Sylvén, L. K. (2016). *Extramural English in teaching and learning: From theory and research to practice*. London: Palgrave Macmillan.
- Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, 67(4), 747-790. DOI: 10.1111/lang.12241
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of HG Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M., & Lapkin, S. (1989). Canadian immersion and adult second language teaching: What's the connection? *The Modern Language Journal*, 73(2), 150-159. DOI: 10.1111/j.1540-4781.1989.tb02537.x
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Sylvén, L. K. (2019). *Investigating Content and Language Integrated Learning: Insights from Swedish high schools. Bilingual education and bilingualism*. Bristol: Multilingual Matters.
- Sylvén, L. K., & Ohlander, S. (2015). The CLISS project: Receptive vocabulary in CLIL versus non-CLIL groups. *Moderna språk*, 108(2), 80-114. Retrieved from <http://130.241.16.45/ojs/index.php/modernasprak/article/view/2789/2614>
- Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly*, 15(3), 285-295. DOI: 10.2307/3586754
- Taylor, J. R. (1998). Syntactic constructions as prototype categories. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 177-202). Mahwah, NJ: Erlbaum.
- Tedick, D. J., & Cammarata, L. (2012). Content and language integration in K-12 contexts: Student outcomes, teacher practices, and stakeholder perspectives. *Foreign Language Annals*, 45(1), 28-53. DOI: 10.1111/j.1944-9720.2012.01178.x
- Tedick, D. J., Christian, D., & Fortune, T. W. (2011). *Immersion education. Practices, policies, possibilities*. Bristol: Multilingual matters.

- Terrazas Gallego, M., & Agustín Llach, M. P. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study. *International Journal of English Studies*, 9(1), 113-133. Retrieved from <https://revistas.um.es/ijes/article/view/90681/87481>
- Terrell, T. D. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal*, 75(1), 52-63. DOI: 10.1111/j.1540-4781.1991.tb01083.x
- Thewissen, J. (2008). The phraseological errors of French-, German-, and Spanish speaking EFL learners: Evidence from an error-tagged learner corpus. In *Proceedings from the 8th Teaching and Language Corpora Conference (TaLC8)* (pp. 300-306). Lisbon: Associação de Estudos e de Investigação Científica do ISLA-Lisboa. Retrieved from <http://anafrankenberg.synthasite.com/resources/TaLCLisbon2008Proceedings.pdf>
- Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal*, 97, 77-101. DOI: 10.1111/j.1540-4781.2012.01422.x
- Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. Heredia & A. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 171-207). Cambridge: Cambridge University Press.
- Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, 31, 101-121. DOI: 10.1017/S0305000903005944
- Torras, M. R., & Celaya, M. L. (2001). Age-related differences in the development of written production. An empirical study of EFL school learners. *International Journal of English Studies*, 1(2), 103-126. Retrieved from <https://revistas.um.es/ijes/article/view/48211>
- Torras, M. R., Navés, T., Celaya, M. L., & Pérez-Vidal, C. (2006). Age and IL development in writing. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 156-182). Clevedon: Multilingual Matters.
- Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61(2), 569-613. DOI: 10.1111/j.1467-9922.2010.00622.x
- Tummers, J., Heylen, K., & Geeraerts, D. (2005). Usage-based approaches in cognitive linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1(2), 225-261. DOI: 10.1515/cllt.2005.1.2.225
- Turnbull, M., Lapkin, S., Hart, D., & Swain, M. (1998). Time on task and immersion graduates' French proficiency. In S. Lapkin (Ed.), *French second language education in Canada: Empirical studies* (pp. 31-55). Toronto: University of Toronto Press.
- Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. *Annual Review of Applied Linguistics*, 30, 270-291. DOI: 10.1017/S0267190510000140

Ullman, M. T. (2006). Language and the brain. In J. Connor-Linton & R. W. Fasold (Eds.), *An introduction to language and linguistics* (pp. 235-274). Cambridge: Cambridge University Press.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 153-172). Amsterdam/Philadelphia: John Benjamins.

Van de Craen, P., Lochtman, K., Ceuleers, E., Mondt, K., & Allain, L. (2007). An interdisciplinary approach to CLIL learning in primary schools in Brussels. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical perspectives on CLIL classroom discourse* (pp. 253-274). Frankfurt am Main: Peter Lang.

Van de Craen, P., Mondt, K., Allain, L., & Gao, Y. (2007). Why and how CLIL works. An outline for a CLIL theory. *Vienna English Working Papers*, 16(3), 70-78. Retrieved from https://www.researchgate.net/profile/Piet_Craen/publication/290828891_Why_and_how_CLIL_works_An_outline_for_a_CLIL_theory/links/5743f23f08ae9ace841b42cb/Why-and-how-CLIL-works-An-outline-for-a-CLIL-theory.pdf

Van de Craen, P., Mondt, K., Ceuleers, E., & Migom, E. (2010). EMILE a douze ans. Douze ans d'enseignement immersif en Belgique. Résultats et perspectives. *Synergies-Monde*, 7, 127-140.

Van de Craen, P., Surmont, J., Ceuleers, E., & Allain, L. (2013). How policies influence multilingual education and the impact of multilingual education on practices. In A. C. Berthoud, F. Grin & G. Lüdi (Eds.), *Exploring the dynamics of multilingualism. The DYLAN project* (pp. 343-372). Amsterdam/Philadelphia: John Benjamins.

Vandeputte, K. (2011). Améliorer l'enseignement en immersion, cela passe aussi par un besoin urgent de matériel didactique. In L. Beheydt & P. Hiligsmann (Eds.), *Met immersie aan de slag/Au travail, en immersion* (pp. 111-118). Louvain-la-Neuve: Presses universitaires de Louvain.

Van Lancker, D., & Canter, G. J. (1981). Idiomatic versus literal interpretations of ditropically ambiguous sentences. *Journal of Speech and Hearing Research*, 24, 64-69. DOI: 10.1044/jshr.2401.64

Van Lancker, D., Canter, G. J., & Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research*, 24, 330-335. DOI: 10.1044/jshr.2403.330

Van Lancker-Sidtis, D. (2003). Auditory recognition of idioms by first and second speakers of English: It takes one to know one. *Applied Psycholinguistics*, 24, 45-57. DOI: 10.1017/S0142716403000031

Van Mensel, L., Bulon, A., Hendrikx, I., Meunier, F., & Van Goethem, K. (submitted). Effects of input on L2 writing in English and Dutch: CLIL and non-CLIL learners in French-speaking Belgium. *Journal of Immersion and Content-Based Language Education*.

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*. DOI: 10.1080/07908318.2019.1571078

VanParys, J., Stumpel, R., Boers, T., & de Bot, V. (2008). *KONNEX: Uitdrukkingen en vaste verbindingen in de Nederlandse taal*. Utrecht: ThiemeMeulenhoff.

Verspoor, M., & Edelenbos, P. (2011). Tweektalig onderwijs zorgt voor een duurzame voorsprong. *Levende Talen Tijdschrift*, 12(4), 5-13. Retrieved from <http://www.lt-tijdschriften.nl/ojs/index.php/ltt/article/view/69/68>

Verspoor, M., & Smiskova, H. (2012). Foreign language writing development from a dynamic usage based perspective. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 17-46). Boston/Berlin: De Gruyter.

Vesterbacka, S. (1991). Ritualized routines and L2 acquisition: Acquisition strategies in an immersion program. *Journal of Multilingual and Multicultural Development*, 12(1-2), 35-43. DOI: 10.1080/01434632.1991.9994443

Vinogradov, V. (1947). *On the basic types of phraseological units in Russian*. Moscow: Izd-vo AN SSSR.

Waibel, B. (2008). *Phrasal verbs: German and Italian learners of English compared*. Saarbrücken: VDM.

Walsh, S. (2010). What features of spoken and written corpora can be exploited in creating language teaching materials and syllabuses? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 333-344). London/New York: Routledge.

Wang, Y., & Shaw, P. (2008). Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME Journal*, 32, 201-232. Retrieved from https://www.researchgate.net/profile/Philip_Shaw6/publication/241014504_Transfer_and_universality_Collocation_use_in_advanced_Chinese_and_Swedish_learner_English/links/53f596530cf22be01c3f892b/Transfer-and-universality-Collocation-use-in-advanced-Chinese-and-Swedish-learner-English.pdf

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *Tesol Quarterly*, 43(1), 55-77. DOI: 10.1002/j.1545-7249.2009.tb00227.x

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocations. *Language Learning*, 63(1), 91-120. DOI: 10.1111/j.1467-9922.2012.00729.x

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16(2), 180-205. DOI: 10.1093/applin/16.2.180

Wesche, M. B. (2002). Early French immersion: How has the original Canadian model stood the test of time? In P. Thorsten, A. Rohde, H. Wode & P. Burmeister (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 357-378). Trier: Wissenschaftlicher Verlag Trier.

Wible, D. (2008). Multiword expressions and the digital turn. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 163-181). Amsterdam/Philadelphia: John Benjamins.

- Wichmann, A. (2008). Speech corpora and spoken corpora. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp. 187-206). Berlin/New-York: Walter de Gruyter.
- Wichmann, A., & Fligelstone, S. (2014). *Teaching and language corpora*. London: Routledge.
- Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, 21(1), 1-48. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/E364B6EBFDED1295B4D9FEEA5E076429/S0272263199001011a.pdf/memory_attention_and_inductive_learning.pdf
- Williams, M. (1994). Motivation in foreign and second language learning: An interactive perspective. *Educational and Child Psychology*, 11, 77-84.
- Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32(4), 430-449. DOI: 10.1093/applin/amr011
- Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, 35(3), 451-482. DOI: 10.1017/S0272263113000107
- Wong-Fillmore, L. (1976). *The second time around*. Stanford, CA: Stanford University.
- Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 20(1), 1-15. DOI: 10.18806/tesl.v20i1.935
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, 32, 231-254. DOI: 10.1017/S026719051200013X
- Wray, A., & Perkins, M. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20, 1-28. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.558.6248&rep=rep1&type=pdf>
- Wulff, S. (2018). Acquisition of formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 19-37). New-York/London: Routledge.
- Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647-668. DOI: 10.5054/tq.2010.235998
- Yan, H. (2010). Study on the causes and countermeasures of the lexical collocation mistakes in college English. *English Language Teaching*, 3(1), 162-165. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1081497.pdf>

- Yi, W. (2018). Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition*, 40(4), 831-856. DOI: <https://doi.org/10.1017/S0272263118000141>
- Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31(2), 236-259. DOI: 10.1093/applin/amp024
- Zhang, T., van Heuven, W., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science*, 22(10), 1237-1242. DOI: 10.1177/0956797611421492
- Zydati, W. (2007). *Deutsch-Englische Zge in Berlin (DEZIBEL). Eine Evaluation des bilingualen Sachfachunterrichts in Gymnasien: Kontext, Kompetenzen, Konsequenzen*. Frankfurt am Main: Peter Lang.
- Zyzik, E. (2006). Transitivity alternations and sequence learning: Insights from L2 Spanish production data. *Studies in Second Language Acquisition*, 28(3), 449-485. DOI: 10.1017/S0272263106060190
- Zyzik, E. (2009). The role of input revisited: Nativist versus usage-based models. *L2 Journal*, 1, 42-61. DOI: 10.5070/l2.v1i1.9056

Appendices

1. Writing task (English/Dutch)

1.1. Guidelines e-mail about the holidays (English)

Write an e-mail in English (15 to 25 lines) to a friend to tell him/her about your holidays.

Describe what you liked most, and tell about the things that were beautiful or funny. Also, tell him/her about things that happened and that you did not like at all.

You may also want to talk about the persons you met, the weather, the problems you encountered, the overall atmosphere, the food, etc. You may also refer to the holiday pictures below for inspiration.

From: anexample@email.com

To: anotherexample@email.com

Object: Holidays

Message:



1.2. Guidelines e-mail about a party (English)

Write an e-mail in English (15 to 25 lines) to a friend to tell him/her about the party you attended last night.

Describe what you liked most and speak about some funny and crazy things that happened last night. Also, tell him/her about about things that happened and that you did not like at all.

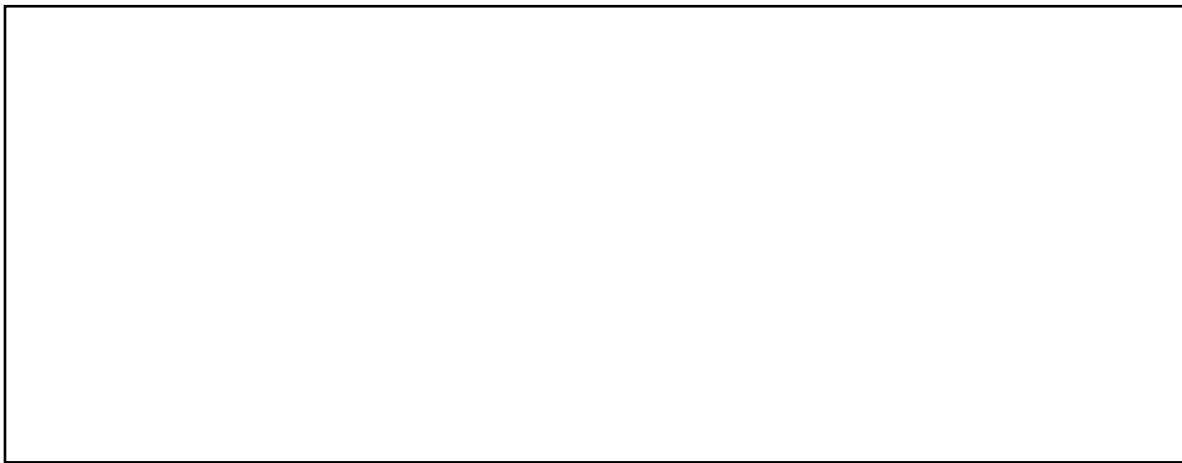
You may want to talk about the music, the atmosphere, the guests, the food, the presents, the performances, etc. You may also refer to the pictures below for inspiration.

From: anexample@email.com

To: anotherexample@email.com

Object: Party

Message:



1.3. Guidelines e-mail about the holidays (Dutch)

Schrijf een e-mail in het Nederlands (15 à 25 regels) aan een vriend(in) en vertel hem of haar over je vakantie.

Beschrijf wat je het allerleukst, het mooiste en het grappigst vond. Vertel ook wat je juist helemaal niet leuk vond.

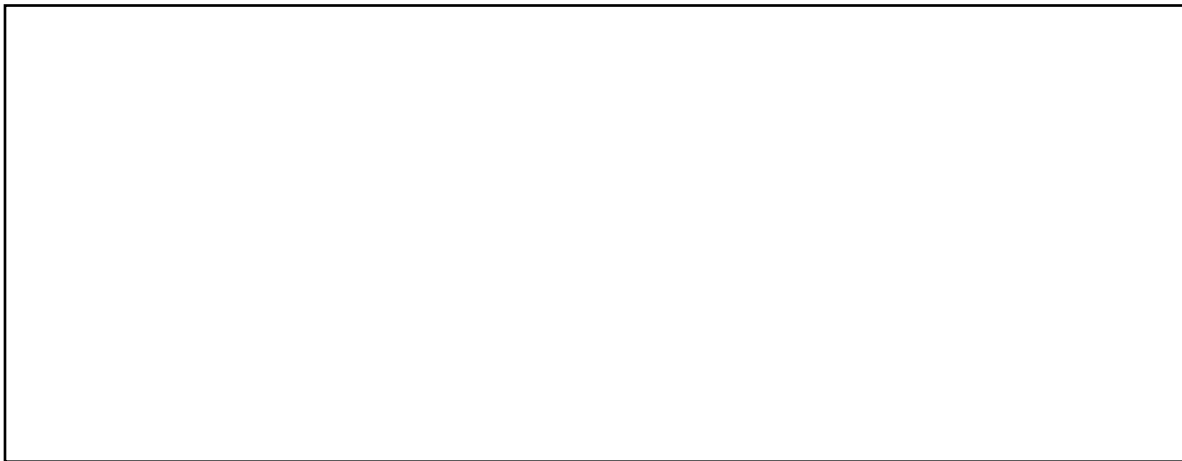
Als je niet zoveel inspiratie hebt, kun je bijvoorbeeld schrijven over de mensen die je ontmoet hebt, het weer, problemen die je had, de sfeer in het algemeen, het eten, etc. Je kan je eventueel ook baseren op de vakantiefoto's hieronder.

Van: eenvoorbeeld@email.com

Voor: eenandervoorbeeld@email.com

Onderwerp: Vakantie

Bericht:



1.4. Guidelines e-mail about a party (Dutch)

Schrijf een e-mail in het Nederlands (15 à 25 regels) aan een vriend(in) en verel hem of haar over een feestje waar je gisteren heen bent geweest.

Beschrijf wat je het allerleukst, het grappigst en het vreemdst vond. Vertel ook wat je juist helemaal niet leuk vond.

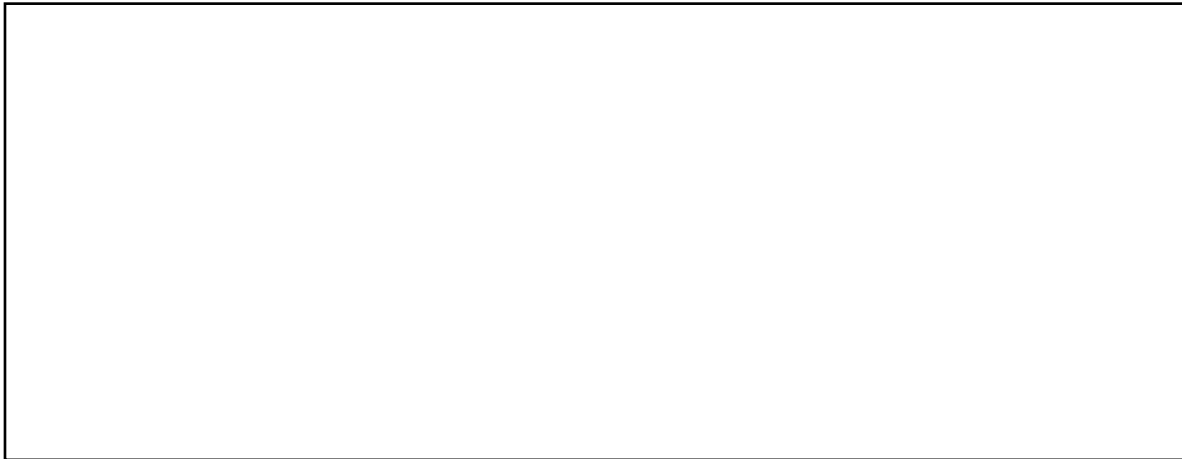
Als je niet zoveel inspiratie hebt, kun je bijvoorbeeld schrijven over de muziek, de sfeer, de gasten, het eten, de cadeaus, de acts/performances, etc. Je kan je eventueel ook baseren op de foto's hieronder.

Van: eenvoorbeeld@email.com

Voor: eenandervoorbeeld@email.com

Onderwerp: Feestje

Bericht:



2. Oral task (English/Dutch)

2.1. Guidelines exercise about the holidays (English/Dutch)

Tu vas participer à un petit exercice de quelques minutes durant lequel tu vas travailler par groupe de deux. Tu vas recevoir une carte avec des images qui sont différentes de celles de ton partenaire. A aucun moment tu ne pourras observer la carte de ton interlocuteur.

Contexte : *La carte est une invitation pour 2 personnes à une fête. Ton ami(e) en a reçu une aussi, mais ce n'est pas la même. Imagine que vous vous êtes donné rendez-vous pour en discuter et décider à quelle fête vous irez.*

Tâche : *Commence par poser des questions à ton partenaire afin d'en savoir plus sur la fête à laquelle il/elle est invité(e). Ensuite, essaye de le/la convaincre d'aller à la fête décrite sur ta carte. Vous devez arriver à un compromis.*

Durant cet exercice, tu utiliseras l'anglais/le néerlandais. Veille à répondre clairement aux questions de ton interlocuteur et à lui demander le plus d'informations possible. Si tu ne comprends pas ses questions, n'hésite pas à lui faire répéter.

Invitation pour 2 personnes

Viens faire la fête avec nous !



2.2. Guidelines exercise about a party (English/Dutch)

Tu vas participer à un petit exercice de quelques minutes durant lequel tu vas travailler par groupe de deux. Tu vas recevoir une carte avec des images qui sont différentes de celles de ton partenaire. A aucun moment tu ne pourras observer la carte de ton interlocuteur.

Contexte : *La carte est une invitation pour 2 personnes à une fête. Ton ami(e) en a reçu une aussi, mais ce n'est pas la même. Imagine que vous vous êtes donné rendez-vous pour en discuter et décider à quelle fête vous irez.*

Tâche : *Commence par poser des questions à ton partenaire afin d'en savoir plus sur la fête à laquelle il/elle est invité(e). Ensuite, essaye de le/la convaincre d'aller à la fête décrite sur ta carte. Vous devez arriver à un compromis.*

Durant cet exercice, tu utiliseras l'anglais/le néerlandais. Veille à répondre clairement aux questions de ton interlocuteur et à lui demander le plus d'informations possible. Si tu ne comprends pas ses questions, n'hésite pas à lui faire répéter.

*Invitation pour 2 personnes
Viens faire la fête avec nous !*



3. Multiple-choice tests verb-noun collocations

3.1. English

1. Conduct one's rights
Exercise one's rights
Employ one's rights
2. Win a purpose
Earn a purpose
Serve a purpose
3. Shed tears
Fell tears
Raise tears
4. Make damage
Do damage
Run damage
5. Perform a speech
Give a speech
Hold a speech
6. Draw a conclusion
Pull a conclusion
Make a conclusion
7. Seize an opportunity
Catch an opportunity
Pull an opportunity
8. Lay a visit
Pay a visit
Do a visit
9. Turn out a fire
Set out a fire
Put out a fire
10. Press charges
Push charges
Run charges
11. Drop count
Lose count
Pass count

12. Hold discussions
Set discussions
Do discussions
13. Keep one's breath
Hold one's breath
House one's breath
14. Gain friends
Create friends
Make friends
15. Take harm
Do harm
Make harm
16. Last a secret
Hold a secret
Keep a secret
17. Give birth
Lay birth
Bring birth
18. Conduct a crime
Comply a crime
Commit a crime

3.2. Dutch

1. De mensen achter mij in de bioscoop zaten de hele tijd commentaar te ... op de acteurs. Heel irritant!
a. bezorgen b. leveren c. praten
2. Bij de brand is gelukkig niemand gewond ...
a. geraakt b. geworden c. gekomen
3. De leerlingen vinden het niet leuk dat de leraar zo veel nadruk ... op grammatica en hen niet laat praten.
a. zegt b. zet c. legt
4. “Lekker in de zon zitten vind ik heerlijk!” “Je ... wel risico op kanker, wist je dat?”
a. hebt b. loopt c. krijgt
5. “Ik heb je gisteren gebeld. Was je niet thuis?” “Jawel, maar ik was zo moe dat ik de telefoon hoorde, maar niet heb ...”
a. opgehaald b. opgenomen c. beantwoord
6. Ik heb slechte punten gekregen. De volgende keer moet ik proberen minder fouten te ...
a. zetten b. maken c. doen
7. Die man zegt elke keer iets anders. Het is normaal dat de politie zijn verklaring in twijfel ...
a. trekt b. duwt c. schopt
8. Als je naar Amsterdam gaat, kun je een kijkje ... in het huis waar Anne Frank heeft gewoond.
a. zien b. doen c. nemen
9. Dit is eigenlijk geheim, je moet dus je mond ...
a. houden b. doen c. maken
10. Wat ik voor jou voel, is moeilijk onder woorden te ...
a. brengen b. zeggen c. maken
11. We ... om zeven uur aan tafel.
a. staan b. zetten c. gaan

12. Piet is heel blij: hij heeft een contract met Inter Milan ..., zijn favoriete club.
a. gedicht b. gesloten c. gemaakt
13. “Hoe vond je de show?” - “Niet zo leuk. Je kon goed zien dat alles in scène was ...”
a. gemaakt b. gedaan c. gezet
14. Wist je dat de pers het schandaal rond de jonge prins en zijn vriendin openbaar ...?
a. maakte b. hing c. legde
15. Eric heeft heel vaak hoofdpijn. Ik ... me zorgen: gaat het wel goed met hem?
a. doe b. maak c. raak
16. “Je bent niet erg enthousiast over je nieuwe camera” - “Nee, hij ... niet aan mijn verwachtingen”
a. denkt b. beantwoordt c. dient
17. Zo kan ik me niet concentreren. Dat lawaai buiten ... op mijn zenuwen.
a. werkt b. drukt c. loopt
18. Ik weet niet precies waarom hij weg moet, maar ik geloof dat dat verband ... met zijn ouders.
a. doet b. hangt c. houdt
19. Het ... geen verschil of je je boodschappen bij Delhaize doet of bij Carrefour: beide winkels zijn even duur.
a. heeft b. raakt c. maakt
20. Ik vind dat de stad maatregelen moet ... tegen het vandalisme.
a. nemen b. maken c. stellen
21. Op een camping in Frankrijk is een dubbele moord ...
a. gepleegd b. gemaakt c. gedaan
22. We moeten geld besparen, dus we moeten prioriteiten ...
a. maken b. geven c. stellen

4. Text fragments illustrating the categorization of phraseological units

CLIL ENGLISH

“Dear Matthew

How are you [*SPEECH ACT FORMULAE*]? I just **have** a **party** [*LEXICAL COLLOCATION – VERB + NOUN*] **last night** [*LEXICAL COLLOCATION – ADJ + NOUN*] and it was great. **As you probably know** [*SPEECH ACT FORMULAE*], yesterday it was my **birthday** [*COMPOUND*]. First we **went to the cinema** [*GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)*] to watch the new james bond. **I thought that** [*ATTITUDINAL FORMULAE*] it was good but less than the previous one, Skifall.

Than we **came back home** [*GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)*], we **put** some **good music** **[on]** [*LEXICAL COLLOCATIONS – VERB + NOUN and ADJ + NOUN*]. It was **really cool** [*LEXICAL COLLOCATION – ADV + ADJ*], I invited all my friend. I **really sad** [*LEXICAL COLLOCATION – ADV + ADJ*] that you couldn t **come with** [*GRAMMATICAL COLLOCATION – VERB + PREP*] us. We ate pizzas.

I also **receave** a lot of **presents** [*LEXICAL COLLOCATION – VERB + NOUN*] my **grandma** [*COMPOUND*] bought me a new iphone my parents **video games** [*COMPOUND*] and my friend **gave** me **either money** [*LEXICAL COLLOCATION – VERB + NOUN*] **or** [*COMPLEX CONJUNCTION*] cd s **and things like that**. [*ATTITUDINAL FORMULAE*]

My dog, tina, **ate** my **birthday cake**. [*LEXICAL COLLOCATION (VERB + NOUN) + COMPOUND*]

You miss me [*SPEECH ACT FORMULAE*], **i hope to see you soon**. [*SPEECH ACT FORMULAE*]

Bye”

NON-CLIL ENGLISH

“Hello!

I **went to** [GRAMMATICAL COLLOCATION – VERB + PREP] the USA **by plan** [GRAMMATICAL COLLOCATION – PREP + NOUN] with my mother, my father and my **best friend** [COMPOUND]. We **stayed in a hotel** [LEXICAL COLLOCATION (VERB + NOUN) + GRAMMATICAL COLLOCATION (PREP + NOUN)] near the beach **for a week** [GRAMMATICAL COLLOCATION – PREP + NOUN]. We **went to** [GRAMMATICAL COLLOCATION – VERB + PREP] California **in fact** [ATTITUDINAL FORMULAE]! The **hotel** was **really beautiful** [LEXICAL COLLOCATIONS – ADJ + NOUN and ADV + ADJ], there was a **swimmingpool** [COMPOUND]. The first day we **had a walk** [LEXICAL COLLOCATION – VERB + NOUN] **in the city** [GRAMMATICAL COLLOCATION – PREP + NOUN] and then we **played volleyball** [LEXICAL COLLOCATION (VERB + NOUN) + COMPOUND] **at the beach** [GRAMMATICAL COLLOCATION – PREP + NOUN]. The second day my parents **stayed at the hotel** [LEXICAL COLLOCATION (VERB + NOUN) + GRAMMATICAL COLLOCATION (PREP + NOUN)] and me and my **best friend** [COMPOUND] had just enjoy the sun **at [the] beach** [GRAMMATICAL COLLOCATION – PREP + NOUN], we met four **teenagers** [COMPOUND]. We **played volleyball** [LEXICAL COLLOCATION (VERB + NOUN) + COMPOUND] together (yaeh again lol) then they invited us to **go in a bar** [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)] for a drink, we accepted. While I was **talking with** [GRAMMATICAL COLLOCATION – VERB + PREP] Mathy, my **best friends** [COMPOUND] (Ines) was **talking with** [GRAMMATICAL COLLOCATION – VERB + PREP] Jake, I was thinking that she **had a crush on** [IDIOM] him she was smiling it was **so cute** [LEXICAL COLLOCATION – ADV + ADJ]! **In the bar** [GRAMMATICAL COLLOCATION – PREP + NOUN] **another** [COMPOUND] boy was **waiting for** [GRAMMATICAL COLLOCATION – VERB + PREP] us, it was Bethany's **boyfriend** [COMPOUND]. We **had a drink** [LEXICAL COLLOCATION – VERB + NOUN], I took a mojito and Ines a tequila. I was drunk so I don't remember **everything** [COMPOUND], but **I know** [ATTITUDINAL FORMULAE] I **danced with** [GRAMMATICAL COLLOCATION – VERB + PREP] Mathy. We **stayed with** [GRAMMATICAL COLLOCATION – VERB + PREP] them and when I **had to go** [GRAMMATICAL COLLOCATION – VERB + TO + INF] I was **so sad** [LEXICAL COLLOCATION – ADV + ADJ].”

CLIL DUTCH

“Dag Lies,

ik weet dat [ATTITUDINAL FORMULAE] ik geen emails sinds 4 weken heb gesturen⁶³ [LEXICAL COLLOCATION – VERB + NOUN] maar ik zou graag [ATTITUDINAL FORMULAE] [over] mijn laaste feest willen vertellen [GRAMMATICAL COLLOCATION – VERB + PREP]. Ik ben aan⁶⁴ de feest van Danny gegaan [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)] en ik heb veel nieuwe ideeën [LEXICAL COLLOCATION – VERB + NOUN] voor jouw verjaardage feest⁶⁵ [COMPOUND] die zal binnen 3 weken [GRAMMATICAL COLLOCATION – PREP + NOUN] zijn⁶⁶ [GRAMMATICAL COLLOCATION – VERB + NOUN]!

Ten eerste [LINKING ADVERBIAL] was het muziek heel leuk [LEXICAL COLLOCATIONS – ADJ + NOUN and ADV + ADJ]! Ik heb heel veel gedanst met [GRAMMATICAL COLLOCATION – VERB + PREP] iedereen. Het herinert me aan [GRAMMATICAL COLLOCATION – VERB + PREP] de Carnaval feest⁶⁷ [COMPOUND] van vorige jaar [LEXICAL COLLOCATION – ADJ + NOUN].

Bovendien, was het eten zo lekker [LEXICAL COLLOCATIONS – ADJ + NOUN] dat [COMPLEX CONJUNCTION] ik denk dat [ATTITUDINAL FORMULAE] ik 2 kilos heb genomen [LEXICAL COLLOCATION – VERB + NOUN]!! Daarentegen had de ouders van Danny geen drankjes gekopen⁶⁸ [LEXICAL COLLOCATION – VERB + NOUN].

Verder, heeft Danny's superleuke [COMPOUND] cadeau's gekregen [LEXICAL COLLOCATION – VERB + NOUN]. Hij was heel gelukkig [LEXICAL COLLOCATION – ADV + ADJ]. Ik had een cd-speler [COMPOUND] gekocht. Mijn ouders waren niet tevreden dat ik mijn geld besteed aan [LEXICAL COLLOCATION (VERB + NOUN) + GRAMMATICAL COLLOCATION (VERB + PREP)] een vriend. Nochtans heb ik gezegd dat ik genoeg zakgeld had [COMPOUND + LEXICAL COLLOCATION (VERB + NOUN)] om een cadeau voor mijn beste vriend [LEXICAL COLLOCATION – ADJ + NOUN] te [COMPLEX CONJUNCTION] kopen.

⁶³ gestuurd

⁶⁴ naar (...) gegaan

⁶⁵ verjaardagsfeest

⁶⁶ plaatsvinden

⁶⁷ carnavalsfeest

⁶⁸ gekocht

Als⁶⁹ conclusie [LINKING ADVERBIAL] besef ik dat [ATTITUDINAL FORMULAE] we help zullen nodig help⁷⁰ [LEXICAL COLLOCATION – VERB + NOUN] om een feest zo aangename als [COMPLEX CONJUNCTION] die van danny te [COMPLEX CONJUNCTION] organiseren [GRAMMATICAL COLLOCATION – VERB + NOUN]. Ik zou graag [ATTITUDINAL FORMULAE] je zien binnen de korste keren [IDIOOM] om alles te [COMPLEX CONJUNCTION] planifiseren voor jouw feest!

Kusje

G raldine”

NON-CLIL Dutch

« best Sara,

Hoe gaat het [SPEECH ACT FORMULAE]? Hoe was je vakantie [SPEECH ACT FORMULAE]?

Ik ben naar frankrijk gegaan [GRAMMATICAL COLLOCATION]. Dat was heel leuk [LEXICAL COLLOCATION – ADV + ADJ].

[Op] Dinsdag [GRAMMATICAL COLLOCATION – PREP + NOUN], ben ik met mijn vader en mijn moeder naar luchthaven [COMPOUND] gegaan [GRAMMATICAL COLLOCATION – VERB + PREP].

We zijn in het mooie hotel [GRAMMATICAL COLLOCATION (PREP + NOUN) + LEXICAL COLLOCATION (ADJ + NOUN)] verbleven [GRAMMATICAL COLLOCATION – VERB + NOUN]. Ik ben naar het strand gegaan [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)] en ik heb veel fritjes gegeten. Dat was heel lekker [LEXICAL COLLOCATION – ADV + ADJ].

Voor mijn verjaardag, zijn we naar het restaurant gegaan [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)].

Hier heb ik een mooi jongeren⁷¹ [LEXICAL COLLOCATION – ADJ + NOUN] ontmoeten.

⁶⁹ ter conclusie

⁷⁰ hulp nodig hebben

⁷¹ jongen

Jack woonde van⁷² [GRAMMATICAL COLLOCATION – VERB + PREP] het hotel waar ik verbleef [LEXICAL COLLOCATION – VERB + NOUN]. Hij had groen ogen [GRAMMATICAL COLLOCATION (VERB + NOUN) + LEXICAL COLLOCATION (ADJ + NOUN)].

We zijn naar de bioscoop gegaan [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)] en we hebben een fantastisch film gezien [LEXICAL COLLOCATIONS – ADJ + NOUN and VERB + NOUN]. Dat was tof.

We zijn samen naar het zwembad ook gegaan [GRAMMATICAL COLLOCATION (VERB + PREP) + LEXICAL COLLOCATION (VERB + NOUN)].

[Op] Een andere dag [GRAMMATICAL COLLOCATION – PREP + NOUN] heb hij gezegen dat hij van me hield [GRAMMATICAL COLLOCATION – VERB + PREP].

Mijn vakantie was tof [LEXICAL COLLOCATION – ADJ + NOUN] maar nu ik heb veel huiswerk [LEXICAL COLLOCATION (VERB + NOUN) + COMPOUND].

Ik ben boos.

Mail me vlug [SPEECH ACT FORMULAE]

Marie-Grâce.”

⁷² woonde in

5. Distribution of the phraseological units produced across individual learners

Number of learners using word combinations per sub-category (English)

	CLIL (/90)	NON-CLIL (/90)
Lexical collocations	90 (100%)	88 (98%)
Compounds	89 (99%)	88 (98%)
Grammatical collocations	90 (100%)	90 (100%)
Phrasal verbs	28 (31%)	6 (7%)
Complex prepositions	22 (24%)	9 (10%)
Complex conjunctions	19 (21%)	5 (5.5%)
Linking adverbials	21 (23%)	22 (24%)
Speech act formulae	69 (77%)	68 (75.5%)
Attitudinal formulae	37 (41%)	28 (31%)

Number of learners using word combinations per sub-category (Dutch)

	CLIL (/132)	NON-CLIL (/100)
Lexical collocations	132 (100%)	99 (99%)
Compounds	117 (89%)	79 (79%)
Grammatical collocations	131 (99%)	96 (96%)
Phrasal verbs	106 (80%)	45 (45%)
Complex prepositions	17 (13%)	4 (4%)
Complex conjunctions	102 (77%)	50 (50%)
Linking adverbials	31 (23%)	12 (12%)
Speech act formulae	100 (76%)	67 (67%)
Attitudinal formulae	91 (69%)	39 (39%)

6. Comparative analysis of global written complexity – statistics (Chapter 5)

	DUTCH CLIL / NON-CLIL	ENGLISH CLIL / NON-CLIL	FRENCH CLIL / NON-CLIL
Sentences per text	$U = 4234,0$ $z = -4,68, p < .05^*$ $r = -0,307$	$U = 3937,0$ $z = -0,32, p > .05$ $r = -0,024$	$U = 21647,5$ $z = -1,10, p > .05$ $r = -0,053$
Words per text	$U = 1909,0$ $z = -9,27, p < .05^*$ $r = -0,601$	$U = 2254,5$ $z = -5,14, p < .05^*$ $r = -0,383$	$U = 20838,5$ $z = -1,73, p > .05$ $r = -0,083$
Words per sentence	$U = 4257,5$ $z = -4,63, p < .05^*$ $r = -0,301$	$U = 2539,5$ $z = -4,32, p < .05^*$ $r = -0,322$	$U = 22350,5$ $z = -0,06, p > .05$ $r = -0,003$
Syllables/morphs per word	$U = 3889,0$ $z = -5,36, p < .05^*$ $r = -0,352$	$U = 3250,5$ $z = -2,28, p < .05^{**}$ $r = -0,170$	
Letters per word	$U = 5034,0$ $z = -3,09, p < .05^*$ $r = -0,203$	$U = 3896,0$ $z = -0,44, p > .05$ $r = -0,033$	$U = 20641,0$ $z = -1,88, p > .05$ $r = -0,091$
TTR content words	$U = 6124,5$ $z = -0,94, p > .05$ $r = -0,062$	$U = 3655,5$ $z = -1,129, p > .05$ $r = -0,084$	$U = 23031,5$ $z = -0,03, p > .05$ $r = -0,001$
TTR all words	$U = 5188,5$ $z = -2,79, p < .05^{**}$ $r = -0,183$	$U = 3762,0$ $z = -0,82, p > .05$ $r = -0,061$	$U = 22869,5$ $z = -0,15, p > .05$ $r = -0,007$
MTLD	$U = 3587,0$ $z = -5,95, p < .05^*$ $r = -0,391$	$U = 2380,0$ $z = -4,78, p < .05^*$ $r = -0,356$	

7. Correlations between predictor and outcome variables (Chapter 8)

Correlations between variables in Chapter 8 – English

		% of correct answers	SES	Informal contact with TL	Duration of TL instruction	Productive vocabulary knowledge	Receptive vocabulary knowledge	CLIL
% of correct answers	Cor.	1,000	0,165	0,366	0,297	0,423	0,533	0,439
	Coeff.			**	**	**	**	**
	Sig. (2- tailed)		0,072	0,000	0,000	0,000	0,000	0,000
	<i>n</i>	165	120	156	151	159	155	165
SES	Cor.	0,165	1,000	0,129	0,159	0,193	0,151	0,175
	Coeff.					*		
	Sig. (2- tailed)	0,072		0,171	0,095	0,036	0,105	0,056
	<i>n</i>	120	120	115	111	118	116	120
Informal contact with TL	Cor.	0,366	0,129	1,000	0,147	0,271	0,311	0,265
	Coeff.	**				**	**	**
	Sig. (2- tailed)	0,000	0,171		0,071	0,001	0,000	0,001
	<i>n</i>	156	115	156	151	150	147	156
Duration of TL instruction	Cor.	0,297	0,159	0,147	1,000	0,209	0,311	0,166
	Coeff.	**				*	**	*
	Sig. (2- tailed)	0,000	0,095	0,071		0,012	0,000	0,042
	<i>n</i>	151	111	151	151	145	147	151
Productive vocabulary knowledge	Cor.	0,423	0,193	0,271	0,209	1,000	0,439	0,336
	Coeff.	**	*	**	*		**	**
	Sig. (2- tailed)	0,000	0,036	0,001	0,012		0,000	0,000
	<i>n</i>	159	118	150	145	159	149	159
Receptive vocabulary knowledge	Cor.	0,533	0,151	0,311	0,311	0,439	1,000	0,550
	Coeff.	**		**	**	**		**
	Sig. (2- tailed)	0,000	0,105	0,000	0,000	0,000		0,000
	<i>n</i>	155	116	152	147	149	155	155
CLIL	Cor.	0,439	0,175	0,265	0,166	0,336	0,550	
	Coeff.	**		**	*	**	**	
	Sig. (2- tailed)	0,000	0,056	0,001	0,042	0,000	0,000	1,000
	<i>n</i>	165	120	156	151	159	155	165

Correlations between variables Chapter 8 – Dutch

		% of correct answers	SES	Duration of TL instruction	Informal contact with TL	Productive vocabulary knowledge	Receptive vocabulary knowledge	CLIL
% of correct answers	Cor.	1,000	0,444	0,331	0,254	0,580	0,717	0,602
	Coeff.		**	**	**	**	**	**
	Sig. (2- tailed)		0,000	0,000	0,000	0,000	0,000	0,000
	<i>n</i>	214	181	200	188	208	201	214
SES	Cor.	0,444	1,000	0,133	0,098	0,303	0,374	0,318
	Coeff.	**				**	**	**
	Sig. (2- tailed)	0,000		0,084	0,203	0,000	0,000	0,000
	<i>n</i>	181	181	170	172	177	175	158
Duration of TL instruction	Cor.	0,331	0,133	1,000	0,251	0,262	0,333	0,259
	Coeff.	**			**	**	**	**
	Sig. (2- tailed)	0,000	0,084		0,001	0,000	0,000	0,000
	<i>n</i>	200	170	200	179	195	190	200
Informal contact with TL	Cor.	0,254	0,098	0,251	1,000	0,275	0,313	0,206
	Coeff.	**		**		**	**	**
	Sig. (2- tailed)	0,000	0,203	0,001		0,000	0,000	0,005
	<i>n</i>	188	172	179	188	183	183	188
Productive voc knowledge	Cor.	0,580	0,303	0,262	0,275	1,000	0,524	0,466
	Coeff.	**	**	**	**		**	**
	Sig. (2- tailed)	0,000	0,000	0,000	0,000		0,000	0,000
	<i>n</i>	208	177	195	183	208	195	208
Receptive voc knowledge	Cor.	0,717	0,374	0,333	0,313	0,524	1,000	0,625
	Coeff.	**	**	**	**	**		**
	Sig. (2- tailed)	0,000	0,000	0,000	0,000	0,000		0,000
	<i>n</i>	201	175	190	183	195	201	201
CLIL	Cor.	0,602	0,278	0,259	0,206	0,466	0,625	1,000
	Coeff.	**	**	**	**	**	**	
	Sig. (2- tailed)	0,000	0,000	0,000	0,005	0,000	0,000	
	<i>n</i>	214	181	200	188	208	201	214