UNIVERSIDAD DE EXTREMADURA

PhD DISSERTATION

TOPIC-BASED TEACHING OF METAPHOR IN AN EFL SYLLABUS: A LONGITUDINAL STUDY OF ACHIEVEMENT AT B2 LEVEL

TESIS DOCTORAL

LA ENSEÑANZA DEL LENGUAJE METAFÓRICO DESDE UN ENFOQUE BASADO EN TEMAS EN LA ASIGNATURA DE INGLÉS COMO LENGUA EXTRANJERA: ESTUDIO LONGITUDINAL EN EL NIVEL B2

MARTA MARTÍN GILETE

PROGRAMA DE DOCTORADO EN LENGUAS Y CULTURAS

Conformidad de la directora y codirectora:

Dra. Ana María Piquer Píriz **Dra. Fiona MacArthur Purdon**

Esta tesis cuenta con la autorización de la directora y codirectora de la misma y de la Comisión Académica del programa. Dichas autorizaciones constan en el Servicio de la Escuela Internacional de Doctorado de la Universidad de Extremadura.

2024

Name and Academic Affiliation

Marta Martín Gilete

Department of English Philology, Faculty of Arts and Letters

University of Extremadura

Cáceres, Spain

Academic Identification

ORCID: https://orcid.org/0000-0001-6065-8304

ResearcherID: H-3670-2018

Scopus Author ID: 58019122100

Contact Information

mmgilete@unex.es

Doctoral Thesis Reference

Martín-Gilete, M. (2024). *Topic-based teaching of metaphor in an EFL syllabus: A longitudinal study of achievement at B2 level*. [Unpublished doctoral dissertation]. University of Extremadura. https://dehesa.unex.es/

To Luis and Eladia.

And for all those who share a passion for metaphors, from whom I draw endless inspiration.

ACKNOWLEDGEMENTS

Dissertations, though borne from one's own research and writing, are never truly solitary accomplishments. There are a number of people who I would like to thank for the support that they have provided over the last six years and whose encouragement has helped me complete this doctoral thesis.

First and foremost are my supervisors, Dr. Fiona MacArthur and Dr. Ana Mª Piquer-Píriz. Their vast knowledge and expertise have been instrumental in shaping this research project. I extend my most sincere thanks for their thoughtful advice, insightful comments, and continuous support throughout my PhD studies. Without their inspiration, this thesis would have remained merely an idea.

I am also grateful to Dr. Elena Semino for her welcome and supervision during my international research stay at the Department of Linguistics and English Language at Lancaster University (UK). Her expertise in metaphor research provided critical feedback that greatly enriched my perspective on my doctoral research project.

Warm thanks to the English Language and Applied Linguistics research team at the University of Extremadura, where research is imbued with a personal touch. My heartfelt thanks to Dr. Irene Castellano-Risco for her collaborative spirit, generosity, and assistance.

Additional thanks are extended to the members of the METCLIL research project team (FFI2017-86320-R). Collaborating on this nationally funded project, focusing on metaphor in academic talk, has been an enlightening experience. It has not only allowed me to develop a diverse set of methodological skills but also enriched my academic training through the expertise of both national and international colleagues.

I am deeply thankful to the teachers and students at the English language school who participated in the study. Their active involvement and unique contributions were invaluable, enriching the findings and my understanding of the subject. My sincerest thanks go to all for their indispensable role in this academic endeavor, particularly during the extraordinary challenges of the COVID-19 pandemic.

I am indebted to the members of RaAM and AELCO, whose wealth of inspiration and scholarly exchange in academic events and beyond has been profoundly motivating. Their formative feedback has both nurtured and propelled my academic ambitions. Many thanks for the enlightening discussions that have followed my presentations, especially to Susan Nacey for her critical feedback and contributions.

Lastly, but by no means least, I acknowledge my family and friends. Although physically distant during the demanding final phase of this piece of work, their support and spirit remained steadfast. A special thanks goes to my parents, Luis and Eladia, whose belief in me and endless encouragement have been the bedrock of my journey. Their love and sacrifices have shaped who I am and have been a constant source of strength and motivation.

To Juan, your companionship and faith in me have been the quiet haven in which this dissertation took shape. The solace found in our daily walks and your support have been the antidote to every moment of doubt. For all this, and more, I am deeply thankful.

Cáceres, Spain Marta Martín Gilete

January 2024

CONTENTS

PART TWO

♦ THE STUDY ♦

ABSTRACT

This doctoral thesis presents a longitudinal study analysing the impact of integrating cognitive-linguistic (CL) oriented methods into a topic-based EFL syllabus, aligned with the CEFR B2 level descriptors. It explores how CL-inspired activities enhance natural metaphor usage among L2 learners as their language skills develop, focusing on conventional metaphors. The quasi-experimental study involves 40 Spanish-speaking secondary school students preparing for B2 level, comparing the vocabulary growth of 20 students receiving explicit metaphor teaching (experimental group) with that of 20 students engaging in a standard communicative vocabulary approach (control group).

The results reveal that metaphor-mediated instruction enhances metaphor usage frequency and diversity in topic-based contexts, increasing learner confidence and vocabulary depth in both speech and writing. However, increased metaphor usage is not always directly tied to the teaching-learning process. Discourse nature and task conventions also play key roles, and may not necessarily indicate a deeper understanding of metaphors or the ability to deploy them in context.

This study highlights the challenges of CL-inspired instruction, which primarily focuses on the conceptual dimension of metaphor, in ensuring consistent accuracy in L2 production of linguistic metaphors. This suggests a need for a more holistic approach that fosters not only deeper cognitive engagement but also teaching lexico-grammatical patterns and pragmatic aspects. Using topic-based metaphors does not automatically lead to higher performance in speech and writing, particularly with unconventional metaphors, as per standard assessment criteria. This suggests a potential mismatch between the importance of metaphor in L2 learning and its assessment in real-world testing environments.

Key words: cognitive-linguistic inspired instruction, metaphor production, standard L2 assessment.

RESUMEN

Esta tesis doctoral analiza el impacto de integrar técnicas pedagógicas inspiradas en lingüística cognitiva (LC) en la asignatura de inglés como lengua extranjera desde un enfoque basado en temas y alineado con el nivel B2 del MCER. Explora cómo actividades inspiradas en LC potencian el uso natural de metáforas convencionales por parte de estudiantes que desarrollan sus habilidades lingüísticas en este nivel. Incluye un estudio cuasi-experimental con 40 estudiantes españoles de secundaria, comparando el crecimiento de vocabulario de 20 alumnos con enseñanza explícita de metáforas (grupo experimental) frente a 20 con un método comunicativo estándar (grupo de control).

Los resultados muestran que la instrucción en metáforas fomenta un uso frecuente y variado en contextos temáticos, mejorando la confianza y el conocimiento de vocabulario en expresión oral y escrita. Sin embargo, estos beneficios no siempre son consecuencia del proceso de enseñanza-aprendizaje. La naturaleza del discurso y las convenciones de las tareas también son clave, aunque no aseguran una comprensión profunda o habilidad para usar metáforas contextualmente.

El estudio resalta los desafíos de la enseñanza inspirada en LC para asegurar una producción precisa de metáforas lingüísticas en L2. Se infiere la necesidad de un enfoque integral que combine la dimensión conceptual con la enseñanza de aspectos léxico-gramaticales y pragmáticos. El uso de metáforas temáticas no garantiza mayor rendimiento en expresión oral y escrita, especialmente aquellas no convencionales, evaluadas según criterios estándar. Esto sugiere una discrepancia entre la importancia de la metáfora en el aprendizaje de L2 y su evaluación en exámenes oficiales.

Palabras clave: instrucción basada en lingüística cognitiva, producción de metáfora, evaluación estándar de L2.

LIST OF TABLES AND FIGURES

Tables

Figures

LIST OF ABBREVIATIONS

ACL – Applied Cognitive Linguistics

AF – Absolute Frequency

BNC – British National Corpus

CAMD – Cambridge Dictionary

CEFR – Common European Framework of Reference for Languages

CL – Cognitive Linguistics

CLAWS – Constituent Likelihood Automatic Word-Tagging System

CLIL – Content and Language Integrated Learning

CM – Conceptual Metaphor

CMA – Conceptual Metaphor Approach

CMT – Conceptual Metaphor Theory

COVID-19 – Coronavirus Disease 2019

DFMA – Discarded for Metaphor Analysis

DMT – Deliberate Metaphor Theory

EFL – English as a Foreign Language

ELF – English as a Lingua Franca

EMI – English as a Medium of Instruction

ESOL – English for Speakers of Other Languages

EuroCoAT – The European Corpus of Academic Talk

IELTS – International English Language Testing System

L1 – First Language

L2 – Second Language

LM – Longman Dictionary of Contemporary English

Ls. – Listening

LU – Lexical Unit

MC – Metaphorical Competence

METCLIL – Corpus of Metaphor in Academic Talk

MFlag – Metaphor Flag

MIP – Metaphor Identification Procedure

MIPVU – Metaphor Identification Procedure Vrije Universiteit

MLP – Metaphor Language Play

MM – Macmillan English Dictionary

MRW – Metaphor-Related Word

non-MRW – non-Metaphor-Related Word

OED – Oxford English Dictionary

OOPT – Oxford Online Placement Test

PoS – Part of Speech

PP – Personification

PPP – Presentation-Practice-Production

Rd. – Reading

RF – Relative Frequency

RI – Relative Increase

RQ – Research Question

Sp. – Speaking

SSL – Spanish as an L2

TEFL – Teaching English as a Foreign Language

TESOL – Teaching English to Speakers of Other Languages

UoE – Use of English

VOICE – Vienna-Oxford International Corpus of English

WiDLii – When in Doubt, Leave it in

Wr. – Writing

PART ONE

♦ FOUNDATIONS ♦

CHAPTER 1

INTRODUCTION

It is now commonplace to state that metaphor — "the phenomenon whereby we talk, and, potentially, think about something in terms of something else" (Semino, 2008, p. 11) — is a fundamental aspect of language use and cognition. Within the theoretical paradigm of Cognitive Linguistics (CL), far from being considered mere rhetorical embellishments in discourse, metaphors are essential cognitive mechanisms inherent to human communication. This perspective highlights that metaphors are pervasive in naturally occurring language use from early infancy (Cameron, 2003; Özçalışkan, 2011; Pouscoulous & Tomasello, 2020), whether one is providing directions (e.g., *take* the second right) or delving into a philosophical discussion (e.g., *take* the dilemma by one its horns).[1] Metaphoricity involves using words with non-literal meaning, conveying conventional meaning into derived meaning, which creates cultural, social, and also psychological realities for language users.

Lakoff and Johnson's (1980, 1999) *Conceptual Metaphor Theory* (CMT) holds that metaphors play a crucial role in shaping our perception of reality by operating primarily at the level of thought through *conceptual metaphors* (CM). CMs help us understand and experience more abstract or complex concepts by mapping them onto more physical, familiar experiences (for further details, see Gibbs, 1994, 2008, 2014; Kövecses, 2002, 2005; Lakoff,1993).

For instance, our concept of "mind" is partly structured by our knowledge of a "container", which underlies the CM of MIND IS A CONTAINER to think and/or communicate about ideas, thoughts, and emotions. In English, we view the "mind" (vehicle/source domain) in terms of a "container" (topic/target domain), with thoughts being stored there or going in and out. Underlying CMs are reflected in language by *linguistic metaphors*, which are often conventional and codified in dictionaries. Hence, vocabulary related to "containers" is effective

[1] Italics are used to signal the items being focused on.

for discussing this mental content. Examples span various word classes, including nouns exemplified by “brain *dump*”; verbs as in “*clear* your mind”; adjectives represented by “have an *open* mind”; adverbs in the phrase “be at the *back* of your mind”; and prepositions as seen in “put a goal *in* your mind”.

The concept of metaphor can be approached from various perspectives in discourse analysis, including *language*, *thought*, and *communication* (Steen, 2008, 2011a). This PhD dissertation focuses on *metaphor in language*, which refers to the metaphorically used words found in discourse, rather than exploring how metaphor is processed in the mind or its contentious deliberate use for communicative functions (Deignan, 2011; Gibbs, 2011; Krennmayr, 2011; Reijnierse, 2017; Steen, 2011b, 2023). However, that debate is beyond the scope of this thesis, which aims to maintain an applied linguistic focus.

Metaphor identification procedures, such as the Metaphor Identification Procedure (MIP; Pragglejaz Group, 2007) and the Metaphor Identification Procedure Vrije Universiteit (MIPVU; Steen et al., 2010), have established that metaphor is ubiquitous in both first language (L1) and second/foreign (L2) language use. Metaphor not only accounts for a substantial proportion of English language use (in relation to the total of words [tokens] analysed) across different genres — covering 17.50% in academic texts, 15.30% in news, 10.80% in fiction, and 6.80% in conversation (Steen et al., 2010, pp. 194–298) — but is also pervasive in learner English. For example, Nacey (2013) reported an overall metaphor density of 15.50% in written essays produced by Norwegian undergraduates.

The prominence of metaphor in learner discourse highlights its important role in the context of Teaching English as a Foreign Language (TEFL), for both L2 learners and teachers (Low, 1988, 2008, 2020). Such centrality presents challenges for L2 learners, requiring mastery in the effective use of metaphorical language in their speech and writing. This ability is crucial, not just in isolation, but in the broader context of standard L2 competence assessment (Littlemore & Low, 2006a, 2006b). In response, instructors should prepare learners to correctly use metaphors in conventional language patterns that are appropriate to

the lexico-grammatical structures of the L2, aiding them in reaching their target proficiency levels (cf. O'Reilly & Yan, 2023).

1.1. METAPHOR IN LEARNER DISCOURSE: SOME IMPORTANT CONSIDERATIONS

Metaphorical competence (MC) can be broadly defined as the ability that "involves the comprehension, production, awareness, and retention of metaphor in speaking, writing, reading and/or listening" (O'Reilly & Marsden, 2021, p. 26). In recent decades, the importance of MC in L2 learning, particularly with the goal of improving language proficiency, has gained recognition from various perspectives (e.g., Danesi, 1995; Littlemore, 2001a; Low, 1988, 2008, 2020; Castellano-Risco & Piquer-Píriz, 2020; O'Reilly & Marsden, 2021, 2023).[2] However, its definition and measurement remain problematic, especially in relation to L2 metaphor production (Littlemore & Low, 2006a). The term "competence" itself carries multiple connotations, open to interpretations ranging from social to psychological dimensions. As Low (2020) notes, "it may emphasise either the process or the outcome, and can relate either directly to an action or result, or more indirectly, serving as a foundational skill" (p. 47).

Prior research has shed light on the benefits of fostering MC among L2 learners of English by primarily investigating comprehension and/or production of metaphor in relation to vocabulary acquisition and retention. Most attention has been devoted to the study of L2 learners' understanding of metaphor in English across both oral and written discourse (Golden, 2010; Littlemore, 2001b, 2004; Piquer-Píriz, 2008a, 2010), as well as research into the interplay between metaphor comprehension and its usage (Charteris-Black, 2002; MacArthur & Littlemore, 2011; O'Reilly & Marsden, 2023).

Studies focusing on metaphor production alone have been relatively limited, frequently analysing L2 writing and providing insights into learners' abilities at specific moments in their learning process (Alejo-González, 2010; Kathpalia & Carmel, 2011; MacArthur, 2010; Nacey, 2013). Research on oral communication in this area has been even more rare, with a particular

[2] See Hoang (2014) and Nacey (2017) for a more comprehensive review.

focus on English as a Lingua Franca (ELF) academic settings (see MacArthur & Alejo-González, 2024, for a comprehensive review).

In recent years, there has been a growing interest in exploring how L2 metaphor production (i.e., distribution and function) develops as proficiency improves, particularly in writing (Cuberos et al., 2019; Hoang & Boers, 2018; Littlemore et al., 2014; Nacey, 2020, 2022; Turner, 2014). These studies have revealed that, as learners progress in their L2 learning, both metaphor density and the frequency of metaphor clusters (i.e., bounded peaks of metaphor density) in their written production increase. This enhanced use of metaphors is not only associated with improved proficiency across all levels of the Common European Framework of Reference for Languages (CEFR)[3] (Littlemore et al., 2014), but it is also linked to higher age and deeper engagement in complex topics across various educational levels and different L1 backgrounds (Cuberos et al., 2019; Hoang & Low, 2018; Nacey, 2013, 2020, 2022; Turner, 2014).

However, distinct patterns have been observed in the type of metaphorical language used by L2 learners concerning open-class (lexical items) vs. closed-class metaphors (grammatical function), and the purpose that metaphor clusters serve, playing discoursal, topical or even unmotivated roles in this intensive metaphor use, especially in learners' written discourse at the upper-intermediate level (see Nacey, 2022). Nacey (2022) suggests that "a preponderance of seemingly unmotivated clusters might be a hallmark of novice texts […], where pupils have not yet mastered their writing skills in either their L1 or L2" (p. 295). Furthermore, such tendencies in novice texts could also stem from the challenges faced by L2 learners, particularly in terms of their limited lexicon and a developing grasp of effectively using metaphors in English (MacArthur, 2010; see below for further elaboration on this idea).

Previous studies have shown that the metaphor density of open-class metaphors (i.e., nouns, verbs, adjectives, and adverbs) increases more rapidly than that of closed-class metaphorical items such as prepositions, while also highlighting variations in the proficiency

[3] Chapter 2 includes more detailed information about the CEFR.

levels at which the use of the former overtakes the use of the latter. Littlemore et al. (2014) found that at B2 level and beyond, open-class metaphors were significantly more frequently used than closed-class metaphors (primarily prepositions), suggesting "a qualitative change in the type of metaphor that the learners need to use", especially at B2 level (p. 128).

In contrast, Nacey's (2020) study revealed that the frequency[4] of open-class metaphors surpassed that of closed-class metaphors far below B2 proficiency level (in year 6),[5] as learners progress through their school years from grade 5 to 13. When examining the individual level, Nacey (2022) demonstrated that as L2 learners progressed through various school grade levels (from years 8 to 11) and their language proficiency improved, there were no significant differences in the rates of increase for open-class and closed-class metaphors over time. The varied findings across these earlier studies raise questions about the behaviour of open-class and closed-class metaphors in learner English discourse, particularly at B2 level.

Recent research into MC has evidenced what instructors can expect L2 learners to be able to do at various proficiency levels, mainly regarding their written expression (e.g., Cuberos et al., 2019; Hoang & Boers, 2018; Littlemore et al., 2014; Nacey, 2020, 2022; O'Reilly & Marsden, 2021). Yet, specific observations related to L2 learners at the B2 level stand out. O'Reilly & Marsden (2023) have observed that at this learning stage, L2 learners "recognise […] metaphors with a moderate amount of accuracy, although they are more likely to struggle when producing metaphor within these types of item contexts" (p. 31).

Littlemore et al. (2014) highlight the need for providing support in metaphor usage, especially at B2 level, as these learners often find themselves "at an experimental stage of language development during which their task demands require them to experiment with new ways of using metaphorical language" (Littlemore et al., 2014, p. 128). Therefore, they tend to make more errors when using metaphors and there is more evidence of L1 influence (p. 140).

[4] By frequency, I refer to the percentage of words (tokens) used metaphorically within the target text. This measurement is then used to determine the "density" of metaphor in specific discourse contexts.
[5] According to Nacey (2020, p. 185), successful completion of year 11 English is generally recognised as the equivalent of the CEFR B2 level, as endorsed by the Norwegian Ministry and Research.

Nacey (2022) builds on this, suggesting that such learners could "have not yet mastered their writing skills in either their L1 or L2" (p. 295). These challenges might be attributed to L2 learners' limited vocabulary and insufficient understanding of how to use metaphors effectively in English (MacArthur, 2010). Littlemore et al. (2014) further note that L2 learners at B2 level "lack the support to do so convincingly, or the confidence to do so without falling back on their native language" (p. 128). Based on their findings, Littlemore et al. (2014) advocate for a teaching approach that places special emphasis on metaphor use, particularly at B2 or upper-intermediate level.

We know little about *where* L2 learners can find support to enhance their use of metaphors in English. Previous research has tackled this issue from various angles. Some studies have implemented CL-informed instruction in short-term studies, which were conducted outside the context of normal classroom activities (e.g., Saaty, 2016, 2020). Others have analysed the production of metaphors in learner discourse, particularly in written expression (see above discussion). Others have explored metaphor usage in L2 textbook discourse with a particular focus on metaphor-related activities (e.g., Alejo-González et al., 2010; Amaya-Chávez, 2010; Millar, 2023).

To the best of my knowledge, a side-by-side investigation of L2 metaphor production, both in speech and writing, in relation to the meaningful discourse learners encounter in mainstream materials is still to be undertaken. This gap is especially evident when considering long-term support of CL-oriented pedagogical practices within regular instructional activities.

Such support from English as a Foreign Language (EFL) instructors in enhancing metaphor usage could improve L2 learners' English proficiency, potentially increasing their chances of success in high-stakes English for speakers of other languages (ESOL) examinations. While previous research indicates a correlation between fluency in elicited MC (written comprehension and production) and overall language proficiency (O'Reilly & Marsden, 2023), the impact of metaphor usage on outcomes naturally produced in real-world spoken and written contexts such as standard L2 competence exams is yet to be investigated.

1.2. PURPOSE AND SCOPE OF RESEARCH

This doctoral thesis examines the long-term effectiveness of a topic-based approach to explicit metaphor instruction within an EFL syllabus, focusing on its impact on the frequency and variety of metaphors in L2 learners' spoken and written production. The research further explores whether these learning gains can lead to improved English language performance at the CEFR B2 level, as measured by the *B2 First for Schools Cambridge English* qualification.[6]

Considering the importance of topic similarity in preparing for standard L2 competence exams, this research integrates metaphor awareness into regular EFL instruction through CL-oriented distributed learning, termed here *metaphor-mediated instruction*. By exploiting the semantic potential of topic-based vocabulary from the goal-oriented textbook used in class, the CL-oriented instructional intervention seeks to enhance the natural metaphor usage of Spanish secondary school students as their English proficiency increases over time. It explores vocabulary growth, particularly in conventional metaphor usage, across various real-life oral and written communication contexts within high-stakes ESOL examination.

It is important to note that this PhD dissertation does not specifically investigate the learning and retention of metaphors as distinct vocabulary items pre-taught in the L2 classroom. Instead, the study takes a different approach by tracking the evolution of L2 learners' organic metaphor usage while preparing for the CEFR B2 level, in the context of metaphor-mediated instruction supported by textbook-based input.

1.3. THESIS OUTLINE

This PhD dissertation is structured into two main parts. Part 1 lays out the foundational background, with Chapter 2 specifically exploring the topic of metaphor in EFL. Section 2.1 introduces a CL approach to metaphor, focusing on the analysis of metaphor in language, especially learner discourse. In Section 2.2, the theoretical framework of the study is presented, delving into key considerations pertaining to the application of CL principles to L2

[6] See Chapter 3 for detailed information about the *B2 First for Schools Cambridge English* qualification.

instruction. Section 2.3 undertakes a critical review of previous research studies that are particularly relevant to the current investigation, exploring effective methods of fostering the understanding and production of metaphorical language by learners in instructed L2 settings. Section 2.4 addresses the challenges associated with integrating CL principles into EFL classrooms and identifies the research gap within pedagogically oriented metaphor studies, thereby justifying the undertaking of this doctoral thesis. Section 2.5 defines its objectives and research questions (RQs), establishing the framework for the current investigation.

Part 2 of this dissertation is devoted to the empirical study and is divided into four chapters. Chapter 3 describes the methodology employed, including the context and design of the longitudinal study (Section 3.1) and the implementation of the teaching method and the procedure followed in metaphor-mediated instruction (Section 3.2). This chapter also details the method used to gather and analyse the use of metaphor in the oral and written discourse of L2 learners, encompassing both the control and experimental groups before and after the CL-oriented instructional intervention (Section 3.3).

The results of the empirical study are reported in Chapter 4, which addresses each of the three RQs. The impact of metaphor-mediated instruction is examined in relation to the participants' use of metaphorical language (Section 4.1), their achievement at the B2 level (Section 4.2), and how their metaphor performance relates to their English language performance at the upper-intermediate level (Section 4.3).

Chapter 5 provides an analysis and discussion of the research findings. It aims to respond to the three RQs explored in this study across three key dimensions: L2 learners' metaphor performance (Section 5.1), their English language performance (Section 5.2), and the relationship between metaphorical language use and their L2 proficiency at B2 level, particularly regarding *Speaking* and *Writing* (Section 5.3).

Chapter 6 presents the conclusions drawn from the study. In Section 6.1, a summary of the key findings and contributions is provided. Section 6.2 delves into the implications and potential applications of the study to the teaching and assessment of metaphor in EFL contexts. Section 6.3 acknowledges some limitations of the study and suggests directions for

future research. In Section 6.4, concluding remarks offer a reflection on the overall significance and implications of the research conducted.

CHAPTER 2

BACKGROUND

2.1. OPERATIONALISING METAPHOR IN LEARNER DISCOURSE

This PhD dissertation delves into metaphor, focusing on its conceptualisation and use in everyday language by L2 speakers. As introduced in Chapter 1, metaphor extends beyond its traditional perception as merely an ornamental device for poets, writers, and orators. The CL approach challenges the notion of metaphor being confined to literary and formal speech events, emerging as a pervasive element in both language and thought. Metaphor plays a vital role across diverse contexts, being essential for human communication.

Prior to the late 1970s, metaphor was primarily associated with figures of speech and rhetoric leading to its disconnection from everyday language usage. The conventional perspective was challenged by the cognitive view, set in train by Michael Reddy in 1979, that recognised metaphor as omnipresent in everyday language and a critical component in cognitive processing (see Reddy, 2012, for more details).

This shift in understanding was further reinforced by the emergence of CMT, following the publication of Lakoff and Johnson's seminal work *Metaphors We Live By* in 1980 and the subsequent works on the topic. These developments called into question the rigid distinction between literal and figurative language. Within this framework, metaphors are not merely linguistic expressions, but conceptualisation of abstract entities according to experiential realised in language. Thus, literal language is typically understood as referring to tangible, physical experiences, while metaphorical language primarily deals with conceptualising abstractions or emotions. As a result, metaphor is seen to manifest in two dimensions: as *conceptual* metaphor, existing in thought before being translated into words, and as *linguistic* metaphor, the external expression of these thoughts in language.

2.1.1. Approaches to Analysing Metaphor

Since the 1980s, there have been three separate research traditions to analysing the concept of metaphor (see Steen, 2023). Drawing on the distinction between CMs and their linguistic realisations, metaphor can be understood from two perspectives in discourse analysis: the *conceptual* approach, which focuses on metaphor in thought (CMs), and the *linguistic* approach, which deals with metaphor in language (linguistic metaphors). Researchers often differentiate between these two facets, separating the psychological processes from the linguistic products. For instance, language-based metaphors such as the phrases "*rich* in health" and "*bankrupt* of health" are distinguished from the underlying thought-based metaphor, HEALTH IS WEALTH, which motivates and shapes these linguistic instantiations (Lakoff & Johnson, 1980, 1999).

At the linguistic level, metaphors are characterised by the "topic" they convey (e.g., well-being) through the "vehicle" terms, such as *rich* and *bankrupt*, the actual words used. This indicates a conceptual mapping between the source domain (WEALTH) and the target domain (HEALTH). Lexicon related to "money" is thus frequently employed to discuss the importance of health, framing abstract concepts in terms of more concrete, experiential associations. In this context, *rich* and *bankrupt* (words typically associated with the "money" lexicon) are understood in terms of "health" to imply good health and a severe lack of health, respectively.[7]

Applied metaphor research often focuses on conventional uses of metaphors, primarily emphasising their linguistic and conceptual properties. In contrast to these general trends, a third approach emerges, focusing on the discourse context or the deliberate use of metaphors[8] (e.g., Cameron, 2003; Charteris-Black & Musolff, 2003; Nacey, 2013; Semino, 2008; Cuberos et al., 2019). This research offers a different perspective, exploring the communicative functions of linguistic metaphors used by language users in discourse.

For example, in a recent longitudinal study, Nacey (2022) explored the role of metaphor clusters in learner discourse among teenage Norwegian L2 learners of English

[7] For a comprehensive overview of CMT, see Lakoff (1993) and Gibbs (2014).
[8] For a comprehensive discussion on *Deliberate Metaphor Theory* (DMT), see Steen (2023).

(aged 13–17), building upon her previous research (see Nacey, 2020). She found that over time, these learners produced more metaphor clusters with increasingly purposeful functions, both topically and discursively. However, many metaphor clusters appeared to lack discernible motivation, even in the later school grades, where students are expected to reach an upper-intermediate level for university admission (see Nacey, 2022, p. 185).

The present study employs a linguistic approach to operationalising the concept of metaphor in learner discourse, encompassing both spoken and written production in an L2. It is important to emphasise, however, that the conceptual aspect of metaphor remains pertinent to this study. While the primary focus is not on analysing metaphor comprehension, the research does address the idea of metaphor as a mental process. This is particularly evident in the implementation of CL-inspired activities, which aim to foster deep cognitive processing and enhance understanding of metaphorically motivated language use. The main goal is to examine conventional linguistic metaphors produced by L2 speakers, primarily focusing on metaphor density as an indicator of lexical *growth*.

2.1.2. Analysing Metaphor in Use: Learner Discourse Perspectives

In the last forty years, there has been a sharp rise in interest in metaphor, leading to its investigation in real world contexts. The linguistic approach to metaphor, focusing on its occurrence in natural language settings, has deepen how metaphors work in both L1 and L2 discourse. Recent research has delved into metaphor use in diverse environments, such as climate change discussions in UK secondary schools (Deignan et al., 2019) and in palliative healthcare (Semino et al., 2018). There is also a growing focus on naturally produced metaphors in L2, particularly in written learner discourse. Some notable studies in this area include those by Hoang and Boers (2018), Littlemore et al. (2014), and Nacey (2013, 2020, 2022). This research is discussed in more detail in Section 2.3 of this chapter.

2.1.2.1. Thematic Elements: Genre and Topic

The linguistic approach to metaphor has shown that metaphor use, in terms of density and types, varies considerably across different contexts due to the nature of genres and topics covered, as observed in both L1 and L2 discourse (Deignan et al., 2013; Semino, 2008).

In standard L2 competence contexts, the specific characteristics of the production task may contribute to increased use of metaphorical language. Research has demonstrated that more abstract topics, in particular, tend to encourage increased use of metaphor (Hoang & Boers, 2018; Littlemore et al., 2014; Nacey, 2013, 2020, 2022). As described in Chapter 3, learners at the B2 level are expected to engage in spoken and written tasks on abstract topics, requiring the expression of opinions and the personal significance for the speakers. This often necessitates more sophisticated language use (Littlemore et al., 2014). Additionally, the effort by L2 learners to excel by employing more complex language, as influenced by task conventions, might also lead to an increased use of metaphors (see Nacey, 2010, 2013).

The complexities inherent in metaphor usage in L2 learning contexts pose unique challenges for research. To ensure topic consistency, previous research has selected texts on related subjects within a broad theme. However, as Nacey (2022) highlights, maintaining topic uniformity is a significant challenge in metaphor research, especially in longitudinal studies where topic choice might not be consistently controlled. Even if learners responded to identical prompts in the different testing measures (i.e., pre-test and post-test), task familiarity with the B2 exam parts could influence their language production (see Aas & Nacey, 2019).

2.1.2.2. Language Modes: Conversations vs. Compositions

These observations from previous studies are typically based on the analysis of naturally occurring metaphors in written discourse. Thus, it is also important to highlight the characteristics of different language modes, as the communicative behaviour of L2 learners may differ between spoken and written production, potentially affecting metaphor use.

MacArthur and Littlemore (2011) analysed the use of metaphorical language in natural face-to-face interactions, focusing on interactive communication between L1 and L2 speakers

of English. They examined metaphors in both semi-structured interviews and spontaneous conversations, finding that metaphors facilitated the expression of ideas and the development of topics. The repetition of words with potential for metaphoric extension proved especially valuable for L2 speakers in building coherent conversations with L1 speakers. The study further highlighted factors such as speaker relationships, linguistic and cultural backgrounds, the purpose of conversations, and topics as key factors that influence metaphor usage in speech. However, MacArthur and Littlemore (2011) emphasised a distinct feature of face-to-face interactions, as opposed to monologic written texts: the dynamic production of discourse through “the joint efforts of the conversational partners” (p. 203).

In a separate study, MacArthur (2016a) explored the use of metaphor in office-hour consultations involving Spanish undergraduates. Echoing her earlier research with Littlemore, she observed that the communicative success of metaphors largely hinged on how conversational partners enacted their roles as collaborative participants in the conversation.

In the context of high-stakes ESOL examinations, it is also essential to acknowledge the significant and multifaceted differences between planned discourse (*Writing*) and unplanned or semi-planned discourse (*Speaking*) across various aspects such as structure, language use, interaction dynamics, and the roles of participants.

Planned discourse, often seen in monologic written communication, is characterised by its structured format, allowing for careful planning, revision, and adherence to formal language and style. This type of discourse is often associated with higher levels of lexical richness and syntactic complexity (Taguchi et al., 2013). In contrast, unplanned or semi-planned discourse, typical in dialogic oral communication, is more spontaneous and fluid, with participants engaging in real-time interactions that are less formally structured and often more informal in language use. Such discourse is marked by immediacy and a focus on interaction, often leading to a more dynamic use of language (Cameron & Deignan, 2006).

These distinctions are evident in the structure of the discourse; written communication is well-organised and cohesive, while oral communication is dynamic and adaptable (Halliday & Matthiessen, 2014). Language use in written form tends to be more complex and formal,

whereas oral language is characterised by colloquialisms and a conversational tone (Biber & Gray, 2016). The interaction dynamics also differ significantly: monologic communication is a one-way process, whereas dialogic communication is interactive, involving continuous exchanges between participants (Webb & Nation, 2017). Consequently, the roles of participants in written discourse are fixed, with a clear distinction between the writer and the audience, while in oral discourse, these roles are more interchangeable and collaborative (Hughes & Heasley, 2018).

Understanding the distinct characteristics of these modes is crucial, particularly in L2 instruction and assessment. Recognising the unique features and challenges of each mode can lead to a deeper understanding of L2 metaphor use, thereby informing the development of more effective pedagogical strategies. This, in turn, could support improvements in learner performance. Such understanding is especially relevant in standard L2 assessment contexts where both *Speaking* and *Writing* skills are critically evaluated. However, to the best of my knowledge, the specific communicative behaviour of L2 learners in testing environments, particularly in relation to their metaphor usage, remains underexplored.

2.1.2.3. Uses of Metaphors: Conventionality in L2 Discourse

Research into real-world contexts has shown that L1 speakers of English typically convey meaning using conventional metaphors, which are deeply entrenched in everyday language and recorded in dictionaries (Cameron, 2003, 2008). These metaphors can manifest in language use as single polysemous words (e.g., *flooded* in the sense of being overwhelmed with tasks), collocations (e.g., “drowning in work”) and idioms (e.g., “burning the candle at both ends”). However, while these metaphors might resonate intuitively with L1 speakers, they can often be obscure and challenging for L2 learners of English.

This linguistic challenge is reflected in Littlemore’s (2023) discussion on construal patterns. She states “the words we use to talk about a particular phenomenon can never reflect a purely objective view of that phenomenon. We can only witness phenomena through human eyes and from a human perspective” (p. 5). This observation underscores the inherent

subjectivity in language use, suggesting that even default ways of describing situations through CMs are imbued with subjective perspectives, thereby lacking complete neutrality.

Consider the difference between “drowning in work” and “managing a workload”. While both phrases describe being engaged with a lot of tasks, “drowning in work” employs a metaphor suggesting being overwhelmed and incapacitated, as if submerged under water. In contrast, “managing a workload” presents a more controlled and orderly scenario, implying a capability to handle tasks effectively. These metaphorical uses of language are not merely stylistic choices; they demonstrate our ability to frame experiences in various ways.

Language often contains conventional ways of construing phenomena and events, which sometimes differ from those in other languages. This variance presents unique challenges for L2 learners in appropriately understanding and using metaphors. For instance, when a Spanish learner of English directly translates metaphorical expressions from Spanish to English, it can lead to unconventional language use. The Spanish phrase *estar en las nubes* literally translates to “to be in the clouds”. In English, a similar expression is “having one’s head in the clouds”. While both use the metaphor of being up in the clouds, implying being lost in thought or daydreaming, the cultural connotations and nuances might slightly differ considering the Spanish phrase often carries a more poetic or whimsical undertone. Furthermore, a Spanish speaker might say, “she is always walking in the clouds”, attempting to convey the idea of *estar en las nubes*. While “having one’s head in the clouds” is more conventional, this direct translation, “walking in the clouds”, is not standard English as the exact words that constitute the shared CM are different in both language (Littlemore & Low, 2006a). This kind of direct translation can lead to unconventional expressions in English, reflecting the learner’s L1 linguistic and cultural background.

Littlemore (2023) highlights that successful L2 learning requires to “develop a degree of cognitive flexibility and openness to new ways of seeing things” on the part of the learner (p. 265). L2 learners need to overcome the cognitive habits developed from speaking their L1 by adapting to how CMs are conventionally used in the L2. Additionally, successful learning may be also influenced by the concept of tolerance of ambiguity (Ely, 1989), which measures

how comfortable a person feels in unfamiliar or ambiguous situations. This trait significantly influences L2 learners' risk-taking ability, as those who can tolerate ambiguity are better equipped to accept that the L2 presents information differently from their L1.

Another learner trait that may affect the acquisition of new construal patterns in an L2 is the individual's cognitive ability for "novelty", which aids learners in overcoming the effects of L1 transfer. This ability, as defined by Littlemore (2023), entails "to spot new patterns in the language input and to use one's existing knowledge selectively, along with analogical reasoning, to work out new form-meaning pairings" (p. 266). It becomes particularly crucial when learners encounter conventional metaphors in the L2, which may seem novel or even creative when the meanings of similar words in their L1 differ significantly (see Picken, 2007). However, Nacey (2013) found relatively few instances of creative metaphor in her investigation into the use of metaphor by Norwegian learners of English. This suggests that the novelty of L2 metaphors does not always translate into creative usage by L2 learners.

For example, the English verb *run*, in its basic (literal) sense, corresponds to the Spanish verb *correr* (as in "run a mile", translated as *correr una milla*). However, in English, *run* also has extended (metaphorical) meanings in expressions such as "run out of time" or "run a business". These metaphorical uses do not directly translate to Spanish. For "run out of time," the Spanish equivalent is *quedarse sin tiempo*, and for "run a business," it is *llevar un negocio* or *dirigir un negocio*. In each case, a different verb or expression is used in Spanish to convey the metaphorical meaning that *run* has in English.

Such divergence in metaphor use between languages presents important challenges in L2 learning, especially in formal instruction. This learning process involves the careful acquisition of norms and standard uses of an L2, with a specific aim to achieve native-like accuracy in standard L2 competence exams. In this context, while there are potential benefits of creative metaphor use for enhancing learners' MC as shown by Littlemore et al. (2023), Littlemore's (2023) recommendation to encourage learners, particularly those below advanced levels, to understand and produce what they perceive as "creative" metaphor may not be advisable.

2.1.3. Identifying Metaphor in Language Use: A Focus on Learner Discourse

Since the 1980s, the growing interest in metaphor has led to the development of robust methodologies for analysing this phenomenon in real-world contexts. A key aspect of metaphor research is the identification of metaphors in discourse, requiring the analysis of linguistic expressions to determine if they are being used metaphorically. One major challenge is developing standardised procedures and protocols that different researchers can consistently apply. Without a replicable method, metaphor identification risks becoming subjective and overly dependent on the analysts' language knowledge. Therefore, establishing reliable methodologies that minimise individual researcher bias and ensure consistency in analysis is crucial. Standardised methods enhance comparability between studies and analysts, leading to more trustworthy and robust conclusions.

Various approaches for rigorously identifying metaphors in empirical research have been proposed (e.g., Cameron, 2003; Cameron & Maslen, 2010; Charteris-Black, 2004; Pragglejaz Group, 2007; Steen et al., 2010). These methods have significantly contributed to metaphor research in real-world contexts, providing insights into metaphor frequencies or densities and identifying bursts of metaphor use at specific points in discourse (e.g., Alejo-González, 2022; Cameron & Stelma, 2004; Littlemore et al., 2014; Hoang & Boers, 2018; Nacey, 2022). The MIPVU procedure (Steen et al., 2010), based on the MIP protocol developed by the Pragglejaz Group (2007),[9] has gained widespread acceptance for its effectiveness in identifying linguistic realisations of CMs across various discourse contexts. It offers a detailed approach for determining the metaphoricity of each lexical unit (LU)[10] in natural discourse, employing analytical criteria and external resources, such as dictionaries, to guide the identification process and reduce bias.

MIPVU has been applied to L1 English by the VU Amsterdam Corpus (VUAMC; Steen et al., 2010) group across diverse genres. As mentioned in Chapter 1, this revealed 10.80%

[9] For nuanced details about the differences between MIP and MIPVU, see Nacey (2013, pp. 69–79).
[10] The term "lexical unit" is defined in Section 3.3.1.2 of Chapter 3.

metaphor usage in fiction (Dorst, 2011), 17.50% in academic texts (Herrmann, 2013), 6.80% in conversation (Kaal, 2012), and 15.30% in news (Krennmayr, 2011).

The effectiveness of MIPVU in analysing written discourse in L2 English across various proficiency levels, L1 backgrounds, ages, and educational stages is further highlighted by research from various studies on learner English. For example, Nacey's (2013) analysis of non-academic essays from undergraduate Norwegian learners, using the NICLE corpus of approximately 20,000 words, reported a metaphor density rate of 15.50%. Littlemore et al. (2014) investigated metaphor use in 200 written essays by adult learners of English who had successfully passed *Cambridge* ESOL exams across the CEFR levels A2 to C2, with 20 texts per level and learner group. At the B2 level, they observed metaphor densities of 9.90% and 11.62% in argumentative essays by 20 Greek- and 20 German speakers, respectively. Hoang and Boers (2018) analysed 257 essays by L1 Vietnamese learners of English, spanning three different undergraduate years. These essays, focusing on modern society and particularly emphasising the relevance of literature, revealed a metaphor density of 13.49%.[11]

Additionally, MIPVU has been employed for analysing spoken metaphors by L2 speakers in ELF contexts. This is evidenced by the EuroCoAT (The European Corpus of Academic Talk) corpus (MacArthur et al., 2014), which reported 11.90% metaphor usage (as observed by Alejo-González, 2022), and the MetCLIL (Corpus of Metaphor in Academic Talk) corpus (Alejo-González et al., 2021), which found 14.22% usage rate (see Alejo-González, 2024). Furthermore, the application of MIPVU extends to languages beyond English (see Nacey et al., 2019a, for further details).

Reporting the methods used for identifying metaphors in analysed discourse is essential, but so is ensuring the reliability of these procedures. Despite MIPVU being widely acknowledged as a reliable method, previous studies have pointed out a potential for analysts to display a positive bias toward identifying certain LUs as non-metaphorical, or overlooking

[11] These researchers employed a mixed methodology for identifying metaphor in discourse.

metaphor instances (Littlemore et al., 2014; Nacey, 2013; Steen et al., 2010). Thus, evaluating the quality of the metaphor identification process is vital for achieving trustworthy results.

One method to test the reliability of MIPVU is the inter-rater approach, involving external researchers independently applying the procedure to the same data sample (Cameron, 2003). This method is widely recognised for identifying discrepancies or inconsistencies in the analysis, thereby increasing the validity and replicability of the results. Robust figures of inter-rater agreement are crucial in this context, as discussed by Nacey et al. (2019a, pp. 155–157).

2.2. METAPHOR IN TEFL: BRIDGING THEORY AND APPLICATION

CL represents a well-established theoretical approach in the field of linguistics that has garnered significant recognition. In the last four decades, the pervasive role of metaphor in our daily communicative exchanges and its influence on the way we conceptualise abstract concepts using more concrete ones have been studied extensively. The theoretical tenets and conceptual constructs of CL have been applied across various disciplines, such as psychology, discourse studies, literature, philosophy, translation studies, and artificial intelligence (see Piquer-Píriz & Alejo-González, 2016, 2018).

One area within Applied Cognitive Linguistics (ACL) that has received a great deal of attention is L2 pedagogy. The principles of CL hold substantial implications for L2 instruction, providing strategies to address challenges in learning, especially features such as vocabulary, including non-literal language. In recent years, the analysis of L2 figurative language use, especially the ability to produce and understand metaphor, has become one of the most fruitful topics in ACL research (Achard, 2018; Achard & Niemeier, 2004; Boers & Lindstromberg, 2008a; De Knop et al., 2010; Ervas et al., 2019; Holme, 2009; Hijazo-Gascón & Llopis-García, 2019; Ibarretxe-Antuñano et al., 2019; Littlemore, 2009, 2023; Littlemore & Juchem-Grundmann, 2010; Littlemore & Low, 2006a; Piquer-Píriz & Alejo-González, 2020; Pütz et al., 2001a, 2001b; Pütz & Sicola, 2010; Robinson & Ellis, 2008; Tyler, 2012). This body of work has shed some light on the relationship between language and thought and it also has

revealed important implications for L2 learning as well as to understanding how learning an L2 also influences cognitive processes (for more details, see Littlemore, 2023).

There are a number of inextricably linked key concepts in CL, which are of particular relevance to L2 instruction. CL offers an alternative approach to the conventional notion of linguistic "arbitrariness" in L2 vocabulary learning. The concept of *linguistic motivation* suggests that our understanding of language stems from our physical, social, and cultural experiences (Radden & Panther, 2004). In this framework, meaning is constructed through conceptualisation, informed by our *encyclopaedic knowledge* of the world. This knowledge extends beyond denotative meanings, encompassing connotations associated with words and expressions after repeated exposure in various contexts. Such insights highlight the *embodiment* of CMs, suggesting that language is shaped by our physical experiences and our interactions with our surroundings.

Boers and Lindstromberg (2006) argue that language can be motivated in three distinct explainable mechanisms: *form-form* connections (e.g., alliteration), *form-meaning* connections (iconicity), and *meaning-meaning* connections (e.g., polysemy). Among these, the last plays a particular crucial role in raising L2 learners' awareness of the motivated aspects of meaning, especially in explaining the non-arbitrary associations between basic and metaphorical meanings.

Polysemy can be defined as a "generalization over related senses of linguistic expressions" (Lakoff, 1993, p. 209). The view that language is motivated (i.e., explainable), rather than entirely arbitrary, extends to the concept of *motivated polysemy*, where words adopt multiple, related meanings. Such linguistic motivation leads to the construction of radial categories, suggesting that the various senses of particular words are viewed as semantic networks, with the more concrete, physical senses lying towards the centre of the category and the most abstract, metaphorical senses lying towards the periphery. The different senses, related through metaphor and metonymy, are interconnected, and derive from our bodily experiences and interactions with the world (Lakoff, 1987; Langacker, 1990; Taylor, 2003).

For instance, the semantic extensions of the terms *up* and *down* when used to describe emotions are grounded in our bodily experience of being higher or lower. In English, the term *up* is often linked to positive emotions, as in “uplifted spirits”, indicating a joyful state, much like the physical feeling of rising or being at a higher elevation. On the other hand, the term *down* is used in expressions such as “feeling down”, suggesting negative emotions, akin to the physical state of being closer to the ground or moving downwards.

The integration of insights from CL into L2 instruction has particularly attracted a lot of attention in EFL settings (see MacArthur, 2017). However, there has been a recent increase in research applying ACL principles to languages other than English, with studies on French (Achard, 2018), German (De Knop & Dirven, 2008), Italian (Danesi, 2008), and Spanish (Ibarretxe-Antuñano et al., 2019). In particular, Spanish as an L2 (SSL) has seen a growing interest in research into the practical applications of CL, especially in the area of Cognitive Grammar (see Hijazo-Gascón & Llopis-García, 2019, for a review). Recent research has also explored the potential of using CL-based multimodal input, such as audio-visual media, for teaching metaphor. This approach has been shown to benefit the spoken and written linguistic performance of English learners of Spanish, as observed by Martín-Gascón (2023).

The ubiquity of metaphor in language use has sparked widespread interest in researching CL-oriented methods to teach and learn metaphors in instructed L2 contexts, due to its crucial implications for L2 learning (Bielak, 2011; Boers & Lindstromberg, 2006, 2008a; De Knop et al., 2010; Littlemore & Low, 2006a; Low, 1988; Piquer-Píriz & Alejo-González, 2020; Piquer-Píriz & Boers, 2019). L2 learners not only engage with metaphor from an early age (MacArthur & Piquer-Píriz, 2007; Piquer-Píriz, 2008a, 2008b, 2010, 2020) but also at various stages of L2 learning (Littlemore et al., 2014; Nacey, 2022; O’Reilly & Marsden, 2023).

Metaphorical language use is pervasive throughout the L2 learning process, affecting both input and output discourse in instructional and assessment contexts. Consequently, learners undoubtedly need to deal with conventional metaphors throughout their L2 instructed acquisition, which can be immensely beneficial yet also very challenging. Thus, support is essential to maximise these advantages and mitigate the difficulties presented by metaphors.

2.2.1. Mastering Metaphorical Meanings in L2 Learning: Benefits and Challenges

In his ground-breaking article on teaching metaphor, Low (1988) emphasised the importance of metaphorical language in TEFL, drawing on the pedagogical insights from Lakoff and Johnson's (1980) conceptual theory. Since the 1990s, a growing number of scholars have highlighted the crucial role of metaphor in L2 instruction (see Piquer-Píriz & Alejo-González, 2020, for a review). Given the pervasive nature of metaphor in oral and written discourse, L2 learners' mastery of metaphor can be hindered by misunderstandings or misuse, often resulting in them being perceived as *outsiders*. Even advanced learners tend to avoid using words in their metaphorical sense, leaning more towards their literal meanings (Littlemore, 2009).

Littlemore (2023) suggests that this hesitance could arise from L2 learners not noticing the metaphorical uses of language in the input they are exposed to. Alternatively, while they might be aware of the metaphorical meanings present in their passive vocabulary, these meanings may not yet have been integrated into their active use, potentially due to a lack of confidence. Nevertheless, mastering metaphorical meanings is crucial for achieving natural and effective communication in the L2. Consequently, it is of paramount importance to assist L2 learners in using metaphors appropriately since, by extension, proficiency in using metaphors will be significant for L2 competence assessment.

Knowledge of metaphorical language is crucial for reaching L2 proficiency, enhancing both receptive and productive fluency. Littlemore and Low (2006a) showed that metaphor is an integral part of speakers' overall communicative competence, contributing to all four language skills: *Listening*, *Reading*, *Speaking*, and *Writing*. Drawing on Bachman and Palmer's (1996) model, they demonstrated how metaphor affects the four dimensions of communicative competence (grammatical, textual, illocutionary, and sociolinguistic) in an L2. They found that L2 learners face challenges with metaphor across all these dimensions.

Littlemore and Low's (2006a) MC construct, which includes informal experiences and suggestions for developing these in the L2 classroom, has been empirically tested through theory-driven, valid, and reliable instrumentation in O'Reilly and Marsden's (2021, 2023)

research. These studies have examined the MC construct in relation to other aspects of L2 knowledge, such as vocabulary and overall proficiency, offering insights to inform both assessment and pedagogical interventions (for more details, see Section 2.3 of this chapter). However, as further discussed in Section 2.3.3, metaphorical language use has not been fully integrated into mainstream communicative competence models, such as the CEFR, which are widely used in standard L2 competence assessment.

Metaphorical language should not be viewed in isolation from other aspects of linguistic competence. Instead, metaphor is embedded in lexis and underpins grammatical structures (e.g., Goldberg, 1995; Langacker, 1987) as well as serving diverse discourse functions (e.g., Cameron, 2003; Semino, 2008). For those aiming to become fluent in the L2, knowledge, comprehension, and production of metaphorical language are indispensable. Metaphor stands as a key indicator of learners' ability to operate at different levels of L2 proficiency (Hoang & Boers, 2018; Littlemore et al., 2014). Consequently, achieving a proficient level in the L2 can also hinge on learners' fluency with metaphors.

Despite the usefulness of metaphor in discourse, L2 learners often struggle with their comprehension and production (Littlemore & Low, 2006a). While the importance of teaching metaphorical language has gained recognition in recent years (Bielak, 2011; De Knop et al., 2010; Littlemore, 2023; Piquer-Píriz & Alejo-González, 2020), addressing metaphors in TEFL presents some challenges shaped by conceptual, experiential and culture-based linguistic aspects. The ubiquity of metaphor in language can be a double-edged sword for L2 learners, adding complexity to their learning process. Various factors in the teaching-learning process, such as lack of basic vocabulary or cultural knowledge, contribute to these difficulties.

2.2.1.1. Metaphor as a Vocabulary Builder in L2 Learning

Mastering metaphorical meanings can serve as a vocabulary builder for L2 learners As Littlemore (2023) emphasises, "learning a second language involves the ability to reorganize our encyclopaedic knowledge and corresponding word association networks, thus deepening our knowledge of L2 vocabulary" (p. 265).

When it comes to mastering a language and/or enhancing vocabulary skills, it is not only about learning thousands of words and their meanings (i.e., vocabulary *breadth* [quantity]) but also about perceiving lexical combinations and relationships (i.e., vocabulary *depth*, [quality]) (Meara, 1996). Even at advanced levels, L2 learners may not have the necessary tools for effective communication as "they know fewer words, [but also] they have a smaller network of semantic or conceptual links" compared to L1 speakers (Littlemore & Low, 2006a, p. 23). Therefore, an awareness of metaphorical meaning can aid both comprehension and production in an L2.

A good deal of research supports that L2 learners can establish *meaning-meaning connections* among the different senses of a word in an L2 at various stages of education, even from a young age (Boers, 2000a; Lindstromberg & Boers, 2006; Piquer-Píriz, 2005, 2008a, 2008b, 2010, 2020). Therefore, these insights have crucial pedagogical implications for L2 instruction, highlighting the need of learners to explore the metaphorical motivation behind the polysemous senses of lexical items.

Many polysemous words are more frequently used in their figurative senses than in their literal ones (Hoffman, 1983). However, vocabulary instruction has traditionally focused on teaching the basic, literal meanings of words, often sidelining the importance of familiarising L2 learners with frequently used forms, especially their metaphorical senses (Littlemore & Low, 2006a). For example, L2 learners might be taught that "head" refers to the part of the body above the neck, but they might not initially learn its other extended meanings, such as someone who leads or oversees a group (Piquer-Píriz, 2004).

However, the teaching of the polysemous senses of frequently used lexical items can foster deeper cognitive processing by effectively linking basic to extended meanings. Verspoor and Lowie (2003) highlight the effectiveness of providing L2 learners with the basic sense of a metaphorically used word as a cue to its figurative meaning in context, thus enhancing long-term retention of this meaning. Echoing this approach, Littlemore and Low (2006a) advocate for a balanced teaching strategy, recommending L2 instructors, "where feasible, teach basic

senses first and balance this against teaching the most frequent forms first" to help learners uncover the metaphorical motivation for polysemous uses of language forms (p. 26).

Building on this approach, MacArthur (2010) suggests that L2 learners should "break down the meaning of component words in figurative expressions [to] consider what is implied by the use of words" (p. 166). This can help learners develop an intuitive grasp of words, which is essential for their metaphorical application. However, MacArthur (2010) also observes that "learners often appear not to possess that 'feel' for or 'grasp' of their basic senses that is necessary for using them metaphorically" (p. 166). To address this issue, MacArthur (2010) stresses the importance of raising *metaphor awareness* in the L2 classroom by revisiting familiar lexical items to understand their basic senses. Making these connections can be seen to help L2 learners consolidate and integrate a word into their lexicon (Verspoor & Lowie, 2003). As Boers (2004) argues, "enhanced metaphor awareness can serve as a vehicle for vocabulary acquisition" (p. 214).

Accordingly, MacArthur (2010) suggests that metaphor plays an important role as a mechanism for semantic extension. She argues that fostering L2 learners' vocabulary depth can enrich their lexical repertoire by enhancing their understanding of lexical relationships. By exploiting the semantic potential of frequently used vocabulary (already known or to be learnt), heightened awareness of metaphors can be particularly beneficial in L2 learning.

However, enhancing the semantic potential of frequently used lexical items in the L2 classroom might not always be appropriate. MacArthur and Littlemore (2008) found that the frequency of word's metaphorical senses does not necessarily correspond to how transparent or easily understood these meanings are for L2 learners. This suggests that the most frequent lexical items should not automatically be given teaching priority. Instead, instructors should consider the clarity and teachability of their metaphorical senses, as well as the practicality of CL-oriented teaching methods, rather than adhering strictly to frequency-based methodology.

MacArthur (2017) further notes the difficulty in determining what constitutes a basic sense, especially given that L2 learners might frequently encounter metaphorical senses early on, such as with classroom management language (e.g., *fill in* "fill in the gaps": "to write all the

necessary information on an official document, form etc" [LM1]), or even before the literal sense, as seen with technical terms (e.g., *module*: "one of the separate units of a course of study" [MM1]). In some other cases, the literal sense may never be encountered (e.g., *culture*: "a group of bacteria or other cells that have been grown in a scientific experiment [MM3]).

Vocabulary growth in relation to the number of new lexical forms L2 learners master (breadth) can be unpredictable. However, MacArthur (2010) emphasises that fostering depth of vocabulary knowledge is "amenable to planning" (p. 159). Becoming deeply familiar with target lexical items provides L2 learners with the support they need to discuss a broader range of topics by using everyday vocabulary. Such planning also helps them express their thoughts precisely in a native-like manner, both orally and in written form, "through the appropriation of target language forms and their associated metaphorical and cultural ideas" (MacArthur, 2010, p. 158). Elaborating on this, MacArthur (2010) explains:

> FL learners do not have to master a huge number of vocabulary items in order to express complex thoughts on abstract topics. Indeed, metaphorical uses of language will become increasingly necessary for FL learners as they move from talking or writing about the here-and-now and move into the realm of abstraction. (p. 161)

Raising awareness of metaphorical expressions among L2 learners can equip them with tools to master metaphorical meanings more efficiently, thereby enhancing their L2 proficiency. Researchers, such as Boers (2000b), have shown that when L2 learners are aware of CMs, they tend to understand and recall them more effectively. However, the potential benefits of heightened metaphor awareness extend beyond mere enhanced retention of the learned vocabulary. MacArthur (2010) suggests that it can be an invaluable asset to L2 learners, acting as their "best ally in the quest for greater expressive powers" (p. 159). This enhanced awareness can not only facilitate more flexible use of familiar English vocabulary but also promote fluency in metaphor production.

2.2.1.2. Enhancing Cultural Awareness through Metaphorical Language in TEFL

The emphasis on metaphorical language in the L2 classroom has also grown largely due to the cultural dimension of metaphor. Engaging with metaphors not only provides insights into the target culture but also fosters cultural awareness among L2 learners through cross-linguistic and cross-cultural comparisons.

Metaphorical mappings, however, are not universally consistent across languages. Lakoff and Turner (1989) assert that CMs are part of the common "conceptual apparatus" shared by members of a culture (p. 51). The dynamic nature of cognitive process allows different language to incorporate various ways of perceiving and describing experience. Each culture has unique associations, shaped by distinct cultural backgrounds that may not always be directly experienced by the members of a speech community. Everyday metaphorical language use, grounded in physical and social experience, can be highly culture specific and becomes widely conventionalised within the linguistic framework of the target culture (Deignan et al., 1997; Kövecses, 2002).

Previous studies have shown that the cultural motivation behind metaphors results in cross-linguistic variation, meaning that speakers of different languages use varied metaphors to describe similar areas of human experience or express with different linguistic realisations the same CMs (Boers, 2003; Gibbs, 2012; Kövecses, 2005, 2009; Littlemore, 2001b, 2003; Littlemore & Low, 2006a; Littlemore et al., 2014; MacArthur, 2010, 2016c; Nacey, 2010; Philip, 2005, 2010). Metaphor may also be culturally motivated, underpinned not only by linguistic factors but also by historical, cultural, and social contexts. Consequently, this embedding can pose important challenges for those learning an L2.

For example, the notion of *help* is realised differently in both English and Spanish linguistic metaphors. In Spanish, the concept of *help* is expressed with the phrase *echar un capote* (lit. "throw a cape"). The term *capote*, meaning "bullfighting cloak", reflects how, in Spanish culture, the idea of *help* is metaphorically understood in relation to a specific cultural event: assisting a bullfighter with an extra bullfighting cloak when necessary during bullfights.

English speakers, by contrast, use the metaphorical expression “to bail someone out” to express the idea of *help*. While the term “bail” is associated with cricket — referring to “one of the two pieces of wood laid across the top of the stumps to form the wicket (MM2) — the phrase “to bail someone out” more commonly denotes the act of rescuing someone from a difficult situation, especially financially.

However, there are areas where English and Spanish metaphorical expressions align. For instance, both languages perceive the concept of *control* in a horse-riding scenario (MacArthur, 2005): CONTROL OF AN UNPREDICTABLE/UNDESIRABLE FORCE IS A RIDER’S CONTROL OF A HORSE. Both languages express this metaphor with similar linguistic expressions: the English “take the reins” and the Spanish *tomar las riendas*.

L2 learners frequently face challenges related to such metaphors inherent to the target language. Lacking the shared cultural and social experiences as native speakers, learners might inadvertently rely on L1 concepts when communicating in the L2, which can lead to misunderstandings and misuse of metaphorical language, as previously discussed. To address this issue, Boers (2003) emphasises the significance of learners grasping the cultural values of the L2 to understand and use metaphors more effectively. By mastering metaphorical meanings, learners can not only deepen their lexical knowledge but also enhance their cultural awareness.

Raising metaphor awareness can offer a way into the L2 culture, as this may aid learners to “discover new, foreign connections between words and concepts, and therefore help them to internalise conceptualisations belonging to the foreign language (Niemeier, 2004, p. 112). Having previous knowledge of a CM can better equip L2 learners to understand its linguistic instantiations, even if they encounter them for the first time. This becomes particularly relevant considering the array of human experiences across different languages that shape how concepts are metaphorically structured.

As MacArthur (2010, p. 159) points out, unlike grammar, there are no “hard and fast rules” for metaphors that determine what is correct or incorrect. L2 instructors face the decision of prioritising either the communicative efficacy of metaphorical language in its discourse

context (e.g., MacArthur, 2016c) or its conventional linguistic form (e.g., Philip, 2010). This inherent ambiguity introduces an added layer of complexity to the teaching and learning of metaphorical language. Raising metaphor awareness could support efforts to mitigate the potential communication issues caused by "hybrid metaphors", i.e., those mixing metaphorical conceptualisations and wordings from the L1 (MacArthur, 2016c). Therefore, there has been a recent spate of interest in how to teach and learn metaphor in L2 classrooms to enhance learners' performance and, thus, overall L2 learning.

2.2.2. Supporting L2 Learners in Using Metaphor: Methods and Applications

CL has explored in increasingly sophisticated ways the intricate relationship between cognition, language, and communication. Despite the vast complexity of linguistic phenomena, which intertwines with factors such as cognition, experience, embodiment, human interaction, society, culture, and history, certain patterns have emerged (Ellis & Robinson, 2008). Rather than being dichotomous, the CL perspective suggests that knowledge of language, whether L1 or L2, is acquired through use, viewing language as an emergent dynamic system (Ellis, 2006).

Language knowledge and learning are, therefore, usage-based. This approach highlights the importance of exposure to language within context. Ellis and Robinson (2008) note, "[w]hat is attended is learned, and so attention controls the acquisition of language itself" (p. 3). This exposure to input enables L2 learners to identify patterns and relationships specific to the language system and to hypothesise about language structures and use. Given that linguistic knowledge arises from language use and in interactive settings, testing their input-driven knowledge in authentic communicative contexts can enrich and deepen L2 learners' understanding of the language system.

While comprehending and producing metaphor is considered a natural process for a native speaker, dealing with the diversity of meanings often becomes laborious for L2 learners as they process metaphor analytically, in a more mechanical way (Kecskes, 2006). Lowery (2013) stresses that native speakers of English have "a lifetime of exposure to English

language and culture that helps them understand […] metaphor" (p. 12). In contrast, learners often have more restricted time and fewer opportunities to interact with the L2 in educational environments compared to natural or immersion settings, as Piquer-Píriz and Boers (2019) observe. Consequently, given the intricacy of language patterns in L2 learning, the incidental acquisition of metaphors can be particularly challenging. Rather, metaphor should be explicitly taught, providing EFL students with the tools to unlock the puzzle of metaphorical language in English.

2.2.2.1. CL-Inspired Instruction: Pedagogical Techniques

CL provides a crucial framework for aiding L2 learners in discerning the multiple yet interconnected meanings of words and enhancing their understanding of metaphor. This encyclopaedic view of meaning (Langacker, 1987) aligns with teaching aimed at the depth of linguistic knowledge and enhanced language awareness.

A large body of research has shown that fostering the notion of linguistic motivation among L2 learners through various CL-oriented teaching methods can facilitate the comprehension and retention of conventional metaphorical language across different areas such as polysemes (MacArthur & Littlemore, 2008; MacArthur & Piquer-Píriz, 2007; Piquer-Píriz, 2008a; Ponterotto, 1994; Saaty, 2020), idioms (Boers, 2001, Boers et al., 2008, 2009), phrasal verbs (Alejo-González, 2010; Condon, 2008; Dirven, 2001; Kövecses & Szabó, 1996; Rudzka-Ostyn, 2003), and prepositions (Lindstromberg, 1996, 2010).

CL specifically focuses on strategies that prioritise revealing semantically motivated relationships, examining the effect of fostering learners' deep engagement with metaphors, and exploiting their imagery to elucidate metaphorical meanings. Central to this is the *focus-on-meaning* approach to teaching L2 vocabulary, which revolves around meaning-oriented activities (Boers & Lindstromberg, 2006, 2008b). This focus on linguistic motivation contrasts sharply with traditional L2 vocabulary techniques, which often treat word meanings as arbitrary without exploring the deeper, often culturally or experientially grounded meanings of words.

The benefits of CL-oriented methods for L2 learners include meaningful learning, reduced memory load, and heightened language awareness due to the focus on semantic links. CL-style instruction equips L2 learners with a deeper understanding of word meanings, aiding the integration of new vocabulary into their active language use. CL-oriented methods can assist L2 learners in building upon pre-existing knowledge, promoting the use of familiar words in extended senses (Verspoor & Lowie, 2003). When learners are exposed to the literal meaning of a metaphor, they are more likely to remember it (Wang et al., 2020). This approach to vocabulary learning becomes more systematic and organised (Boers & Lindstromberg, 2008b), promoting recall as words are remembered within meaningful contexts, rather than as isolated entities (Boers, 2004). Such enriched learning allows for greater flexibility in using existing vocabulary, ultimately enhancing L2 proficiency. Additionally, raising metaphor awareness with CL-oriented methods can foster a more critical view of texts (Holme, 2004).

Various pedagogical techniques have been explored to raise metaphor awareness, as evidenced in controlled quasi-experimental studies carried out in instructed L2 settings (see Boers, 2013). Many of these strategies, particularly teacher-led explanations, visual tools, and kinaesthetic approaches, are familiar to L2 practitioners and learners. While these techniques are not presented as being a totally new methodology, CL aims to support and amplify certain aspects of L2 instruction that deserve further attention, giving them a unique twist. Boers and Lindstromberg (2006, p. 313) highlight the allure of a CL-inspired approach to L2 vocabulary instruction, for both learners and teachers. It can be more engaging compared to traditional methods, as it ties L2 learning to the learners' own experiences and cognitive processes.

Indeed, studies have shown that deep cognitive processing, aided by *semantic and/or etymological elaboration*, contributes to effective learning of metaphorical language. For example, this is evident when L2 learners receive explicit instruction of metaphors through guided *verbal explanation* about the motivation behind extended meanings (Condon & Kelly, 2002; Littlemore et al., 2013) or discussions on their etymology (Boers et al., 2004b). Active engagement with the language by identifying underlying CMs through *conceptual grouping* (Boers, 2000a; Beréndi et al., 2008) or *guessing strategies*, such as brainstorming and

speculation, complement explanatory input, solidifying understanding of metaphors (Skoufaki, 2008; Verspoor & Lowie, 2003).

Meaning-meaning connections between the basic and extended senses of words can be also enhanced through *pictorial elucidation*, using either imagined or illustrated visual images (Dual Coding Theory, e.g., Paivio, 1990). Techniques such as employing photographic visuals (Boers et al., 2008) and drawings (Lindstromberg, 1996, 2010) have proven effective to make metaphors more memorable.

Given the experiential basis of CMs, L2 learners can also infer meanings through physical enactment or Total Physical Response (TPR; Asher, 1981), relating extended meanings to their basic ones (Lindstromberg & Boers, 2006). For instance, in a CL-style embodiment, the phrase *grasp an idea* can be physically connected to the action of seizing an object in one's hand. Metaphorically, this conveys achieving clarity or understanding on a particular concept or thought.

While CL-oriented methods can prove beneficial in mastering conventional metaphorical meanings, they should not be viewed in isolation. Instead, they ought to be integrated into a broader understanding of the overall L2 learning process, encompassing its social and contextual areas (Littlemore and Juchem-Grundmann, 2010). In fact, context is pivotal in enhancing metaphor awareness. In this regard, a syllabus that lacks contextualisation can constrain material design, and potentially complicate proficiency in EFL.

It should be acknowledged that a great deal of language may not be motivated and must be learned as such. However, Littlemore (2023) highlights the crucial role of L2 instructors in raising learners' awareness of the motivated aspects of language within the target input source, when feasible, to aid in mastering metaphorical meanings.

Matching metaphorical expressions to precise and unique CMs is not always straightforward, potentially resulting in confusion for learners. To illustrate, the phrases *shoot down* (ARGUMENT IS A BATTLE) and *blow up* (ANGER IS AN EXPLOSION) are suitable candidates for raising metaphor awareness in the L2 classroom since they are clear instances of general CMs. In contrast, CMs often intersect, with multiple metaphors at play simultaneously. For

example, an expression such as “on top of the world” may be presented by instructors as an instantiation of HAPPINESS and SUCCESS metaphors. Besides, specific CMs can be subsumed under more general ones. For instance, an expression such as “breaking the glass ceiling” can be presented as an instantiation of a CAREER metaphor as well as a more general SUCCESS metaphor. To address these issues, Boers (2004) recommends drawing learners’ attention mainly to “clear” cases, i.e., expressions whose source domain can be pinpointed unambiguously, considering that those CMs typically require less cognitive effort for learners to understand and for teachers to explain. Additionally, Boers (2004) suggests giving preference to the more specific CMs over the more general ones, as they will be useful in explicit metaphor processing.

Furthermore, it is vital to understand that the concept of linguistically motivated relationships should not be mistaken for being “predictable”, considering that languages are motivated in different ways (Boers & Lindstromberg, 2006). Boers (2011) points out that exploiting the linguistic motivation of language in the L2 classroom does not guarantee that L2 learners are equipped with a “foolproof tool” to autonomously work out metaphorical meanings and usages. Littlemore (2023) stresses that “our ability to understand linguistic metaphors (when they are first encountered) may rely on the successful identification of a relevant conceptual metaphor, at other times it may not” (p. 130). These types of observations are critical when designing pedagogical practices, as noted by Piquer-Píriz and Boers (2019).

Further, Boers (2011) notes that the goal of “educated guesses” is not to predict the meanings of metaphors but to aid learners in establishing meaningful connections between the basic and extended senses to provide some support to deal with metaphor uses (p. 240). As Littlemore (2023) highlights, “[w]hen language is viewed through the lens of cognitive linguistics, its motivated elements become much more apparent” (p. 265).

However, Boers (2004) argues that metaphor-awareness raising activities, while useful in helping L2 learners comprehend expressions, may be more limited in their effectiveness for generating metaphor production. In addition, Low (2008) notes that simply teaching learners about metaphor does not, per se, ensure long-term memory retention of the form or meaning,

nor does it necessarily enhance their productive use. Previous research conducted by Philip (2011) has shown that CMs may not effectively aid learners in acquiring the phraseological patterns that accompany metaphors, which are crucial for determining their meaning.

Building on this, Littlemore (2023) stresses that while linguistic motivation may assist L2 learners in understanding language input, it does not guarantee their ability to produce "appropriate target language forms", potentially leading to unconventional metaphors (p. 224). She points out that while some metaphorical extensions are widely accepted and understood in the language, other seemingly similar extensions may not be as effective or conventional, despite a similar underlying metaphorical concept.

For example, in English, the term *digest* can metaphorically describe the process of understanding and assimilating information, as in the phrase "I need time to *digest* this information". This metaphorical usage is explainable through the CM IDEAS ARE FOOD, drawing a parallel between the physical process of digesting food and the mental process of breaking down and absorbing information. However, extending this metaphor to other similar bodily processes, such as *chew*, results in expressions like "I need to *chew* on this information". While still understandable due to the concept of processing, this phrase is less conventional and might not be as commonly accepted or intuitive as "digest this information". This example highlights the variability in the effectiveness and conventional acceptance of metaphorical language, demonstrating the limitations of linguistic motivation in L2 metaphor production.

Cameron and Deignan (2006) argue that metaphoremes encode three types of information: linguistic (lexico-grammatical patterns), conceptual (CMs), and pragmatic (usage), which entails that learners are faced a complex task when using metaphor in the L2. Littlemore (2023) further comments on this, stating, "conceptual metaphor theory could help them with the second of these aims, but in terms of effective communication, this is arguably the least important of the three" (p. 137). However, the extent to which raising metaphor awareness can assist L2 learners in dealing with the dynamic nature of metaphor, including interactive settings, remains an open question.

2.2.2.2. CL-Inspired Instruction: Pedagogical Implementations

Considering metaphor as a prevalent dimension of language use that enhances L2 proficiency (Littlemore & Low, 2006a), it becomes evident that metaphorical language should receive special attention in teaching materials and pedagogical approaches for L2 instruction. However, its concept and pedagogical potential seem to be underrepresented in EFL classroom practice (MacArthur, 2017). While coursebooks from beginner to proficient levels incorporate metaphorical language, it is frequently isolated and out of context, conveyed "as anything other than the basis of colourful idiomatic phrases" (Littlemore & Low, 2006b, p. 268).

L2 learners, especially at advanced levels, often encounter metaphors in *Reading* and vocabulary-building activities, including reassembling broken-up collocations, selecting from multiple options, or completing gap-filling exercises. More advanced L2 learners might also find metaphors in specialised dictionaries or training materials designed for self-study, which are often distinct from the mainstream textbooks used in class (MacArthur, 2010). While metaphor is explicitly addressed in certain coursebooks (see Littlemore 2023), its presentation is almost invariably unstructured. This lack of systematicity often results in metaphors being presented as unanalysed chunks meant for memorisation, implicitly directing learners to "learn by heart", potentially impeding deeper understanding of metaphorical language in L2 learning.

Although considerable research has been done on how to teach metaphor (e.g., Boers, 2013; Boers & Lindstromberg, 2006, 2008a; De Knop et al., 2010; Low, 1988; MacArthur, 2010; Piquer-Píriz & Alejo-González, 2020), the important findings from studies exploring the effectiveness of CL-oriented methods appear to have made limited inroads into ELT materials or syllabus design. As MacArthur (2017) highlights, "[t]he role of metaphor […] finds virtually no echo in current models for language teaching and assessment, and is conspicuously absent from the guidelines laid down in the Common European Framework of Reference for Languages (CEFR)" (p. 414). Thus, metaphor is not explicitly included in TEFL curricula.

Scholars in the field of metaphor have advocated for a heightened emphasis on metaphorical language within the CEFR (e.g., Nacey, 2013; Golden, 2021; MacArthur, 2021). Nacey (2017) suggests that the widespread dependence of instructors on the framework might

inhibit the full realisation and efficacy of research findings in L2 classrooms. Many EFL teachers globally might not be fully aware of the thematic relatedness of many conventional metaphors or recognise the metaphorical motivations behind many everyday language expressions encountered by L2 learners. Such an oversight could lead to students missing essential guidance on the role of metaphor in achieving L2 proficiency. In this vein, other reasons have been observed, one of which, noted by Piquer-Píriz and Martín-Gilete (forthcoming), is the need for L2 instructors to undergo training in the applications of CL to L2 instruction.

The limited impact of pedagogically oriented CL research, i.e., transferring theory into practice, has also been attributed to methodological weaknesses in the design of some of the experiments. In his 2004 study, Boers first critically examined the "tentative" findings from previous research on the role of enhanced metaphor awareness in L2 lexical development (p. 228). He highlighted several underexplored areas in metaphor instruction, such as the optimal amount and intensity required, its potential to enhance metaphor production, its effectiveness across different proficiency levels, its impact on individual learner differences, and the types of CMs suitable for teaching. In addition, Low (2008) emphasised the importance of statistical measures in research, stating that "[w]hat is needed now are studies with larger, mixed-level samples, delayed post-tests and where effect sizes are reported" (p. 226).

Boers (2011, 2013) identified additional methodological shortcomings in his review of CL-inspired studies spanning from 1996 to 2010 that evaluated the effectiveness of this approach. For example, he noted that some studies might lack precise testing measures or fail to include comparable teaching interventions between the control and experimental groups, particularly in short-term studies not representative of standard instructional activities. Boers (2011, 2013) also emphasises the need for a larger body of well-structured research, including more fine-tuned longitudinal studies, to rigorously validate the benefits of such an approach to vocabulary acquisition in L2 instruction.

Low (2017) further describes some previous research as being "methodologically unsound, unnecessarily indirect and unlikely to generate viable real-world solutions" (p. 250).

He acknowledges the challenges inherent in conducting classroom-based research, especially with large, randomised samples committed to long-term studies. This is particularly the case when investigating productive skills. For example, when examining written production, Low (2017) points out that it "might well take a long time to generate a change that was large enough to be statistically significant (or lead to a large enough effect size)" (p. 250).

2.3. METAPHOR IN TEFL: CURRENT STATE AND APPLICATIONS OF RESEARCH

Pedagogically oriented metaphor research can focus on the educational *process* or on the *products* of education, or use *evidence* from training and examination materials to make recommendations regarding the educational process and better understand *assessment* practices (MacArthur, 2021). The critical review of research is divided into three subsections, each focusing on a different perspective: the teaching-learning process (Section 2.3.1), learning outcomes (Section 2.3.2), and materials and assessment practices (Section 2.3.3).

2.3.1. Teaching-Learning Process

2.3.1.1. CL-Oriented Teaching Methods Explored

Metaphor researchers tend to prefer semantic and/or etymological elaboration as the main CL-oriented approach explored in L2 vocabulary instruction. Much of the research into metaphor teaching has involved verbal explanation (Condon & Kelly, 2002; Kövecses & Szabó, 1996; Littlemore et al., 2013), conceptual grouping (Beréndi et al., 2008; Boers, 2000b), and guessing strategies (Boers, 2001; Boers et al., 2004a; Skoufaki, 2008; Verspoor & Lowie, 2003; Niemeier, 2017). In contrast, alternative methods such as pictorial elucidation (Boers et al., 2008, 2009; Szczepaniak & Lew, 2011) or TPR (Lindstromberg & Boers, 2005; Saaty, 2016, 2020) have received less attention.

However, these CL-oriented pedagogical techniques are not without their challenges. MacArthur (2017) emphasises the complexities involved in using specific types of CMs to elucidate the meanings of target words, which is crucial for enhancing metaphor learning. For

instance, the study by Condon and Kelly (2002) revealed that CMs used in explaining phrasal verbs, especially those involving somewhat concrete and relatively abstract elements, were less effective in enhancing recall when limited to verbal explanations alone.

While pictures can enhance cognitive engagement among L2 learners, Boers (2011) cautions that "pictures do not directly elucidate words; they elucidate concepts" (pp. 244–245). This suggests that images might not always assist in learning the exact wording or appropriate use of metaphorical phrases. Supporting this, studies by Boers et al. (2008, 2009) have shown that images are not particularly effective in terms of prompting recall of the precise language of an idiom. Additionally, there is a risk that L2 learners may remember the visuals more clearly than the phrases themselves, particularly when the vocabulary is new to them. Yet, when learners are already familiar with the word forms, Szczepaniak and Lew (2011) found that an association with pictures can be mnemonically effective. In such cases, recall of the picture can indeed prompt recall of the corresponding word forms, enhancing vocabulary retention.

Other researchers (Low, 2008; Boers & MacArthur, 2009; MacArthur, 2010) have also broached the challenge of whether to incorporate pictures in teaching metaphorical language. This issue, particularly in the context of idioms, was not primarily with recalling the exact wording, but rather with confusing and inaccurate details the pictures might introduce (MacArthur & Boers, forthcoming). Niemeier (2017) observes that EFL textbooks might include ineffectual illustrations, for instance, actual dogs and cats falling from the sky to represent the idiom "it's raining cats and dogs". While visual aids can be beneficial for learning, such depictions might amuse learners without adequately explaining the origins or meanings of the idioms. These observations collectively highlight the complex nature of using visual aids in L2 vocabulary instruction, particularly regarding metaphorical language.

When it comes to inferring meanings through physical enactment, TPR similarly presents its own set of challenges. It is important to note that not all metaphors lend themselves to physical imitation, as some concepts are too abstract or complex for such representation (Casasanto & Gijssels, 2015; Gibbs, 2021). In a recent study, however, Saaty (2020) demonstrated that enactment-based metaphor awareness can effectively familiarise

L2 learners with the embodied motivations behind the CM LIFE IS A JOURNEY and its metaphoric expressions. Importantly, Saaty's (2020) research moves beyond prior studies, such as Lindstromberg and Boers' (2005) focus on action verbs. Saaty (2020) showed that L2 learners could successfully act out not only action verbs, but also more complex constructions such as nouns, collocations, and lexical phrases.

Furthermore, Piquer-Píriz and Martín-Gilete (forthcoming) observed in their study on fostering lexical knowledge of frequently used particles in phrasal verbs with secondary students at B1 level, that teenage learners might feel hesitant or self-conscious about engaging in physical activities associated with TPR. These observations underscore the need for incorporating a diverse range of teaching methods that can address the challenges presented by different CL-oriented pedagogical techniques but also cater to different preferences in the L2 classroom.

A holistic approach to teaching metaphorical language, employing various CL-oriented teaching methods, could enhance metaphor awareness among L2 learners more effectively. As MacArthur (2010) suggests, "[v]isual illustrations and physical enactment may thus be used regularly in the classroom to support the verbal explanations of how concrete scenarios may motivate metaphorical uses of words and phrases" (p. 166). Littlemore (2023) further highlights the importance of combining methods by acknowledging that, for example, TPR may not be a "communicative" approach to L2 instruction, "as the learners are never really involved in genuine communication with one another" (p. 199).

By integrating these methods, there is potential to steer research towards identifying techniques that are most effective in helping learners master metaphorical meanings, thereby enriching the overall learning experience. This approach is particularly important when considering the diverse learning styles present in a classroom setting. In this context, Boers (2004) hypothesised:

> [A]n enhanced metaphor awareness will probably be most beneficial to learners with an analytic and imager cognitive style. Analytics are probably most capable of recognising the metaphoric nature of an expression or the figurative use of a

> polysemous word by comparing it with a distinct source domain or literal usage. The identification of distinct source domains (or metaphoric themes) behind sets of expressions provides a framework of vocabulary organisation, which is known to facilitate memory storage. Imagers are probably most capable of associating a novel figurative expression with a mental picture or concrete scene, and imaging and concreteness have also been shown to facilitate retention. (Boers, 2004, p. 224)

To illustrate the implications of method integration, Piquer-Píriz and Martín-Gilete (forthcoming) found that developing CL-inspired activities using distinct teaching techniques separately seemed to impede the students' learning process. The teaching methods explored — including TPR, motivation of source domains, and pictorial elucidation — were not integrated during the execution of the CL-inspired activities at any instructional stage. L2 instructors noted that this distinction between techniques might have shifted students' focus from the content to the methodology of learning. Feedback from the classroom revealed that students occasionally felt disoriented during the sessions, not because of the intricacy of the content, but due to the decontextualised nature of the activities. While the CL-inspired activities were based on the EFL syllabus, they were not incorporated into the textbook's lesson plans.

These observations underscore the importance of contextualising CL-oriented methods within the overall instructional framework, ensuring that they are seamlessly integrated into regular activities aligning with the official curriculum. Such contextualisation seems to be vital for enhancing the effectiveness of these pedagogical strategies.

2.3.1.2. Type of Metaphorical Language Researched

Research on metaphors in EFL is primarily based on two sources of evidence: polysemy and specific conventional metaphorical expressions, including idioms and phrasal verbs. Numerous studies have attempted to exploit the notion of linguistic motivation to teach various vocabulary types, generally yielding favourable results (see Boers, 2013). Yet, the research emphasis appears varied, with certain areas receiving more attention than others.

While there is a significant research focus on metaphorical expressions such as idioms (Boers, 2001; Beréndi et al., 2008; Boers et al., 2004a, 2004b, 2008, 2009; Eyckmans & Lindstromberg, 2016; Guo, 2007; Li, 2009; MacArthur & Boers, in press; Skoufaki, 2008; Vasiljevic, 2011) and phrasal verbs (Alejo-González, 2010; Boers, 2000b; Condon & Kelly, 2002; Condon, 2008; Kövecses & Szabó 1996; Li, 2002; Rudzka-Ostyn, 2003; Yasuda, 2010; Millar, 2023), polysemy seems to be a less explored territory (Deignan et al., 1997; Boers, 2000b; MacArthur & Littlemore, 2008; MacArthur & Piquer-Píriz, 2007; Morimoto & Lowen, 2007; Piquer-Píriz, 2008a; Ponterotto, 1994; Verspoor & Lowie, 2003).

As previously discussed, Saaty (2016, 2020) conducted a study examining the effectiveness of teaching metaphorical expressions (i.e., polysemes) through enactment-based awareness, specifically focusing on the CM LIFE IS A JOURNEY, as opposed to conceptual metaphor awareness and semantic clustering methods. This five-week study, involving 60 female Saudi EFL university students at the B2 level, explored their comprehension, recall, and productive use of pre-taught metaphorical expressions in writing, standing apart from regular classroom activities. A key finding was that enactment-based teaching more effectively enhanced the learners' comprehension and retention of metaphorical expressions than conceptual metaphor awareness alone. However, despite these gains in comprehension and production, the study observed that learners could effectively express themselves without necessarily using the taught metaphorical expressions. The study echoes the broader challenges identified in metaphor teaching in instructed L2 settings, as noted by Boers (2004) and Low (2008). These challenges include the observation that while metaphor awareness-raising activities can aid in comprehension and retention, they may not sufficiently promote metaphor production.

CL-oriented methods for teaching diverse types of metaphorical language have not been uniformly explored, with research focusing on specific, isolated aspects of language via CL-inspired activities that are independently designed and implemented (e.g., Piquer-Píriz & Martín-Gilete, forthcoming). While this research approach allows for in-depth exploration of specific language elements, it also presents some limitations such as risk of overlooking the

interconnectedness and complexity of metaphorical expressions in natural language use and lack of integration into broader instructional context. Consequently, this limited scope of inquiry might have constrained the broader understanding of metaphor in language, as it might not sufficiently prepare L2 learners to understand and use metaphors across various contexts.

Metaphor scholars such as MacArthur (2010) and Boers (2013) advocate for in-context, distributed learning of metaphors in CL-oriented teaching interventions. MacArthur (2010) proposes a more holistic, awareness-raising approach in teaching metaphors, transitioning from a focus on specific teaching techniques to a broader appreciation of metaphor's pervasive role in language:

> In this approach, the specific pedagogical techniques employed […] are seen to be less significant than the general foregrounding of metaphor and the effect(s) this may have on learners' growing awareness of how metaphor permeates language (their own and the L2 to be learnt) as reflected in the growing felicity of the metaphorical language used in their written work. (MacArthur, 2010, p. 157)

Boers (2013) calls for a more dynamic, learner-centred approach to metaphor learning in L2 instruction. He advocates for a shift in focus in pedagogically oriented metaphor research following the "teach as the need arises" approach, strategically incorporating a judicious selection of target language throughout the EFL syllabus:

> It would also make the approach more faithful to the usage-based nature of language. It is remarkable that in the majority of the studies […], the target vocabulary was poorly contextualised, so the input was lacking in cues regarding common usage patterns: just the kind of cues that fuel L1 acquisition. (Boers, 2013, p. 217)

Extending this idea, Boers (2022) recommends training L2 learners to become "independent collectors" of metaphors. The goal is not just to raise awareness of the importance of metaphors but also to instil a consistent habit of noticing metaphorical language by creating learning conditions that enhance the uptake of lexical items from input texts, an important consideration given the often limited time available in educational contexts for implementing L2 training programmes. This approach aligns with current communicative

models in L2 pedagogy, which prioritise learner engagement and promote learner autonomy over the learning process as well as real-world application of language skills (e.g., the CEFR; Council of Europe, 2020).

2.3.1.3. Timing of Studies

It is reasonable to assume that for L2 learners to achieve enduring changes in their mastery of metaphorical language, the teaching approach should be sustained over time, akin to a full-term EFL course. However, while there are some notable exceptions (Condon, 2008; Gao & Meng, 2010; Radić-Bojanić, 2013; Shokouhi & Isazadeh, 2009; Piquer-Píriz & Martín-Gilete, forthcoming), empirical research on the effectiveness of CL-inspired language pedagogy often overlooks long-term effects (see Boers, 2013).

The prevailing trend is towards short-term experiments, which primarily report the immediate benefits of raising metaphor awareness in one-off interventions. These studies particularly highlight the retention of a small number of pre-selected metaphors taught outside the range of regular classroom activities (Beréndi et al., 2008; Boers, 2000b; Boers et al., 2007; Chen & Lai, 2011; Guo, 2007; Herrera & White, 2000; Yasuda, 2010). Littlemore (2023) observes that “the most lasting learning effects are most likely to result when explicit knowledge is brought to be on implicit knowledge and vice versa” and suggest that “learners need to see evidence of it in authentic discourse and test out their own hypothesis in as natural as setting as possible” (p. 225). However, it is unknown whether raising metaphor awareness influences learners’ approach to mastering metaphorical meanings and their success beyond the materials used in the experiments. It remains to be studied if this insight can be transferred to new metaphorical expressions in subsequent applications and/or future contexts.

Boers (2004) suggests that explicit instruction of metaphors over an extended period can be particularly effective in fostering productive vocabulary knowledge. However, when learners apply their metaphor awareness to independently generate figurative expressions, the outcomes can be unexpected, often leading to “marked language”, as highlighted by Boers

(2004, p. 217). He notes that this deviation may result in non-standard or unconventional forms, a consequence of the unpredictable nature of motivated language.

Although fostering independent metaphor usage among L2 learners may risk erroneous direct translations due to L1 transfer, Boers (2004) argues that recognising shared CMs in both the L1 and L2 can expedite the learning process. This phenomenon, also observed by Littlemore (2023), poses challenges in achieving proficiency in an L2. Adherence to conventional language norms is often required for mastery, especially in formal training aiming at native-like linguistic accuracy as opposed to intercultural communication, or testing environments where deviations might be deemed inappropriate.

In a pilot experiment conducted with English teachers at the University of Birmingham, Boers (2004) highlights that native speakers seem to be more tolerant of (written) unconventional metaphorical expressions derived from an established CM, compared to deviant versions of idioms. Drawing on this apparent fair degree of tolerance demonstrated by the native speakers attested in the pilot experiment reported above, Boers (2004) concludes that “it seems that it may be feasible to take metaphor awareness beyond the realm of receptive vocabulary learning, after all” (p. 220). However, it remains uncertain whether the same degree of tolerance by L1 speakers extends to unconventional metaphorical expressions generally as opposed to incorrectly used idioms. Evidently, L1 speaker feedback on unconventional metaphorical usages is provided by assessors in standard L2 competence assessment, as will be seen.

Along with the degree of monitoring required to effectively foster metaphor awareness, another concern may be the amount of class time invested to obtain long-term effects. In this context, MacArthur (2010) reiterates the significance of sustained efforts in awareness-raising in L2 classroom practice to ensure a lasting impact. As an example, Piquer-Píriz and Martín-Gilete (forthcoming) determined that a three-month CL-oriented instructional intervention was insufficient to establish teaching and learning routines fostering cognitive engagement in the teaching interventions, which were quite time-consuming. Despite having incorporated

distributed learning from a longitudinal perspective, learners only received CL-style instruction for approximately 10 hours in the three months.

Previous empirical studies appear not to have sufficiently addressed the long-term effects of the CL-oriented instructional interventions on broader metaphor instruction. This includes exploring how increased metaphor awareness might influence L2 learners' organic use of metaphors contributing to their lexical development, and how these potential learning outcomes could enhance their productive skills, namely, *Speaking* and *Writing*. Consequently, a more comprehensive understanding might be derived from long-term, sustained research conducted in the real L2 classroom.

2.3.1.4. Participant Profile

While teaching metaphor is of importance for L2 instruction, there has been an unequal distribution of empirical studies conducted in pedagogically oriented metaphor research, with varying emphasis on different ages and educational stages. Most studies have targeted learners at the higher education level or similar educational contexts (Beréndi et al., 2008; Boers & Demecheleer, 2001; Boers et al., 2008; Kamberi, 2014; Lindstromberg & Boers, 2005; Radić-Bojanić, 2013; Skoufaki, 2008; Szczepaniak & Lew, 2011). University-age students have been frequently the participants in these studies, perhaps because they are readily accessible to researchers.

In contrast, secondary education students (Golden, 2010; Meissner, 2010; Niemeier, 2017) and young learners (Piquer-Píriz, 2008a, 2010; Pan, 2019) have received less attention. The underrepresentation of these age groups may be attributed to several factors, including the additional ethical considerations associated with conducting research with younger populations, logistical challenges of working with students at early stages, and the general concerns about the perceived cognitive demands of metaphorical language.

In Pan's (2019) study with Chinese learners of English aged 10–12 at A1 level, the regular teacher described the CL-oriented method as particularly demanding. Despite this, the young L2 learners in her study remained motivated, a sentiment echoed in Niemeier's (2017)

research with German learners of English at A2 level studying at lower secondary education. At the tertiary level, Condon (2008) observed that students experienced fatigue when dealing with metaphorical language, particularly towards the end of the day.

These observations are consistent with recent research, which suggests that implementing CL-inspired instruction can be both challenging and time-consuming (Martín-Gilete, 2022a; Piquer-Píriz & Martín-Gilete, forthcoming). Factors such as implementation timing, class timetables, student-teacher ratios, duration of activities, and student attitudes towards the activities were observed to be crucial for the successful integration of CL-inspired instruction. These researchers also draw attention to concerns such as fatigue and diminished sustained attention, issues affecting not only students but teachers as well.

To successfully conduct the CL-oriented activities, which often consume a significant portion of class time, the instructor's guidance is essential. This implies that the procedure can be exhausting for teachers too. Yet, this approach is often seen as appealing to L2 learners, including teenagers, and is highly valued by teachers in actual TEFL settings (see Piquer-Píriz & Martín-Gilete, forthcoming). As MacArthur (2010) suggests, CL-inspired instruction is likely to promote a more engaging perspective on L2 instructed acquisition, potentially revitalising students' interest and enhancing their enjoyment of the learning process.

Regarding the level of English proficiency, the empirical studies have included participants from all levels as defined by the CEFR. Yet, beginners' levels, specifically A1 (Piquer-Píriz, 2004) and A2 (Niemeier, 2017) as well as advanced levels, such as C1 (Eyckmans & Lindstromberg, 2016) and C2 (Boers et al., 2004b), have been comparatively less explored. The intermediate level, particularly B1, has garnered the most attention (Boers, 2000b; Condon, 2008; Gao & Meng, 2010; Littlemore, 2004; Littlemore et al., 2011; Pourdana et al., 2014). However, pedagogically oriented metaphor research at the B2 level within the context of TEFL remains less extensive (e.g., Martín-Gilete, 2022b; Littlemore et al., 2014).

The scarcity of research focusing on B2 level learners, particularly regarding metaphor production, underscores the need for enhanced support at this critical learning stage. Littlemore et al. (2014) investigated metaphor production in adult Greek and German learners

of English who had successfully passed *Cambridge* ESOL exams across CERF levels A2 to C2, as already mentioned. They noted that at B2 level, where L2 learners are expected to express more sophisticated thoughts, there is an increase in metaphor usage, particularly in tasks that require the expression of opinions and reflections on abstract topics. Additionally, the study found that the role of metaphors at B2 level, in relation to the CEFR standards for writing, becomes more complex, particularly in persuasive academic arguments as learners' proficiency advances. It should be recalled, of course, that Littlemore et al.'s (2014) participants had not received any metaphor instruction, as a far as it could be ascertained.

The study, which examined texts on topics such as politics and government for thematic consistency, noted a significant use of open-class metaphors over closed-class ones (mainly prepositions) at the B2 level. This pattern signifies a qualitative shift in metaphor usage as learners advance in proficiency, especially beyond the upper-intermediate level. However, with this increase comes challenges, including a rise in metaphor-related errors and the influence of learners' L1, underscoring the need for improved metaphor training at B2 level. Littlemore et al. (2014) found that while B2 learners employ metaphors to articulate complex ideas, their use of metaphors often strays from conventional standards, which can adversely affect communication. This insight underlines the importance of including explicit metaphor instruction in the formal training of learners at the B2 level.

Regarding participants' L1 backgrounds, although previous studies have showcased a diverse range of L1s, the primary focus has been on Dutch-speaking participants (Boers, 2000). Nonetheless, other studies have involved Chinese (Gao & Meng, 2010); Hungarian (Beréndi et al., 2008); Spanish (Piquer-Píriz, 2008a, 2010); Persian (Pourdana et al., 2014); French (Boers & Demecheleer, 2001); Japanese (Yasuda, 2010); Norwegian (Nacey, 2013); and Greek (Skoufaki, 2008).

In terms of the number of participants, most research studies have used small-scale groups (Boers & Demecheleer, 2001; Eyckmans & Lindstromberg, 2016; MacArthur & Littlemore, 2008; Verspoor & Lowie, 2003), as opposed to large-scale groups of participants (Boers et al., 2004a; Li, 2002). Many of these studies have not consistently included both

control and experimental groups. However, among those that have, most employed two groups: one control and one experimental (Beréndi et al., 2008; Boers, 2001; Lindstromberg & Boers, 2005; Yasuda, 2010). Fewer studies have used three groups, with one control and two experimental groups (Hashemian, 2013; Saaty, 2016, 2020).

The trend towards such studies can be seen as a pragmatic response to the challenges inherent in conducting classroom-based research, especially in the context of L2 instruction (Low, 2017). Factors such as resource limitations to large numbers of participants, which also pose logistical challenges, can lead researchers to prefer smaller-scale studies with fewer groups, as they offer a balance of depth, control, and practicality, which, in turn, is essential for classroom-based research.

However, in many of these research studies, there is a lack of parity in teaching interventions between control and experimental groups. The activities for the control group were not particularly motivating and often relied on rote learning from word lists (MacArthur, 2017). As Boers (2013) argues, any improved retention observed in learners taught with awareness-raising exercises might be due, in part, to the presentation of vocabulary in smaller chunks to the experimental groups, as opposed to the control groups who are typically given long, undifferentiated lists of vocabulary, resulting in potentially unpair comparisons.

As discussed above, much of the empirical research evaluating the effectiveness of explicit metaphor instruction has focused on a specific participant profile: undergraduate students who are Dutch-speaking learners of English with a B1 level of English proficiency. These participants have often featured in small-scale studies, typically divided into control groups (using a memorising/translation approach) and experimental groups (using a CL-inspired approach).

As a result, this consistent focus highlights gaps in our understanding of metaphor instruction across various contexts. Arguably, different profiles of L2 learners might encounter varied challenges with metaphors, especially in authentic classroom settings. This is particularly true for those whose L1 has significant linguistic or cultural differences from the L2. These differences may require alternative teaching approaches and can significantly

influence both quantitative and qualitative results. Yet, Boers (2004) emphasises the need for caution in treating a group of L2 learners as a homogeneous population. He notes that, even if all participants share the same level of proficiency, raising metaphor awareness may not be equally effective for everyone due to various affective and cognitive factors.

2.3.2. Outcomes of Instruction

2.3.2.1. Research Focus on Exploring Metaphor Learning Outcomes

The central role of metaphor in communicative competence underscores the need for metaphor research to address both receptive (comprehension) and generative (production) skills, which are crucial for reaching proficiency in an L2 (Littlemore & Low, 2006a). Indeed, raising learners' awareness of metaphorical aspects of language has been shown to benefit their MC, enhancing both their comprehension and production, as discussed below.

Yet, much of the pedagogically oriented metaphor research has mostly focused on the benefits of applying the CL notion of motivation to enhance comprehension and retention of conventional English metaphors (Beréndi et al., 2008; Boers, 2000a; Boers & Demecheleer, 2001; Danesi, 1992; Golden, 2010; Guo, 2007; Lindstromberg & Boers, 2005; Littlemore, 2004; MacArthur & Littlemore, 2008; Piquer-Píriz, 2020; Verspoor & Lowie, 2003), as well as exploring the interplay between metaphor comprehension and its usage (Alsadi, 2016; Boers, 2001; Charteris-Black, 2002; Littlemore, 2001b; Littlemore, 2010; MacArthur & Littlemore, 2011; Morimoto & Loewen, 2007; Radić-Bojanić, 2013; Saaty, 2016, 2020).

O'Reilly and Marsden (2021, 2023) developed an extensive set of tests to elicit and reliably measure metaphor-related skills and competences, drawing on the constructs of L2 MC as defined by Low's (1988) and Littlemore and Low's (2006a) work. Their 2021 article involved the development and analysis of this large, theory-driven battery of L2 MC tests, which was completed by 112 L1 Mandarin learners of English and 31 native English speakers.

In their 2023 article, O'Reilly and Marsden extended their previous work (2021) on MC by applying the MC Test battery in a study with 108 Mandarin learners of English. This

research aimed to explore the relationship between MC, vocabulary knowledge, and overall L2 proficiency. Their correlation analyses revealed significant positive relationships between specific MC construct measures and two high-stakes proficiency measures: the Oxford Online Placement Test (OOPT) and the International English Language Testing System (IELTS). These relationships were stronger in overall comprehension (*Reading*) compared to overall production (*Writing*). Although a clear association was found between vocabulary depth and MC, the role of MC in higher-stakes testing contexts was less clear. The study's findings highlight the expected metaphor comprehension and production abilities at various CEFR levels. As anticipated in Chapter 1, learners at the B2 level and below showed moderate accuracy in recognising metaphors but faced challenges in producing them (O'Reilly & Marsden, 2023, p. 31). These observations provide research-based insights into the infrequent references to metaphor in major language descriptors such as the CEFR.

In contrast, the independent usage of metaphor has received less attention. Research focusing on L2 metaphor production has primarily explored written expression, often observing spontaneous responses from L2 learners to tasks (Alejo-González, 2010; Kathpalia & Carmel, 2011; Kövecses & Szabó, 1996; Nacey, 2013). For example, Nacey (2013) analysed metaphor usage in argumentative essays on "technology and industrialisation" written by undergraduate Norwegian learners of English, predominantly at CEFR levels C1 and C2. Her research revealed that these non-native speakers produced a greater number of metaphors compared to native speakers, with a density of 15.50% vs. 13.30%. A notable finding was the prominence of prepositions as the primary metaphorical word class. Nacey observed that learners often made efforts to employ more complex language as a means to excel in their writing (see also Nacey, 2010; Nacey & Uri Jensen, 2019).

Kathpalia and Carmel (2011) analysed linguistic metaphors in 113 written essays produced by Singaporean learners of L2 English, focusing on their relation to grammatical, illocutionary, textual, and sociolinguistic competences as defined by Littlemore and Low (2006a). Their findings reveal that while these learners frequently attempted to use metaphor in their writing, they often lacked the appropriate, pre-fabricated language structures for

effective metaphor expression. This observation raises important questions about the focus of L2 metaphor instruction: whether it should prioritise teaching specific target forms of linguistic metaphor that are directly tested, or emphasise enhancing the natural, broad use of metaphors, which, however, might lead to unconventional metaphorical expressions (Boers, 2004). In this context, O'Reilly and Marsden (2021) argue that if communicative success, rather than adherence to a constrained set of pre-fabricated language, is the primary aim, then the ability to use metaphor should not rely solely on a fixed set of metaphors.

Recent research has adopted a longitudinal perspective to examine how learners actually use metaphor in L2 contexts. As previously discussed, this approach has yielded empirical data about the frequency and types of metaphor use across various CEFR levels and educational stages including learners from different L1 backgrounds, yet predominantly in written discourse (Cuberos et al., 2019; Hoang & Boers, 2018; Littlemore et al., 2014; Nacey, 2020, 2022).

In her longitudinal study, Nacey (2022) explored the development of metaphor usage among teenage Norwegian L2 learners of English. The study tracked five secondary school students aged 13–17 over four years, analysing 40 authentic end-of-semester written exams from the TRAWL corpus (Dirdal et al., 2022), with 10 compositions per grade level and two per student, covering a variety of topics. Findings revealed that despite advancing through different school grades and improving L2 proficiency, there was no significant variation in the rates of increase between open-class and closed-class metaphors. Nacey (2022) inferred that this might suggest "these learners have not reached the B2 level" (p. 294). These insights offer valuable perspectives on individual language learning trajectories over time.

Furthermore, Cuberos et al. (2019) investigated metaphor production by native and non-native speakers of Spanish. Their study explored the oral and written expository and narrative texts by L1 Arabic and L1 Chinese learners of Spanish across three age groups: grade school, junior high, and university students. While focusing primarily on deliberate metaphor, they also observed that metaphor use in L2 discourse tends to develop with age.

This longitudinal approach moves beyond the earlier qualitative insights derived from L2 classroom practice, aiming to support and enhance metaphor learning. For example, MacArthur (2010) explored metaphor usage by L2 learners in their written work, focusing on the need for metaphors in discussing abstract topics at more advanced levels. She also emphasised the necessity of user-friendly ways of talking about metaphor in the EFL classroom and the importance of providing corrective feedback on learners' metaphor production, while also considering cross-linguistic and cross-cultural aspects.

However, despite this evolution in metaphor research approach, as MacArthur (2021) remarks, scholars often focus on "analyses of the written texts produced at the end of a learning period in exams or other types of assessment" (p. 353), indicating a continued emphasis on written production.

Recent research has expanded its scope to include L2 metaphor production in speech, yet particularly focusing on ELF academic settings. These studies primarily analyse density, specific linguistic aspects, and the discursive functions of metaphor in academic discourse in higher education, especially in contexts where English serves as a Medium of Instruction (EMI). In their comprehensive review, MacArthur and Alejo-González (2024) identify key research areas including monologic academic talk in lectures (Littlemore et al., 2011; Low, 2010; Low et al., 2008), dialogic interactions in academic mentoring such as office-hours consultations from the EuroCoAT corpus (Alejo-González, 2022; MacArthur, 2016a, 2016b, 2020; MacArthur et al., 2015) and discussions seminars from the MetCLIL corpus, where metaphors emerge in various interactions between lectures and students and among peers (Alejo-González, 2024; Castellano-Risco et al., 2023; Fielden-Burns & Piquer-Píriz, 2022; Littlemore & Fielden-Burns, 2023; Krennmayr et al., 2022; Philip, 2023). Research conducted on these corpora reveals that in spoken academic contexts, prepositions are the metaphor type most frequently used (e.g., Alejo-González, 2022, 2024). This finding underscores the importance of this word class in facilitating academic oral interaction, actually a not surprising finding considering that spatial prepositions are highly polysemous (Lindstromberg, 2009).

Previous research has shown the considerable use of metaphors by L2 speakers in ELF academic settings. Alejo-González (2024) revealed that metaphor density in interactive discussion seminars was 14.22%, exceeding the 6.80% observed in L1 English conversation by Kaal (2012) but less than the 17.50% found in L1 English academic written discourse by Herrmann (2013). Additionally, the density observed in discussion seminars is also greater than that found in another interactive oral genre involving L2 speakers, office-hours consultations, with a density of 11.90% as reported by Alejo-González (2022).

Furthermore, both discussion seminars and office-hours consultations showed that prepositions (i.e., closed-class items) were the most frequently used type of metaphor during face-to-face interactions, whether with L1 speakers or among L2 peers. The analysis of the EuroCoAT and MetCLIL corpora indicates a high degree of overlap in the use of metaphorically used prepositions (*in*, *about*, *on*, *with*, *from*, *at*, *by*) in these settings, indicating a foundational set of metaphorical prepositions in academic oral communication. The findings also suggest that the interactive nature of the setting may influence metaphor usage. However, research on L2 spoken metaphor production in broader contexts, such as EFL instruction and its formal assessment, remains largely unexplored.

Littlemore and Low (2006a) provided insight into this notable imbalance in metaphor research (comprehension vs. production), suggesting that L2 learners “probably need to understand metaphor more often than they need to produce it” (p. 46). Yet, L2 proficiency is not solely about receiving and processing information (input); it also crucially depends on the ability for effective expression (output). The importance of balancing these aspects becomes particularly evident in testing environments, where both productive (*Speaking* and *Writing*) and comprehension skills (*Listening* and *Reading*) are evaluated equally (see UCLES, 2019a).

Understanding how L2 learners employ metaphors to reach the required proficiency levels in both oral and written discourse is essential. This importance stems not only from its contribution to overall L2 proficiency but also from its potential impact on performance in standard L2 competence exams (Littlemore & Low, 2006a). Littlemore et al. (2014) highlight that L2 learners often produce metaphors under communicative pressure, especially in testing

environments. In such contexts, the use of metaphors can be either necessary or advantageous for L2 learners. Nonetheless, MacArthur (2021) raises a critical point, stating that "learners use metaphor in the languages they have learned, but it is not clear whether this is a consequence of the educational process they have undergone or quite independent of it" (p. 356). This observation highlights the need for more research into how explicit metaphor instruction in L2 classrooms influences learners' use of metaphors, particularly in enhancing their speech and writing performance.

While there has been a robust focus on the teaching-learning process in metaphor research, as previously discussed in Section 2.3.1 of this chapter, the exploration of learner discourse has focused on metaphor appearing spontaneously in response to task, rather than emerging as a result of any pedagogical focus on metaphor awareness (see Piquer-Píriz & Alejo-González, 2020; and Littlemore 2023). Despite advancements in our understanding of metaphor usage, the complete scope of L2 metaphor production, particularly as it develops in tandem with L2 proficiency, remains underexplored. This gap is significant, particularly in considering whether and how enhancing metaphor awareness in L2 classrooms directly contributes to improved levels of L2 proficiency. This is especially pertinent in instructed L2 settings and testing environments, which not only include written assessments but also oral examinations. Understanding the impact of metaphor instruction in these varied contexts is crucial for a holistic view of L2 proficiency development.

2.3.2.2. Testing Learning Outcomes: A View on L2 Metaphor Production

Previous research has underscored the importance of using *naturalistic* data when examining L2 metaphor production. This approach facilitates the analysis of authentic and spontaneous language use. Studies such as those conducted by Hoang and Boers (2018), Littlemore et al. (2014), and Nacey (2013, 2020, 2022) show how focusing on naturalistic data, mostly obtained from written assignments, can shed light on how metaphors are used in real-life contexts, thus offering a more genuine representation of L2 metaphor production.

In contrast to this approach, *elicitation methods* using comprehension/production tests and experimental stimuli (e.g., Aleshtar & Dowlatabadi, 2014; Azuma, 2005; Castellano-Risco & Piquer-Píriz, 2020; Littlemore, 2001a; O'Reilly & Marsden, 2021, 2023) offer a different perspective. These methods typically involve structured tasks or specific prompts designed to elicit metaphor usage, facilitating a controlled analysis of particular aspects of metaphor use. While they provide a focused view on specific linguistic phenomena, such elicitation methods may not capture the full range and spontaneity of organic use of metaphors found in natural conversations and writings.

Studies using naturalistic data often apply specific metaphor identification procedures, such as MIP (Pragglejaz Group, 2007) and MIPVU (Steen et al., 2010). However, it is crucial to acknowledge the limitations of data naturally produced in regular classroom activities, as it may not fully represent the breadth of L2 learners' metaphor production abilities. For instance, if a learner fails to produce a specific metaphor or type of metaphor, it is unclear whether this is due to a lack of understanding or simply because they had no opportunity or inclination to use it. Furthermore, as noted by O'Reilly and Marsden (2021), distinguishing clearly between naturalistic and elicited data can be challenging, since both methods can overlap in situations such as "in a role-play job interview eliciting spontaneous productions or a written essay eliciting pre-prepared content" (p. 27). These scenarios are similar to those used in standard L2 competence assessment, for example, employing both prompted discussions and essays.

How metaphor production can be tested as part of general L2 proficiency and vocabulary knowledge remains problematic. While Bachman and Palmer's (1996) model of communicative competence has been widely adopted as a basis for designing L2 competence assessment exams, CL-oriented instructional interventions often employ specific methods designed by metaphor scholars to test L2 learners' metaphor knowledge (Boers, 2013). These ad-hoc tests stand apart from conventional L2 testing methods that assess productive skills through free, spontaneous responses within a controlled environment. Instead, the emphasis often shifts to activities such as gap-filling or multiple-choice clozes, which may not adequately

gauge whether L2 learners have truly mastered the nuances of metaphorical meanings or not necessarily assess control of form, which is a very important aspect of L2 assessment.

Low (2008) argues that "forced-choice and even constrained-response tests have been shown to overestimate learning in key areas of language […], so we might assume that metaphoric competence is best tested by some forms of free-response direct communicative test" (p. 222). This perspective highlights that conventional assessment formats may not be the most suitable means to measure learning gains, as they do not align with the instructional methods in TEFL settings, where the communicative approach appears to be widely used.

Piquer-Píriz and Martín-Gilete's (forthcoming) study addresses the discrepancy between statistical significance and observed learning outcomes, especially in the context of L2 metaphor production. Despite L2 learners being exposed to CL-inspired instruction, they were assessed using a formal gap-fill test with decontextualised sentences requiring the target vocabulary. The quantitative results were disappointing, suggesting a misalignment between the testing method and teaching approach, resulting in no apparent learning gains. This misalignment might stem from a broader discrepancy between CL principles and the standard practices outlined in official EFL syllabi. The study raised concerns that the gap-fill test might not have effectively gauged the intended learning outcomes, given its lack of adaptation to classroom-based research.

Echoing this sentiment, Martín-Gascón et al. (2023) highlight the importance of aligning assessment methods with the instructional approaches used in L2 instructed settings. They emphasise the need for "adapted tests" to effectively measure CL-oriented practices, addressing the shortfall in validated tests for ACL. The authors advocate for using CL-based tests over traditional quantitative testing techniques. Indeed, their research demonstrates that pairing a CL-oriented instructional approach with a matching CL-inspired assessment design (i.e., elicitation method) results in greater learning outcomes compared to the more pervasive notional-functional approaches.

In my view, analysing authentic, spontaneous oral and written outputs in controlled testing environments — especially those linked to the CEFR descriptors such as official

practice tests — may provide an understanding of L2 organic metaphor use in *Speaking* and *Writing*. This approach aligns closely with the EFL syllabus and the tasks performed in instructional L2 settings (cf. Piquer-Píriz & Martín-Gilete, forthcoming). In this respect, authentic exam papers that cover a wide range of broad topics are essential. They help to mitigate the topic-dependence of language, a key factor influencing the use and variety of metaphors in discourse (Deignan et al., 2013; Golden, 2012, 2021; Semino, 2008). They also highlight the importance of "ecological validity" in findings, reflecting real-world language use in standard L2 assessment contexts, as suggested by Boers (2011, p. 237).

Adopting this alternative approach to traditional gap-filling activities in classroom-based research ensures an assessment method that is congruent with the teaching approach. This not only helps reduce the chance factor but also establishes consistent assessment conditions for both control and experimental groups (cf. Martín-Gascón et al., 2023). This approach offers a more holistic view of L2 metaphor production, thereby enabling a more precise evaluation of the real-world impact of CL-inspired language pedagogy.

2.3.3. Addressing the Neglect: Metaphor in L2 Instruction and Assessment

The challenge of incorporating CL-inspired instruction becomes particularly evident when metaphor-awareness activities are not a standard element of L2 classroom practice. The situation is further complicated by the limited availability of teaching materials designed from a CL perspective, coupled with a scarcity of clear guidance for L2 instructors in official curricula (MacArthur, 2021).

Despite the importance of metaphor in L2 learning and the work by CL linguists, the significant findings from pedagogically oriented metaphor research have largely remained peripheral to official L2 competence descriptors, such as the CEFR. Consequently, there has been little impact on mainstream textbooks and syllabus design, as well as on major examination boards (MacArthur, 2017; O'Reilly, 2017; O'Reilly & Marsden, 2023).

2.3.3.1. Teachability of Metaphors in ELT: Materials and Syllabus Design

Very few attempts have been made to develop specific CL-oriented teaching materials; yet they have focused on the areas most emphasised in metaphorical language research. Boers and Lindstromberg (2009) as well as Lindstromberg and Boers (2008) have proposed techniques specifically for L2 instructors, with a particular focus on teaching idioms and collocations. On the other hand, materials have been specifically tailored for L2 learners. Lazar's (2003) suggestions for CL-inspired instruction in the EFL classroom included lesson plans structured around CMs, uncovering the metaphorical motivation behind the polysemous senses of general metaphorical language. Similarly, Rudzka-Ostyn's (2003) work addressed prepositions and particles, especially in conjunction with verbs, as seen in phrasal verbs.

Since the early 2000s, when these proposals were made, there has been a noticeable lack of new contributions or advancements in the field. This gap is particularly surprising given the growing empirical evidence that underscores the positive effects of CL principles in L2 instruction (see Piquer-Píriz & Alejo-González, 2020). As current exceptions, Llopis-García et al. (2022) emphasises the robust demand for quality teaching materials and provides guidelines on their design, while also highlighting potential educational technology tools that can be employed.

Scholars have also shown interest in analysing mainstream ELT materials to provide recommendations for the teaching-learning process. In the absence of teacher awareness of the CL view of metaphor and/or training, the materials L2 instructors select will become crucial in determining how learners acquire metaphorical language in L2 classrooms. Thus, a particular area of focus within this research has been the teachability of metaphorical language in training materials, assessing their support for enhancing L2 learners' metaphor knowledge.

Studies such as those by Alejo-González et al. (2010) and Amaya-Chávez (2010) suggest that metaphor-related activities are seldom included in coursebooks. In a more recent analysis, Millar (2023) examined the influence of CL on L2 instruction, specifically regarding phrasal verbs in mainstream ELT coursebooks. Her findings suggest that the potential of a CL-oriented approach to exploit the semantic motivations of polysemous meanings and

promote figurative thinking is not fully realised. This is particularly evident in teaching methods and metalinguistic discourse that provide explicit guidance, pointing to a need for improvement in ELT materials, especially at the B2 or upper-intermediate level (but see MacArthur, 2017, for discussion of the treatment of phrasal verbs in recent textbooks).

In their analysis of popular EFL textbooks and sets of supplementary materials designed after 1994, Littlemore and Low (2006a) identified three broad approaches towards the teaching of metaphor in mainstream materials, which continue to be prevalent today (see Littlemore, 2023). The first approach treats metaphor as if it is something unusual, presented in dedicated books on idioms (e.g., McCarthy & O'Dell, 2017a) or phrasal verbs (e.g., McCarthy & O'Dell, 2017b), often as self-study activities. Such materials tend to offer decontextualised activities that seem to test rather than to teach the metaphorical language in question, providing few opportunities for in-depth learning.

The second approach observed by Littlemore and Low (2006a) involves offering supplementary vocabulary teaching material in the form of lessons or units structured around CMs. Along with Lazar's (2003) proposal, there is a set of internet materials on the *Onestop English* site (Clandfield, 2003) offering downloadable lesson plans with a communicative focus. However, even though metaphor is given a more prominent place in L2 instruction through these materials, it remains unincorporated in the primary textbook used in class.

Finally, the third approach involves attempts to incorporate explicit teaching of metaphors within textbook activities. However, these activities are predominantly found in advanced-level materials and are often confined to isolated sections in the textbook. For specific details, see Littlemore and Low (2006a, pp. 208–209). As a result, they remain unintegrated with the broader contents of the units, ultimately presenting metaphor-related work as an auxiliary component of the L2 classroom.

Thus, while metaphor may be incorporated into EFL instruction, it often lacks the appropriate focus. It is typically found in reference works, metaphor-led lesson plans, and incidental exercises. Ideally, the emphasis on teaching metaphor should be prominent in coursebooks themselves as well as L2 classroom practice. Metaphor has not yet been

sufficiently incorporated into L2 classroom practice, indicating a need for a more systematic focus on metaphor within the EFL syllabus (MacArthur, 2010; Littlemore, 2023).

To address this gap, Andreou and Galantomos (2008) proposed the development of a conceptual syllabus "arranged around certain conceptual domains, while their instruction will put an additional emphasis on morphosyntactic and communicative cues" (p. 72). In contrast, Littlemore (2023) suggests designing teaching materials that incorporate a radial category approach. However, as Low (2020) notes, "it would be hard for teachers to integrate […] metaphor teaching into a regular or full syllabus/curriculum, unless publishers created extensive course materials using it, with explanations for teachers" (p. 52). That is, despite the presence of CL-oriented activities in mainstream materials, it would remain essential for instructors to be knowledgeable about how to effectively implement them in regular teaching practices (see further discussion below).

The design of EFL syllabi is not the sole factor influencing the integration of metaphor in L2 instruction. MacArthur (2010) recommends using monolingual dictionaries side by side with bilingual dictionaries to provide "supplementary information about how words are typically used in context" (p. 168). Consistent with this, the *MacMillan English Dictionary for Advanced Learners* (MM) includes special enhanced entries termed "metaphor boxes" alongside specific word entries. These boxes showcase examples of words and phrases representing the target CMs, thereby aiding the understanding and teaching of metaphorical language.

Recent research has explored the type and frequency of metaphors L2 learners are exposed to in textbooks. While Alejo-González and García-Bermejo (2020) focused on Content and Language Integrated Learning (CLIL) environments, other studies have analysed how mainstream materials, especially in the context of official exam preparation, serve as a primary input enhancement for L2 instruction due to the goal-oriented nature of formal training. For instance, Martín-Gilete (2022b) found that material designers expect EFL learners at B2 level to be familiar with a relatively high number of metaphors for exam preparation across oral and written input texts included in topic-based lessons. Despite this, many textbooks do not directly address metaphor, particularly at this level, as previously discussed.

MacArthur (2010) remarks that L2 learners “can be helped to use [metaphor] better, but only if it is given a prominent place in classroom discussion” (p. 159). To further this point, incorporating explicit metaphor instruction into regular EFL syllabi through CL-oriented distributed learning might offer a starting point to support L2 learners in using metaphors. However, due to the lack of clear guidance for instructors in mainstream materials or syllabus design, training them in CL would be beneficial.

Piquer-Píriz and Martín-Gilete (forthcoming) advocate for raising teacher awareness of the CL view of metaphor. They suggest that, while CL-inspired techniques might be appealing, especially compared to sheer memorisation, involving L2 instructors in the process is essential to effectively bring CL-inspired proposals to L2 classroom practice, considering they are time-consuming and require additional efforts. Training EFL teachers, both pre-service and in-service instructors, on how to apply CL principles in L2 instruction under the mentorship of pedagogy-oriented researchers could prove invaluable (see Giessler, 2012). This would better equip instructors to tackle this aspect of language in L2 classrooms, fostering a deeper understanding of metaphor and formulating teaching strategies that enhance learners’ metaphor skills in the L2.

Such teacher training in CL should present clear guidelines on enacting CL-inspired activities, and potentially lesson plans linked to the EFL syllabus. In this light, the adoption of CL-inspired pedagogical practices could become a mainstay in the teacher’s toolkit, aligning with current communicative models in L2 pedagogy, such as the CEFR. As Boers (2022) describes it, these can be part of their “mixed bag of tricks”, not merely additional techniques but an integral and judiciously implemented component of the EFL teaching-learning process.

In this context, research by Martín-Gilete (2022b) provided preliminary insights into the benefits of raising metaphor awareness by an EFL instructor trained in CL, employing the textbook used in class as the primary input source. However, since her study lacked a control group, any observed learning gains, especially those related to metaphor use and topic similarity between learner discourse and textbook input, should be approached with caution.

Yet, comprehensive studies that focus on the benefits of enhancing metaphor awareness and are conducted by CL linguists actively involved in the real L2 classroom remain limited.

2.3.3.2. Metaphor Use in L2 Competence Assessment

The Common European Framework of Reference (CEFR), formulated by the Council of Europe in 2001 and updated in 2020, provides a comprehensive framework for learning, teaching, and assessing an L2. This document includes a series of "can-do" statements, which are descriptors of language ability that define levels of L2 proficiency. These guidelines are crucial for syllabus designers and language testers, as they establish clear benchmarks for evaluating communicative competence, which instructors often use as a basis for their training methods. As highlighted by Nacey (2013), "what is written there — and left unwritten — has immediate consequences for what is supposed to happen in contemporary language classrooms across the continent" (p. 40).

The CEFR breaks down proficiency into a 6-point scale, ranging from A1 level to C2 level. These levels are categorised into three broad thresholds: Basic User (A1 [Beginner] to A2 [Elementary]), Independent User (B1 [Intermediate] to B2 [Upper Intermediate]), and Proficient User (C1 [Advanced] to C2 [Proficiency]). Each level specifies the communicative linguistic abilities expected of an L2 learner in *Listening*, *Speaking*, *Reading*, and *Writing*. The updated edition expanded the framework beyond the traditional four-skill model, categorising communicative activities into four domains: *reception*, *production*, *interaction*, and *mediation*.

Teaching materials and standard L2 competence exams are often developed based on the descriptors outlined in the CEFR, aligning more closely with real-life language use and providing a more nuanced framework for language assessment. For instance, the structure of oral tests in CEFR-linked exams typically includes tasks that differentiate between spoken *production*, which involves delivering a sustained monologue or long turns, and *interaction*, focusing on conversational dialogue or discussions among candidates in short turns. For specific examples and more detailed information, see Appendix H, which discusses the *B2 First for Schools Cambridge English* examination.

While the CEFR provides a robust framework for L2 instruction, it may not comprehensively address every linguistic aspect, particularly metaphorical language. The CEFR seems to regard metaphors as unusual and extraordinary, which contrasts with the contemporary understanding that metaphors are pervasive in both language and thought (Lakoff & Johnson, 1980, 1999), motivating polysemous use of highly frequent vocabulary items. This perspective mirrors that of Bachman and Palmer (1996), who view metaphors primarily as topics for advanced L2 learners, often entwined with oblique cultural references. In the CEFR, metaphors are typically framed as rhetorical embellishments or obstacles, especially at the C1 and C2 levels. However, this stance appears to overlook the observations by Littlemore and Low (2006a) that metaphor usage is a key indicator of L2 learners' ability to function across various proficiency levels as defined by the CEFR.

As MacArthur (2010) notes, even though CL offers important insights for L2 learning, metaphor remains "a relative newcomer in ELT" (p. 162). This might explain why, in the 2001 edition of the CEFR (prior to the proliferation of empirical evidence supporting CL in L2 instruction), detailed guidance on metaphor instruction and assessment is notably absent. As Nacey (2013) observes, "neither awareness of metaphor […] nor any measure of proficiency in its use appears in the illustrative descriptors of any of the framework's descriptive scales of assessment" (p. 54). Consequently, a closer look at how metaphor is treated within the CEFR can thus provide insight into its perceived importance within L2 instruction and formal assessment, revealing the expectations for L2 learners' acquisition of metaphors.

Regarding linguistic competence, the term "metaphor" is mentioned in the context of lexical competence. Metaphor receives brief attention in the description of a specific type of fixed expression: "phrasal idioms, often semantically opaque, frozen metaphors" (Council of Europe, 2001, pp. 110–111). This particular type of metaphor, excluding other forms of metaphorical language, does not prominently appear on the CEFR assessment scales until the proficient user levels of C1 and C2. At these levels, metaphor is associated with vocabulary range, indicating "a good command of idiomatic expressions and colloquialisms" (Council of Europe, 2001, p. 112). The CEFR implies that L2 learners are not expected to actively use

metaphorical language, limited to idioms and proverbs, until they achieve advanced L2 proficiency levels. However, at the B2 level, idiomatic expressions are identified as potentially challenging, particularly in *Reading*. Learners should have "broad active reading vocabulary, but may experience some difficulty with low frequency idioms" (Council of Europe, 2001, p. 69). For a more detailed discussion on this topic, see Nacey (2013, pp. 43–55).

The importance of metaphor in L2 competence continues to be understated in the recently published CEFR Companion Volume (Council of Europe, 2020). Although the updated edition introduces the description of metaphor skills at the advanced levels, it primarily treats them as ornate or complex rather than integral aspects of communicative competence. Yet, metaphors receive limited attention at the lower levels of L2 proficiency within these influential guidelines on L2 use (MacArthur, 2021).

At the C2 level, the CEFR specifically addresses metaphor in the context of building a plurilingual repertoire. Learners should be able to "explore similarities and differences between metaphors and other figures of speech in the languages in their plurilingual repertoire, either for rhetorical effect or for fun" (Council of Europe, 2020, p. 128).

Additionally, the CEFR discusses metaphor in the context of *mediation*. At the C2 level, learners are expected to interpret creative texts, ranging from literature to film, theatre, recitals, and multimodal installations, which are considered works of imagination and cultural significance. Learners should be able to "recognise the finer subtleties of nuanced language, rhetorical effect and stylistic language use (e.g. metaphors, abnormal syntax, ambiguity), interpreting and "unpacking" meanings and connotations" (Council of Europe, 2020, p. 107).

Metaphors are also linked with creativity in L2 production at advanced levels, particularly in writing. At C1, learners should be able to "incorporate idiom and humour, though use of the latter is not always appropriate"; at C2, they should be able to "exploit idiom and humour appropriately to enhance the impact of the text" (Council of Europe, 2020, p. 67).

The recent study by O'Reilly and Yan (2023) has shed light on the role of metaphor in L2, with a specific focus on Metaphor Language Play (MLP), a key construct in L2 MC. This research analysed the understanding of humour in US sitcoms among 69 advanced L1

Mandarin L2 English learners. Using the MC Test battery developed by O'Reilly and Marsden (2021), the researchers employed Exploratory Factor Analysis to investigate the relationships between various MLP measures, including both receptive and productive skills in written and spoken forms. A thematic analysis was also conducted to identify linguistic, conceptual, and metalinguistic themes in the learners' MLP. Their findings highlighted the interconnectedness of different MLP modes/knowledges, and revealed a rich array of linguistic, conceptual, metalinguistic themes underpinning learners' MLP. They further discussed the potential pedagogical implications of their findings. They proposed adapting certain study elements for use in classroom tasks and assessing learners' MLP abilities. However, they noted that these recommendations remain tentative, given the study's non-interventional nature.

It is important to note that the CEFR clearly states that its categories and examples are intended as suggestions rather than prescriptive rules and should be tailored to individual pedagogical practices. Nevertheless, descriptors for CEFR levels A2–C2 regarding the use of metaphor, as proposed by Littlemore et al. (2014, pp. 142–143), have not been widely adopted in L2 classroom practices. This limited uptake might be partly due, as Nacey (2017) points out, to the fact that "far more practitioners consult the CEFR itself, rather than scholarly articles" (p. 510). Consequently, L2 teachers may favour the CEFR's direct recommendations, frequently incorporated by material writers in syllabus design, over academic proposals, which could restrict the integration of CL research findings into teaching curricula.

Littlemore et al. (2014) highlight the importance of a washback effect, where changes in language testing influence the teaching and learning of metaphor in L2 classrooms, especially at B2 level. This underscores the significance of understanding how metaphorical productions are evaluated in L2 assessments. MacArthur (2021) emphasises the crucial nature of such evaluations, stating "[i]t is thus a matter of some importance to know how these metaphoric productions are judged by the assessors entrusted with awarding a pass or a fail" (p. 356). Furthermore, MacArthur (2021) discusses the implications for teaching, adding:

> Knowing how metaphors are assessed could, in turn, guide teachers entrusted with preparing learners for these examinations to incorporate attention to metaphor in their

> classes, in order to help learners to deploy metaphor in their language production in communicatively successful ways. (MacArthur, 2021, p. 356)

A potential barrier to integrating these practices in L2 assessment could be the lack of specialised training for examiners in evaluating metaphorical language in L2 production. This gap may explain why L2 assessment organisations have yet to incorporate these descriptors in their ESOL examination criteria. Littlemore et al. (2014) argue for increased leniency from L2 assessors towards deviations from conventional L1 language patterns, particularly emphasising the developmental phase of L2 learners at the B2 level. This perspective aligns with the need for more explicit training, not only for assessors but also for instructors, as Piquer-Píriz and Martín-Gilete (forthcoming) suggest. Such training is essential if CL-proposals are to be effectively implemented in the L2 classroom.

Beyond changes in assessment, it is imperative to provide clear instructional guidelines for teaching metaphors across the various CEFR levels. MacArthur (2017) notes that current training materials do not offer sufficient guidance for teachers or learners in this regard. While instructors may be trained in the applications of CL to L2 learning, it is essential to integrate teaching guidelines for explicit metaphor instruction into official L2 competence descriptors, such as the CEFR. Its “can-do” statements notably influence the design of mainstream materials and, by extension, the assessment criteria of major examination boards.

Given these observed gaps in training and curriculum, numerous scholars in metaphor research have been promoting the benefits of integrating metaphor awareness in L2 instruction (Boers, 2013; Boers & Lindstromberg, 2006, 2008a; De Knop et al., 2010; Low, 1988; MacArthur, 2010; Piquer-Píriz & Alejo-González, 2020). Yet, MacArthur (2021) emphasises the necessity for clear assessment methods to inform teaching guidelines, stating that “[e]ducational processes are indissolubly linked to educational products” (p. 357). Despite this widespread call for raising metaphor awareness in L2 classrooms, there remains a dearth of comprehensive studies exploring how metaphor is assessed within examination materials.

The role of metaphors in high-stakes ESOL examinations, particularly in relation to the CEFR descriptors, remains somewhat ambiguous (see Golden, 2021; Littlemore et al., 2014;

Nacey, 2013). While correlations have been identified between both metaphor comprehension and production fluency and L2 proficiency using elicited methods (see O'Reilly & Marsden, 2023), the relationship between naturally produced metaphor (both spoken and written forms) in real-world contexts and L2 competence assessment is still undefined.

A central concern is whether the ability to use metaphors is deemed essential for L2 competence assessment by major examination boards. This raises questions about the validity of assessing metaphor use in high-stakes examinations, pertaining to whether such an assessment truly measures a meaningful and necessary component of L2 competence. It also prompts examination of the value of instructional time dedicated to teaching metaphors in EFL settings, addressing whether the investment in teaching this aspect of language is justified in terms of learning outcomes. As MacArthur (2021) remarks, "what kind of creativity is applauded or indeed tolerated by those entrusted with evaluating the metaphors used in written texts remains largely an unanswered question" (p. 355).

2.4. Challenges and Prospects in Pedagogically Oriented Metaphor Research

Pedagogical applications of CL have proven valuable in L2 instruction, offering significant insights to the field (Piquer-Píriz & Alejo-González, 2020). However, CL-inspired approaches have not effectively reached L2 classrooms (MacArthur, 2017). Major instructional frameworks that influence EFL syllabus design and mainstream materials, such as the CEFR, do not explicitly address metaphor instruction (Nacey, 2013; MacArthur, 2021). This omission may partly account for the limited transfer of key research findings into L2 classroom practices.

The restricted impact of metaphor research on influential guidelines for L2 instruction and assessment may be attributed to methodological shortcomings in some applied metaphor studies (see Boers, 2013; Low, 2017). This critical review further highlights an imbalance in research efforts, which could also contribute to the marginalisation of metaphor in TEFL. The research output has been conducted irregularly by a limited group of scholars, when compared with other fields of research in ELT. The lack of robust and conclusive findings, compounded by the inherent complexities of classroom-based research, may impede the implementation of

CL insights in L2 classroom practices. Additionally, although they are appealing for learners and valued by teachers, the lack of clear guidelines, substantial investment of time, and both cognitive and training efforts required by these teaching methods make them unlikely to be adopted by L2 instructors (Boers, 2022). Furthermore, real L2 classrooms, with their variable conditions, differ significantly from language labs, which further exacerbates these challenges.

There is a pressing need for comprehensive research, aimed at strengthening previous findings and at assessing the feasibility and benefits of a CL-inspired approach to teaching metaphor in real-world EFL contexts. Despite three decades of sustained and fruitful efforts to secure metaphor's role in L2 instruction, significant gaps remain, especially in determining the extent to which motivated language can be enhanced in the real L2 classroom.

Pedagogically oriented metaphor research has predominantly focused on *how* to enhance metaphor awareness among learners in instructed L2 settings through CL-inspired instruction. This body of research has emphasised exploring the teaching-learning process, with a specific goal of optimising the short-term comprehension and retention of conventional metaphorical language, often presented out of context in the materials used in the experiments (Boers, 2013). However, our understanding of the long-term benefits of employing a variety of CL-oriented methods within regular teaching practices to foster lexical development remains incomplete, particularly regarding the sustainability of learning gains beyond the scope of the instructional intervention, from quantitative and qualitative perspectives.

A significant gap exists in our knowledge concerning *what* metaphorical language should be taught at each L2 proficiency level, especially for non-adult students. How learners actually produce metaphors over time in their L2 has received less attention, especially in oral discourse. This issue is particularly acute at the B2 level, which is a critical stage for a significant increase in metaphor use in learner discourse, requiring explicit metaphor instructional support (Littlemore et al., 2014). Nonetheless, how far direct metaphor instruction actually facilitates depth of vocabulary knowledge remains singularly unclear. Additionally, methods for assessing learning gains associated with metaphor still need to be clearly identified (O'Reilly & Marsden, 2021).

There is a significant shortfall in empirical intervention studies that investigate the outcomes of CL-inspired instruction in real-world L2 classroom settings, particularly those providing comparable conditions for control and experimental groups. The challenge of translating CL insights, often derived from small-scale studies with outcomes that do not reflect typical classroom environments, into long-term teaching programmes also remains largely unexplored (MacArthur, 2017). The effectiveness of systematically integrating distributed teaching of metaphorical language into the EFL syllabus to help L2 learners enhance their organic metaphor use, including all areas of language beyond idioms and phrasal verbs, and enhance their autonomous learning is still unclear (Boers, 2013).

Furthermore, how concepts such as “metaphor”, “basic meaning” or “extended meaning” can be seamlessly integrated by CL-trained instructors into L2 classroom discourse is still not clear. Detailed observations from researchers actively involved in the complexities of the real classroom setting remain scarce. Therefore, increased attention to the actual dynamics of teaching and learning interactions is required (MacArthur, 2021).

Additionally, it is uncertain whether enhancing metaphor awareness through input enhancement can accelerate learning for students in instructed L2 settings, which are often characterised by learners with diverse abilities and attitudes. Further research is essential to assess how increased metaphor awareness impacts overall L2 competence, particularly in terms of proficiency in productive skills. How metaphor-related learning gains are reflected in formal assessment practices remains unclear, adding to the uncertainty about the role of metaphor usage in enhancing performance on L2 competence exams (MacArthur, 2021).

While research evaluating CL-inspired instruction may have its limitations, there is compelling evidence supporting the integration of CL principles in L2 classroom practices (Boers & Lindstromberg, 2006; Boers et al., 2010; Bielak, 2011; Boers, 2011). Boers (2013) encapsulates this perspective:

> [C]ollectively, the studies begin to constitute a body of evidence that is hard to dismiss out of hand. As always, more research would definitely be welcome to confirm and fine-tune the findings. In addition, it would be worth finding out if the effectiveness of

> CL-informed interventions could be enhanced by aligning them better with insights from 'mainstream' approaches to L2 vocabulary. (Boers, 2013, p. 216)

Therefore, a fundamental shift to broaden the research focus is essential to address real-world educational problems. To the best of my knowledge, to date, no detailed investigation has been conducted that encompasses both the pedagogical process and the outcomes of instruction concerning L2 metaphor production. Such a study would be useful to shed light on L2 assessment practices and could provide guidance to refine teaching methods. It may enhance our knowledge of how metaphor awareness contributes to L2 learning. Specifically, the study could shed light on the extent to which CL-inspired instruction aids L2 learners in mastering the linguistic dimension of metaphor, i.e., lexico-grammatical patterns. This insight is crucial for achieving native-like accuracy in standard L2 competence exams.

2.5. Objectives and Research Questions

In this PhD dissertation, I conducted a study on the effects of incorporating metaphor awareness into topic-based teaching of metaphorical language in an EFL syllabus linked to the CEFR descriptors (specifically, B2 level) from a longitudinal perspective. This study explores CL-oriented approaches, adopted by an EFL instructor trained in CL, to distributed learning of metaphor and its impact on the oral and written production of L2 learners of English who are receiving private tuition at a language school to reach a B2 level. In turn, the present research study examines whether the learning gains associated with using metaphorical language result in improved performance as assessed in standard L2 competence exams. Specifically, three research questions are addressed:

- RQ1: To what extent does incorporating CL-oriented approaches to distributed learning of metaphor in the L2 classroom affect metaphor use in the oral and written production of L2 learners of English studying at B2 level?
    a. What is the effect of metaphor-mediated instruction on the amount (measured in metaphor density) of metaphorical language use in oral and written learner discourse at B2 level?

b. What is the effect of metaphor-mediated instruction on the type (open-class vs. closed-class metaphors) of metaphorical language use in oral and written learner discourse at B2 level?

- RQ2. How does raising metaphor awareness in topic-based instruction affect L2 learners' achievement at B2 level, as measured by the *B2 First for Schools Cambridge English* qualification, compared to not using CL-oriented methodological techniques?
- RQ3: What relationship, if any, can be established between using metaphorical language in the L2 with English language proficiency at B2 level?

To answer the research questions above, this study adopted a topic-based approach to analysing L2 discourse, focusing on lexico-grammatical elements. It should be noted that, although prepositions were included in metaphor analysis, other grammatical aspects were not considered when identifying metaphorical language (see Section 3.3.1 of Chapter 3).

PART TWO

♦ THE STUDY ♦

CHAPTER 3

METHODOLOGY

Chapter 3 describes the methodology employed to conduct the present empirical study in three main sections. The first section (3.1) provides an overview of the context and design of the longitudinal study. The second section (3.2) explains the implementation of the teaching method and the procedure followed in metaphor-mediated instruction. The third section (3.3) illustrates the method used to gather and analyse the use of metaphor in the learners' discourse, including oral and written production.

3.1. STUDY DESIGN

This section presents the context and design of the longitudinal study, organised as follows. Section 3.1.1 introduces the context of the study and Section 3.1.2 provides information on the participants, including ethical considerations. Section 3.1.3 describes the research instruments used, and Section 3.1.4 details the types and sources of research data collected. Section 3.1.5 presents an overview of the timing and procedure followed in the longitudinal study, and Section 3.1.6 illustrates the transcription method used for the oral data collected. Section 3.1.7 provides some concluding remarks.

3.1.1. Context of the Study

The study was carried out at an English language school[12] situated in Cáceres, a town located in the western region of Extremadura, Spain, with a population of around 96,000. In this region, students often participate in after-school activities to supplement their formal education. Learning English is particularly popular among students due to its potential to improve their future job prospects and chances of admission to international higher education

[12] The name of the English language school has been anonymised.

institutions. As a result, teenagers in this area frequently attend extracurricular English classes that are specifically designed to prepare them for official English language proficiency exams.

The English language school involved in this study is an authorised centre for preparing L2 learners for *Cambridge English Qualifications* and administers mock exams. Its EFL teachers receive specialised training to use official teaching resources and exam preparation materials, including goal-oriented coursebooks and authentic examination papers. *Cambridge English Qualifications* are globally recognised for their reliability in assessing English language proficiency in all four language skills (*Reading*, *Writing*, *Listening*, and *Speaking*), as well as vocabulary and grammar knowledge. These exams are aligned with the Council of Europe's (2020) CEFR, which is widely accepted as the international standard for benchmarking language ability in real-life contexts (see Section 2.3.3 of Chapter 2).

All participants involved in the study reported in this PhD dissertation were enrolled in English courses at CEFR B2 level to prepare for the *B2 First for Schools* examination. This *Cambridge English* diploma certifies mastery of the language skills necessary to communicate confidently in written and spoken English-speaking environments at the upper-intermediate level. At this level, students are expected to follow conversations, discuss general interests, and express and justify their opinions. The B2 exam is available in paper-based and computer-based formats, with this *Cambridge English* qualification specifically designed for school-aged learners, with topics and tasks tailored to their needs. L2 learners receive both overall performance and individual skill scores, and achieving a Grade A, B, or C in the exam earns them the *B2 First for Schools* certificate, which confirms their proficiency in English at B2 level.

The reason for collecting data in this instructed L2 setting was my affiliation with the English language school at the time of the investigation, where I was one of the EFL teachers responsible for providing comprehensive B2 training. This learning environment offered an optimal context for carrying out the study due to several favourable factors, including flexible scheduling, low student-teacher ratio, and homogeneous groups of students.

3.1.2. Participants

This study used a convenience sample of 40 L2 learners of English (22 females and 18 males) aged between 14 and 18 (*M* = 16.13, *SD* = 0.86) whose L1 is Spanish. All participants were high school students enrolled in B2-level English courses at the abovementioned language school. The participants received one-hour English lessons three times a week from October 2019 to May 2020 and had no prior experience participating in studies on the applications of CL to TEFL. All participants had a similar background in *Cambridge Assessment English*, having all passed the *B1 Preliminary for Schools* examination in 2018, two school years before the start of this research project. The average score on the qualification was 148.99 (*SD* = 7.20), with 70% of L2 learners receiving the lowest grade (Grade C) for B1 level, according to the *Cambridge English Scale* (UCLES, 2015).

Participants were selected from those in their second year preparing for the *B2 First for Schools* examination. To ensure effective B2 training, the English language school created homogeneous groups of L2 learners with similar proficiency levels and demographic characteristics. Four groups, each comprising 10 students, were selected as the most comparable for the study. The first and third groups (GR0101 and GR0203) consisted of upper-secondary education students (years 10 and 11), while the second and fourth groups (GR0102 and GR0204) included students in university preparation education (years 12 and 13), forming a relatively homogenous population.

The experimental group comprised 20 participants (*M* age = 15.97, *SD* age = 0.97), randomly selected from two B2-level groups in each age classification in the sample (GR0101 and GR0102). The control group included 20 L2 learners (*M* age = 16.28, *SD* age = 0.73) randomly selected from the other two groups in the sample (GR0203 and GR0204). More detailed information on participants can be found in Table 62 in Appendix A.

The Ethics Committee at the University of Extremadura granted research ethics approval prior to the investigation (see Appendix B). Additionally, a request form for research was submitted to the English language school to inform them about the study (see Appendix C). Upon receiving a positive response, the English language school and this researcher

signed a research commitment agreement (see Appendix D). All participants were informed that they would participate in a study conducted by the University of Extremadura to enhance vocabulary acquisition and develop the oral and written production of L2 learners at B2 level. However, the precise purpose of the study was only explained to students and EFL instructors after the study was completed. It is worth noting that I had planned to introduce CL-oriented approaches to the control group after the study to provide them with the same benefits of raised metaphor awareness in the L2 classroom as the experimental group. However, this plan was disrupted due to the Coronavirus Disease 2019 (COVID-19) outbreak (see Section 3.1.5 of this chapter).

Parental or legal guardian authorisation was requested to ensure compliance with ethical standards. Participant informed consents were signed before the study (see Appendix E). To protect participants' privacy, all data collected was anonymised and a unique code was assigned to each participant, which was used throughout the investigation (see Appendix A). Students' codes will be used to refer to participants in this study.

The present research was carried out with the collaboration of three experienced EFL practitioners. As shown in Table 1 below, INS0101 and INS0202 were responsible for training participants for B2 level, while INS0303 solely assessed the L2 learners to ensure unbiased exam performance marking. None of the study participants had received instruction from any of the three EFL instructors during their first year of B2 preparation.

Table 1

Overview of EFL Teachers: Linguistic Characteristics and Training Backgrounds

Role	Code[13]	Gender	Age	L1	TEFL training	TEFL expertise (yrs.)
Treatment teacher	INS0101	Female	29	Spanish	Master's	5
Control teacher	INS0202	Male	25	Spanish	Master's	2
Participants' assessor	INS0303	Female	47	English	TESOL certificate	22

[13] Each instructor was assigned a code using the following format: the abbreviation of "instructor", the code of the group they interacted with during the study (experimental = 01; control = 02; both = 03), and their position in the list of instructors in alphabetical order.

INS0101 and INS0202 shared similar linguistic characteristics, training backgrounds, and TEFL expertise. Throughout the study, all participants received B2 training using the same goal-oriented textbook and teaching resources, which were provided by these practitioners. The experimental group was instructed by INS0101, an EFL teacher who was trained in CL and the author of the present study. INS0101 implemented CL treatment to teach the metaphors in the textbook and carried out participant observation as part of the research (see Section 3.2 of this chapter for a detailed description of the teaching treatment).

Participants from the control group received B2 training from INS0202, a different EFL teacher at the English language school who lacked training on the applications of CL to L2 instruction. To gain a deeper understanding of the control group teacher's profile, INS0202 was asked to complete an online questionnaire prepared in *Google Forms*, offering details on his participant-related information, training background, and TEFL method used in the L2 classroom. The questionnaire responses confirmed that the control group teacher had not received training in CL and provided the necessary information to establish his approach to preparing L2 learners for the B2 exam (see Section 3.2 of this chapter for further details).

The third instructor, INS0303, was solely responsible for assessing the oral and written production of L2 learners during both the pre- and post-testing phases. As the administrator of the language school, INS0303 brought over 20 years of teaching experience as a native speaker of English. Additionally, she was a certified *Cambridge English Assessment* specialist with extensive experience and knowledge in evaluating learners at B2 level and above, especially in *Speaking* examination work in Extremadura for over 15 years. She also had a wealth of experience in grading L2 learners' progress in *Writing* through formal mock exams.

3.1.3. Instruments

To investigate the three research questions, three instruments were employed: (a) the B2-level textbook used in class for examination training, as well as (b) the pre-test and (c) post-test measures. As will be seen, these instruments served different research purposes,

including conducting teaching treatment, assessing the B2 English proficiency of participants, and identifying metaphor use in L2 learners' oral and written outputs before and after the study.

3.1.3.1. Instrument 1: Topic-Based Textbook

The first research instrument employed in this study was the B2-level textbook used in the L2 classroom. It served two main purposes in carrying out the teaching treatment: (a) preparing all participants for the *B2 First for Schools* examination, and (b) identifying metaphors in the input texts that these L2 learners encountered during B2 training in order to develop the CL-inspired activities implemented with the experimental group.

While this mainstream material was a significant input source for L2 learners, other forms of support also contributed, including natural language input that they encountered inside and outside the instructed L2 setting, the English input received in the formal educational context (i.e., their schools), and teacher or peer language. However, given the goal-oriented nature of *Cambridge English Qualifications*, official exam preparation materials designed for English courses are considered the primary support for examination training.

3.1.3.1.1. Preparing for *B2 First for Schools* Examination

The textbook selected to prepare the 40 participants for the *B2 First for Schools* qualification was *Complete First for Spanish Speakers* (Brook-Hart, 2014).[14] This coursebook, adhering to the principles of the communicative approach (Richards & Rodgers, 2001), was the official preparation material employed to train L2 learners for the upper-intermediate level in the EFL setting described in the study. Based on the *Cambridge English Corpus*, a collection of over two million words gathered from real situations and conversations of native English speakers, this mainstream resource provided L2 learners with the required language while preparing them to reach B2 level. The textbook's topic-based approach enabled students to consolidate the language skills and knowledge, including vocabulary and grammar, that are

[14] *Cambridge University Press* granted the author of this PhD dissertation non-exclusive permission to reproduce text extracts from this source, solely for the purpose of non-commercial publication.

necessary to reach B2 level in the CEFR. The topics and tasks are tailored to the B2 exam requirements and designed to align with the interests and experiences of the age group.

The topic-based textbook consists of 14 units for classroom study, which were intended to be taught over two school years of B2 training at the English language school. In the first year, all 40 participants completed the first seven units. The remaining seven units were scheduled to be taught during their second year of B2 preparation when the study reported in this PhD dissertation was conducted. However, only four of these units were actually covered due to the reasons explained in Section 3.1.5 of this chapter. For a map of the units included in the textbook used in class, see Table 63 in Appendix F.

3.1.3.1.2. Selecting Target Lexis for Metaphor-Mediated Instruction

To identify the target metaphors to be taught through CL-inspired activities as part of the teaching treatment, this study analysed the oral and written input texts of the B2-level textbook used in class. The metaphor analysis was conducted on a total of 23 texts, which comprised 7,200 words and were collected from various comprehension and production tasks.

Tables 64–67 in Appendix G provide a detailed description of the transcripts of the eight oral input texts (3,582 words) and the 15 written input texts (3,618 words) that participants were formally exposed to across the four topics covered in the teaching treatment. The study explains how these topics were taught to all participants and how metaphors were selected to develop CL-inspired activities for the experimental group in Section 3.2 of this chapter.

3.1.3.2. Instruments 2 and 3: Pre-Testing and Post-Testing Measurements

The present study used instruments 2 and 3 as the pre-test and post-test measures, respectively, to assess the B2-level achievement of participants before and after the longitudinal study. These measures also served as sources of input texts that participants were exposed to during both testing phases to produce their oral and written outputs before and after conducting the present research.

3.1.3.2.1. Assessing English Language Proficiency at B2 Level

The pre-test and post-test were two different timed practice tests administered as part of the official B2 mock exams run during L2 learners' second year of B2 training at the language school. These tests were designed to identify areas of strength and weakness in overall and individual skills performance at the beginning and the end of the school year. The *B2 First for Schools* examination assesses L2 learners' proficiency at different CEFR levels based on their achievement. Table 2 displays the *Cambridge English Scale* (UCLES, 2015) scores for the B2-level exam results.

Table 2

Cambridge English Scale Scores: B2 First for Schools

Cambridge English Scale scores[15]	Grade/CEFR	CEFR level
180–190	Grade A	C1
173–179	Grade B	B2
160–172	Grade C	B2
140–159	B1 level	B1

Source: Own elaboration from data extracted from the *Cambridge English Scale* (UCLES, 2015).

The study employed two authentic examination papers from the *B2 First for Schools 3* collection (UCLES, 2018)[16] as pre-test and post-test measures. These papers offered comprehensive coverage of all exam tests and were accompanied by mp3 audio files, answer keys, and detailed assessment guidance. This collection was selected for its up-to-date and reliable exam preparation materials. To avoid familiarity bias, the instructors who taught the participants during their first year of B2 preparation were consulted before the investigation. Accordingly, Test 3 (pp. 51–72) and Test 4 (pp. 73–94) from the *B2 First for Schools 3* (UCLES, 2018) were semi-randomly assigned as pre-testing and post-testing, respectively.

[15] The *B2 First for Schools* examination only reports scores above 140.
[16] *Cambridge University Press* granted the author of this PhD dissertation non-exclusive permission to reproduce text extracts from this source, solely for the purpose of non-commercial publication.

3.1.3.2.2. Input Sources for Output Production

The pre-test and post-test measures involved exposing participants to four input sources, comprising a total of 273 words, which L2 learners used to produce their oral and written outputs in the B2 mock exams. The input texts included two oral texts (230 words) and two written texts (43 words) that served as prompt questions and task instructions for participants to complete the *Writing* and *Speaking* tests.

Table 68 in Appendix G presents a full description of the two input texts used in the pre-test (142 words) and the two input texts used in the post-test (131 words). Further information on the specific focus of these tasks is provided next in the description of output texts collected to identify uses of metaphor before and after the longitudinal study.

3.1.4. Data

The study collected data from the pre-test (instrument 2) and post-test (instrument 3). Data consisted of participants' overall, *Speaking*, and *Writing* B2 proficiency scores, which were examined to evaluate their performance before and after the longitudinal study. Additionally, data was obtained from the oral and written outputs of L2 learners at B2 level to measure differences in their use of metaphors before and after conducting this research.

3.1.4.1. Participants' B2 First for Schools Scores: B2 Level of English Proficiency

The data collected from the testing measures used in this study included the English language proficiency scores of the participants at B2 level, as determined by the four well-known components of the *B2 First for Schools* examination: *Reading and Use of English*, *Writing*, *Listening,* and *Speaking*. Each exam paper was equally weighted, meaning that achieving the CEFR B2 level required an overall performance across all components.

Appendix H details the objectives and format of each exam paper, while Appendix I summarises the assessment criteria for each paper. For a more comprehensive description of the *B2 First for Schools* examination, see UCLES (2019a).

3.1.4.2. Participants' Outputs: Metaphor in Learner English Discourse

The data collection of learner discourse for metaphor analysis involved participants' oral and written production from both pre-testing and post-testing measures, resulting in 160 texts with a total of 30,280 words. The pre-test analysis included 80 output texts with 15,283 words, and the post-test analysis comprised 80 output texts with 14,997 words.

As is well-known, the topic-dependence of words and expressions in language has been shown to affect the use of metaphors in discourse, both in terms of the type (open- vs. closed-class metaphors) and the number of metaphors used (Deignan et al., 2013; Golden, 2012, 2021; Semino, 2008). Accordingly, this investigation focused on analysing discourse that included broad topics, as will be illustrated below.

In this study, differences in text length were not used as exclusion criteria for metaphor analyses (see, however, Hoang & Boers, 2018). Following Nacey's (2020) approach, all output texts collected from the research instruments were accepted for metaphor identification, underlining "what the texts are intended to represent" (p. 178). Thus, all texts collected in the study were deemed representative of the L2 learners' oral and written production skills and use of metaphor, regardless of variations in word length.

For comparison purposes, the identification of metaphor use in the learner discourse only included exam parts in the *Speaking* and *Writing* tests that shared similar task performance and were completed by all participants. While the topics of the pre-test and post-test tasks were relevant to the interests and experiences of the participants, there was no guarantee that the topics would promote metaphor use in their oral or written production. The following subsections provide an overview of the learner discourse for each type of output text.

3.1.4.2.1. Oral Output Texts (Transcriptions)

The study extracted a total of 80 oral output texts, totalling 15,909 words, from the learner production in the *Speaking* tests performed in the pre-testing and post-testing measures. Section 3.1.6 of this chapter describes the transcription method employed to transcribe learners' oral outputs in the study.

The oral texts were collected as part of the *Speaking* task provided in Part 4 of the exam paper in the pre-test and post-test (see Appendix H). This task required participants to engage in an in-depth discussion on a given topic, demonstrating various interaction patterns that reflect authentic conversation. Participants were instructed to express and justify their opinions, agree or disagree, and speculate. Similar to *Writing* Part 1 (see description below), this task primarily focused on evaluating issues related to the given topic rather than providing general information. Table 78 in Appendix J presents an overview of the oral output texts analysed in both testing measures, including text codes and length per participant.

Pre-Test Oral Output Texts (Transcriptions). The pre-test analysis of the 40 oral productions yielded a total of 8,486 words. The participant's ability to engage in a discussion was assessed based on the topic of *helping children with homework* (UCLES, 2018, p. 97). The stimulus provided by the interlocutor is listed below:

- *Do your parents give you a lot of help with your homework? What do they help with?*
- *Some people say that children shouldn't do any homework during the school week. Do you agree? (Why? Why not?)*
- *Do you think the school day should be longer so that students don't have to do work at home? (Why? Why not?)*
- *Some people say that homework isn't a good thing for children because they stay up too late doing it. What do you think?*
- *Do you think it would be a good idea for students to do homework online so that the teachers can see what they've done? (Why? Why not?)*
- *Do you think it's true that giving children prizes is the best way to encourage them to work harder? (Why? Why not?)*

Post-Test Oral Output Texts (Transcriptions). The post-test analysis comprised 40 oral outputs consisting of 7,423 words. Participants were required to discuss the topic of *spending time outdoors* (UCLES, 2018, p. 106) and were provided with spoken prompts to guide their answers. They were instructed to provide full answers to some of the following questions:

- *Do you and your friends spend a lot of your time doing outdoor activities? (Why? Why not?)*
- *Should students have to do outdoor sports lessons at school? (Why? Why not?)*
- *Do you think it's a good idea for families to go on holidays where they spend a lot of time outside, for example camping holidays? (Why? Why not?)*
- *Is growing up in the countryside better for children than living in cities? (Why? Why not?)*
- *Some people say that it is important for big cities to have good parks. Do you agree? (Why? Why not?)*
- *Do you think it's true that if people spent more time outside, they'd care more about protecting the environment? (Why? Why not?)*

3.*1.4.2.2. Written Output Texts*

In this study, a total of 80 written output texts containing 14,371 words were extracted from the learner production in the *Writing* tests administered in both pre-testing and post-testing measures. Written output texts consisted of opinion essays that were part of the Part 1 *Writing* task (see Appendix H). Similar to *Speaking* Part 4 (see above), the compulsory task set for the *Writing* component involved argumentative discussion, requiring participants to express their opinions and justify their personal views on a given essay question or statement.

The *Writing* instructions included a mandatory structure for all learners, requiring them to address two given topics linked clearly to the essay question or statement and discuss a third additional idea of their own. The resulting discursive written texts evaluated the learners' ability to agree or disagree with the essay statement, provide supporting reasons and

explanations, offer examples, compare and contrast ideas, and draw conclusions. Additionally, participants were expected to use a neutral tone and formal style, employing complex sentence structures and infrequent use of pronouns while avoiding colloquial or slang English. Table 77 in Appendix J provides an overview of the 80 written output texts analysed in the pre-test and post-test, including text codes and length per participant.

Pre-Test Written Output Texts. The pre-test analysis included 40 written output texts, which contained a total of 6,797 words. Participants were asked to express whether they agreed with the statement that teenagers currently experience more favourable circumstances compared to some years ago: *Young people nowadays have many more opportunities than young people in the past. Do you agree?* (UCLES, 2018, p. 64). This prompt required participants to address two given topics (*education* and *travel*) and a third idea of their own.

Post-Test Written Output Texts. To examine the discourse in the post-test production, a total of 7,574 words were analysed from 40 written texts. The essay prompt asked participants to express their views of schools in the future, specifically on the possibility of important changes: *Some people say that schools will be very different in the future. What do you think?* (UCLES, 2018, p. 86). Participants were instructed to discuss the topic by addressing *the teachers* and *the subjects*, and an additional third idea of their choice.

3.1.5. Timing and Procedure of the Longitudinal Study

The study was designed as a longitudinal investigation conducted during the 2019–2020 school year, with an initial plan to run for six months between 15th November 2019 to 15th May 2020, totalling 27 weeks. However, the outbreak of COVID-19 disrupted the research schedule, and on 14th March 2020, the Spanish government imposed a nationwide lockdown. Consequently, the teaching treatment was terminated six weeks early, and the post-test administration was moved forward due to the nearly three-month home confinement. A five-month study was then carried out from 15th November 2019 to 25th April 2020, resulting in a

total of 24 weeks of research. The longitudinal investigation underwent three research stages that were adapted to the instructed L2 setting and aligned with the study objectives. Table 3 illustrates the timeline of each stage, which is detailed in the following subsections.

Table 3

Experiment Time Frame for Participants

Study phase	Research period	Groups of participants		Place administered
		Control group	Experimental group	
Phase 1: Nov. 2019	Week 1	Informed consent form Speaker-related characteristics form Pre-test: Written tests		Language school
	Week 2	Pre-test: Oral test – Part 1		
	Week 3	Pre-test: Oral test – Part 2		
Phase 2: Dec. 2019 – Mar. 2020	Weeks 4–18	Control teaching	Treatment teaching	Language school
Phase 3: Apr. 2020	Weeks 19–23	B2 level examination training		Online
	Week 24	Post-test: Written tests Post-test: Oral test – Part 1 Post-test: Oral test – Part 2		

3.1.5.1. Study Phase 1

In the first week of the study, participants filled out informed consent forms and an online questionnaire in *Google Forms*, providing information about their speaker-related features and background. Following this, a pre-test was administered to measure learning and collect data on learner discourse. All participants completed the written papers of the mock exam at the exam venue arranged for the pre-testing event in one afternoon, following the order of the real *B2 First for Schools* examination. The researcher invigilated the pre-test with the assistance of INS0202 and INS0303.

During weeks 2 and 3, the oral pre-test was conducted in a quiet classroom at the English language school. Participants were randomly paired and took turns undertaking the *Speaking* test, which was overseen by the researcher. One examiner (INS0303) acted as both interlocutor and assessor, and audio-recordings of students' performance were used for evaluation, thereby eliminating the need for a second examiner.

3.1.5.2. Study Phase 2

In the second phase of the study, the experimental group received teaching treatment by implementing metaphor-mediated instruction in the L2 classroom. However, the COVID-19 pandemic disrupted the research during week 18 of the longitudinal study, resulting in the originally planned 21-week teaching treatment being shortened to 15 weeks to run alongside the five-month investigation.

Starting from week 4 of the longitudinal study, all four B2 groups underwent topic-based training using the same textbook. The instructors of both the control and experimental groups coordinated weekly to ensure consistent coverage of textbook contents at the same pace. For more information on the teaching treatment sessions' specific method, timeframe, and procedure, see Section 3.2 of this chapter.

During weeks 19 to 23 of the longitudinal study, all participants were instructed online through the videoconference software programme *Zoom*[17] due to lockdown measures imposed in response to the worsening COVID-19 pandemic. The Spanish confinement measures had been extended since mid-March 2020, leading the English language school administration to bring forward the end of the school year and continue with online classes until late April 2020. However, the online sessions differed from the regular English classes offered at the language school. In the time between the teaching treatment at the language school and the post-testing, participants received intensive online examination training, which included weekly coordination with the control group instructor. All participants received the same exam practice using authentic examination papers rather than completing the remaining three topics in the textbook (see Section 3.1.3 of this chapter). Therefore, metaphor awareness was not further introduced in the normal range of the topic-based activities performed in class. Nonetheless, CL treatment was provided for lexicogrammar elements contained in the preparation sample papers where possible.

[17] The videoconference platform *Zoom* can be accessed at https://zoom.us/.

3.1.5.3. Study Phase 3

During the third phase of the study, the primary goal was to measure the effectiveness of the teaching treatment through a post-test. However, due to conflicts arising from the COVID-19 pandemic, the post-test was delayed until the announcement of the lifting of confinement conditions. To avoid bias, the post-test was conducted remotely six weeks after the teaching treatment and slightly over five months after the pre-test, given the worsening epidemiological situation and the prolonged period of mandatory home confinement.

The delayed post-test was administered using *Zoom* as the examination environment. Before the online post-test, participants were given clear instructions, a schedule, an access link to the *Zoom* event, and computer graphics containing instructions on navigating the computer-based tests in the examination environment. The post-test was divided into two parts: the oral post-test was conducted from the 20th to the 23rd of April, while the written post-test was administered on the 25th of April 2020 in a single morning.

The oral post-test was divided into four online sessions held over several days to accommodate scheduling constraints. Each participant was paired with their pre-test partner and evaluated by INS0303, who also served as both the speaking examiner and interlocutor. The researcher oversaw the *Speaking* test, audio-recorded the participants' performance, and provided technical support on *Zoom*. Although some participants experienced internet connectivity issues that affected their sound quality, no biases were detected.

The written post-test was conducted in a computer-based format using *Google Forms* with the "quiz" setting. Participants received a link to the *Google Forms* through *Zoom* chat and were given detailed instructions to complete the post-test. All written tests (*Reading and Use of English*, and *Listening*) were completed online, except for the *Writing* test, which required participants to handwrite their answers and upload pictures to the *Google Forms* task. The post-test was conducted under strict examination conditions, with time limits for completing all exam papers. To prevent cheating, participants were required to keep their webcams on throughout the post-test, and the "gallery view" layout on *Zoom* enabled simultaneous supervision of participants.

After conducting both tests, they were anonymised and scored manually. The *Reading and Use of English* and *Listening* tests were scored by the researcher using the provided answer keys, while the *Writing* and *Speaking* tests were evaluated by the official examiner (INS0303) using analytical assessment scales. To calculate the overall performance, individual scores for each skill were averaged by adding up the separate scores received for each exam paper and dividing them by five. The overall score was then rounded to the nearest whole number. Finally, the test scores obtained were converted to *Cambridge English Scale* scores (UCLES, 2019b). Individual scores were recorded in a *Microsoft Excel* (2021) spreadsheet and processed using various methods. Hand-written learner discourse was entered through keyboarding, while oral learner production was orthographically transcribed.

3.1.6. Transcription Method

It is commonplace to say that transcribing oral data poses challenges due to the specific characteristics of spoken language that require careful attention during transcription (Biber, 1988). These challenges are amplified in the case of L2 spoken English due to non-native variability (Baese-Berk & Morrill, 2015; Wade et al., 2007). Standardised transcription procedures ensure a detailed and informative representation of recorded oral data, promoting objectivity and facilitating analysis of written formats that capture various features of oral interaction, such as intonation, (mis)pronunciation, backchannels, and truncated utterances.

To investigate metaphor use in L2 oral discourse, this study used the transcription system designed for non-native spoken English in the *Vienna-Oxford International Corpus of English* (VOICE, 2021) following previous research (e.g., Alejo-González, 2022, 2024; Alejo et al., 2021; Castellano-Risco et al., 2023; Fielden-Burns & Piquer-Píriz, 2022; Krennmayr et al., 2022; Littlemore & Fielden-Burns, 2023; MacArthur, 2016a, 2016b; MacArthur et al., 2015; Philip, 2023). The present research involved transcribing 80 discussions that were audio-recorded from both experimental and control groups (with 40 transcripts for each testing measure) in order to identify metaphor use before and after the longitudinal study.

In my study, the VOICE Project (2007) spelling and mark-up conventions were adapted to suit the specific research objectives, as presented in Table 4 below. For the purpose of this study, non-linguistic or paralinguistic features, such as hesitations, filled and unfilled pauses, laughs, emphasis, and speaking modes, were not annotated in the transcription process.

Table 4

Modifications of the VOICE Project (2007) Transcription Spelling and Mark-up Conventions

Conventions	Modification	Example
Non-English words	Non-English words were rendered in the standard variant of the original language. No decisions were made concerning the representation of diacritics, as these were not found in the transcripts. For instance, consider the discourse marker *bueno* (*well* in Spanish) in example (1).	(1) <POSTB2.ST0404> well yes we love to go outside and see new things and experience new experience new **bueno** things […] </POSTB2.ST0403>
Unconventional constructions	Unconventional constructions due to the use of coinages and possible L1 influence or transfer/calque were not corrected in standard orthographic form. An example is the word *acostumbrate* (*acostumbrarse* in Spanish) instead of "get used to" or "to accustom" in (2).	(2) <PREB2.ST0101> I think maybe yes but not all the time because maybe they can **acostumbrate** to take get prizes [...] </PREB2.ST0101>
Mispronounced words	Mispronounced words were not represented in standard orthographic form as, although L2 learners might have known the word, the word produced was differently, incorrectly uttered. Hence, mispronounced words were not corrected in the standard orthographic form, whether incorrect utterances were due to grammar mistakes affecting correct standard pronunciation, as shown in (3), or unconventional phonological constructions, as illustrated in (4). Similarly, minor instances of L1 accent were not represented in standard orthography. For instance, consider (5), where the final letter 's' was dropped in *nowadays*.	(3) <PREB2.ST0203> yeah I agree with you because children must do the things but if you have if you for example **childrens** [*children*] are more more difficult to do the things […] </PREB2.ST0203> (4) <PREB2.ST0408> […] if not they they wouldn't know how to to take the time in the future when they will be **basy** [*busy*] (.) </PREB2.ST0408> (5) <POSTB2.ST0302> I totally agree with you because I think that **nowaday** [*nowadays*] is better to go outside to do some lessons of any class (.) </POSTB2.ST0303>
Lexicalised reduced forms	Standardization of lexicalised reduced forms was applied. The only example found in the transcription process was the phonological reduction of *going to* as "gonna", as shown in (6).	(6) <POSTB2.ST0201> […] I think you are **going to** have the same interest or or the things that you are **going to** think in your life (.) </POSTB2.ST0201>
Contractions	Orthographical contractions were rendered but not segmented as (7) illustrates.	(7) <POSTB2.ST0310> I **don't** think so because **I'm** I **don't** like

		the sport that the the school teach teaches [...] </POSTB2.ST0310>
Discourse markers	Discourse markers were represented in orthographic form. However, only close sound-acknowledgement sounds were represented in the case of backchannel and positive feedback with such discourse markers as *yes* in (8), *yeah* in (9), *okay* in (10), or negative minimal feedback with the discourse marker *no*, as shown in (11).	(8) <POSTB2.ST0304> **yes** we always go to the to the sport centre and we practise football or bask- or basketball at the weekend (.) </POSTB2.ST0304> (9) <POSTB2.ST0107> **yeah** absolutely and also that a lot of people like me usually stay all day home (.) </POSTB2.ST0107> (10) <POSTB2.ST0409> **okay** I think I don't know if it's better but I think it's very different because in a city they can make more friends […] </POSTB2.ST0409> (11) <PREB2.ST0405> **no** never they if I have a doubt they don't have any problem to solve me but I usually do it my myself (.) </PREB2.ST0405>
Pauses	Neither brief nor long pauses were registered. However, the use of a full stop between parentheses was employed to mark the end of speaker's utterances in turns. For instance, consider example (12).	(12) <PREB2.ST0303> from my opinion I'm I don't agree with with that because school have a lot of hours in my case I have six hours **(.)** and I know people in other cities that have school in the afternoon and I think that is not a good idea because it **(.)** </PREB2.ST0303>
Intonation	Question intonation was used to disambiguate utterances produced as a question rather than a statement. An example is shown in (13).	(13) <POSTB2.ST0105> […] I prefer just to be with my parents at home playing board games or anything but it depends what about you**?** </POSTB2.ST0105>
Repetitions	Repetitions of words and phrases were transcribed, as presented in (14).	(14) <PREB2.ST0107> it's **less less** tedious (.) </PREB2.ST0107>
Word fragments	Example (15) illustrates how the use of hyphen was used to mark where a part of the word was missing.	(15) <PREB2.ST0403> well I don't I I think I don't agree with that because it's the **respon-** it's the responsibility of the children […] </PREB2.ST0403>
Anonymisation	Speakers' names were anonymised when participants involved in the interaction were addressed or referred to; see an example in	(16) <PREB2.ST0207> from my point of view I think there there is not is a good idea because

	(16). Their names were replaced by participant's code and put into square brackets.	you can copy or share something in internet and the teacher never know what you what do you think **[ST0202]**? </PREB2.ST0207>
Unintelligible speech	Unintelligible speech was not represented in the transcription process, involving no decisions made on how to represent utterances not understood completely. However, uncertain transcription of words that cannot be reliably identified was put in parentheses, as shown in (17).	(17) <PREB2.ST0104> […] if they can solve me the problem is good but if not I need to **(search)** information in the internet […] </PREB2.ST0104>

The transcription process underwent three stages to ensure the accuracy and reliability of the 80 oral discourse recordings. The first stage involved creating a basic transcription using alphabetic roman characters and decapitalisation, except for words that require capitalisation, such as proper nouns. British spelling was used to represent learners' naturally occurring speech, and the *Oxford English Dictionary Online* (OED) was used as the primary reference for standard spelling, hyphenation, and abbreviations. The first version of the transcripts included the complete oral production of all participants in the topic discussion.

In the second stage, the orthographic representation of the spoken discourse was divided into different transcripts separating the participants' utterances. The second result of transcription was then checked for accuracy. Finally, all transcripts were proofread and corrected to ensure the systematic application of transcription conventions across the oral output texts. Contextual information, such as the topic covered in the discussion and the questions asked by the examiner during the *Speaking* task, was added between curly brackets at the beginning of each transcript. For a transcript sample, see Figure 12 in Appendix K.

Each transcript was labelled with a code indicating the recording source (PREB2 for pre-test and POSTB2 for post-test), followed by the participant's code. For example, "PREB2.ST0206" corresponds to the transcript of participant six in the B2 class GR0102 (experimental group) in the pre-test. Furthermore, each transcript includes an additional source identification that specifies the discourse type, mode, provenance, text source, *Speaking* task, participant group, and participant code.

3.1.7. Concluding Remarks

This section (3.1) has provided an overview of the context and study design. The following key points are highlighted:

- The longitudinal, quasi-experimental study investigates the impact of incorporating metaphor awareness into distributed learning of metaphorical language in an EFL syllabus on the oral and written production of 40 Spanish L2 learners of English preparing for CEFR B2 level. In turn, it explores whether the benefits of metaphor use result in improved performance in standard high-stakes ESOL examinations.
- The study addressed the RQs by adopting a topic-based approach that focused on lexico-grammatical elements to analyse the use of metaphor in the learner discourse of participants from the control and experimental groups.
- The Ethics Committee at the University of Extremadura granted research ethics approval. Parental or legal guardian authorisation was obtained, and participants signed informed consent forms. All data collected for the study were anonymised.
- The study employed three different instruments, each serving specific research purposes to address three research questions: (a) the topic-based textbook used in class was employed to conduct the teaching treatment; both (b) pre-testing, and (c) (delayed) post-testing measures (consisting of two *Cambridge* B2 level mock exams) were used to assess participants' achievement at B2 level and identifying metaphor use in learner discourse before and after the longitudinal study.
- Data included participants' oral and written outputs collected from the B2 *Speaking* and *Writing* tests covering broad topics. No specific criteria were applied for the length of the 160 collected texts, totalling 30,280 words. An adapted version of the VOICE (2021) transcription procedure was used to transcribe non-native spoken English.
- The study lasted for five months and comprised three research stages, namely, pre-testing, teaching treatment, and post-testing, which were adapted to address the impact of COVID-19 during the study period.

3.2. TEACHING TREATMENT: METAPHOR-MEDIATED INSTRUCTION

The aim of the teaching treatment was to enhance the semantic potential of new words introduced in the L2 classroom (MacArthur, 2010). Further, the study sought to foster learners' autonomy in effectively handling newly encountered metaphors in English over the learning process (Boers, 2013), thus increasing their chances of success in high-stakes ESOL exams.

B2 preparation classes followed a goal-oriented approach where all participants had access to the same materials and resources in line with communicative language teaching (Richards & Rodgers, 2001). In contrast, the experimental group received an alternative method of vocabulary instruction within this approach, which included metaphor awareness integrated into classroom activities through CL-style techniques. The treatment teacher acted as a mediator, providing explicit guidance and instruction using CL-inspired methods in the L2 classroom. In metaphor-mediated instruction, learners were encouraged to explore the deeper meaning of lexical items and to think about why the L2 expresses things the way it does.

Concurrently, the control group relied solely on the textbook's construction-based vocabulary activities, which included Spanish equivalents and/or synonyms for enhanced understanding. As described in Section 3.1, the control group's instructor was unfamiliar with the thematic connection of metaphors or the motivation of language use (Lakoff, 1993; 1997; Lakoff & Johnson 1980, 1996). Thus, CL tenets were absent from the L2 vocabulary instruction given to the control group, which was exposed to other techniques following a communicative language teaching approach but excluding opportunities for in-depth learning of vocabulary.

For example, in Unit 11 "Medical Matters", participants from the experimental group, explicitly guided by their teacher, were made aware of the semantically motivated relationships in metaphorical language, such as *keep* (verb) and *fit* (adjective), within the collocational phrase "keep fit". They cognitively engaged with both the basic and extended meanings of such metaphors using a variety of CL-oriented methods (see below for further details). In contrast, participants from the control group followed a more conventional approach, focusing on textbook communicative activities without integrating metaphor awareness, thus not

emphasising the deeper metaphorical or semantically motivated relationships of language to enhance their understanding.

It is important to note that the present study sought to support L2 learners in using English-specific conventional metaphors that are targeted at the B2 level. This focus differs from enhancing creative metaphor use, which is typically expected of advanced learners with language abilities at the C1 and C2 levels, as previously observed (see Littlemore et al., 2014; O'Reilly & Fan, 2023) and as specifically described in the latest edition of the CEFR standards targeted at the advanced levels (see Section 2.3.3 of Chapter 2).

The following sections provide an overview of the treatment teaching method and its procedure. Section 3.2.1 describes the design criteria for developing CL-inspired activities, while Section 3.2.2 provides an account of the procedure for implementing these activities with the experimental group, including details about the time and amount of exposure. Section 3.2.3 concludes with some remarks.

3.2.1. Designing CL-Inspired Activities Adapted to an EFL Syllabus

MacArthur (2010, p. 168) highlights that raising metaphor awareness in the L2 classroom largely depends on the resources available to learners for exploring metaphors, the tasks they undertake and the quality of feedback they receive.

In my role as EFL instructor and researcher for this study, I developed pedagogical activities grounded in CL principles. They were designed to enhance learners' ability to notice metaphors in L2 language use and to understand the motivation behind the metaphorical meaning of vocabulary. In the following subsections, I explain the specific criteria that informed the design of these CL-inspired activities, which were tailored to suit the EFL syllabus.

3.2.1.1. Development of CL-Inspired Activities

The CL-inspired activities were developed using communicative language methodology with an input-driven and usage-based approach (Robinson & Ellis, 2008). The

lesson plans aligned with the EFL syllabus and transformed communicative tasks in the topic-based units into meaning-oriented activities.

The CL-inspired activities were designed following the widely recognised Presentation-Practice-Production (PPP) approach, which is known for its flexibility and adaptability to diverse teaching contexts (Harmer, 2001). This teaching model is structured in three stages: (a) the *Presentation* stage introduces the target language; (b) the *Practice* stage provides learners with opportunities to practice the language; and (c) the *Production* stage involves freer practice activities that encourage learners to use the target language spontaneously.

The PPP approach emphasises meaning-based learning, which provides learners with various opportunities to engage deeply with metaphors and use them in context. The CL-inspired activities were designed to encourage L2 learners to explore the use of metaphorical language in situated learning, consistent with the recommendations set forth by Low (1988). During the *Presentation* stage, the EFL teacher primarily employed CL-oriented methods to raise learner awareness, with students encouraged to apply these techniques in the controlled *Practice* stage. While the *Production* stage may not always elicit the target language as intended, the PPP approach offers learners a systematic and dynamic learning experience that allows them to develop their language skills and apply the target language in communicative contexts. While Masuda (2023) suggests adding a creative practice phase, the decision to omit this stage aligns with the EFL syllabus guidelines.

3.2.1.2. Distributed Learning of Metaphor

The topic-based textbook provided an ideal foundation for designing activities focused on metaphorical themes related to the lexical items covered in each unit dealing with abstract topics (MacArthur, 2010). The Conceptual Metaphor Approach (CMA) was used to explain the link between vocabulary and the underlying CM themes. When possible, learners were exposed to the CM relevant to the textbook topic to enhance their comprehension of metaphors in the oral and written input texts. Following Littlemore (2009), the key target vocabulary was organised and taught within frames to foster a deeper understanding of the

language. For instance, the CM CAREER IS A BUILDING was used to support L2 learners in dealing with the metaphorical language encountered throughout Unit 8, “Dream of the stars”, which focuses on the topic of *careers and aspirations* (see Figure 13 in Appendix L).

3.2.1.3. Target Lexis

Metaphor-mediated instruction consistently focused on general metaphorical language from the input texts of the textbook. The metaphor-driven activities aimed to engage L2 learners cognitively and deepen their understanding of metaphorically motivated English language use (see Boers, 2013; MacArthur, 2010). These activities included linguistic elements such as polysemes, phrasal verbs, collocations, and idioms that were relevant to the topics covered in the textbook tasks at B2 level.

Given the lack of clear guidelines on teaching metaphors (Low, 2008), particularly at different CEFR levels, I carefully selected metaphors from the oral and written input texts provided in the B2 level textbook. To identify instances of metaphorical language, I conducted a thorough analysis of these texts using the MIPVU method (Steen et al., 2010).[18]

To ensure that the selected language targets would be appropriate for learners’ proficiency level and learning outcomes, I applied four key criteria: (a) topic relevance, (b) frequency of occurrence in the textbook, (c) task support, and (d) feasibility of teaching using CL-oriented approaches (Boers & Lindstromberg, 2008b; MacArthur, 2017; Littlemore, 2023). For example, some metaphors selected for raising metaphor awareness in Unit 8, “Dream of the stars,” included the polyseme *star*, the phrasal verb *climb up*, the collocational phrase *build a career*, and the idiom *hit the big time*.

3.2.1.4. CL-Oriented Methods

To enhance metaphor awareness among L2 learners, I integrated CL-oriented pedagogical approaches into regular B2 training activities. Following MacArthur’s (2010)

[18] Section 3.3 of this chapter explains the metaphor identification procedure in detail.

proposal, I used a range of CL-oriented techniques to accommodate different types of metaphors and learning styles. For this purpose, I combined CL-oriented methods that have proven effective (see Boers, 2011, 2013).

In my study, pictorial elucidation and TPR were "used regularly in the classroom to support the verbal explanations of how concrete scenarios may motivate metaphorical uses of words and phrases" (MacArthur, 2010, p. 166). Pictorials were used as clues to interpret metaphors both before and after verbal explanations to help learners understand the meaning of metaphorical language (Boers et al., 2008). Figure 14 in Appendix L illustrates a visual prompt used in Unit 9, "Secrets of the mind", on the topic of *happiness and the mind* to elucidate the meaning of *cheer up* and *feel down* underlying the CM HAPPINESS IS VERTICALITY.

L2 learners were occasionally asked to create their own drawings to promote deep processing of the selected metaphorical language with mental imagery (Boers & Lindstromberg, 2006, 2008b). Figures 15 and 16 in Appendix L present examples of learners' meaning-meaning connections created by breaking down the meaning of the constituent words of the phrase *walking on air* (Unit 9, "Secrets of the mind") and establishing meaningful associations between the metaphorical sense of the individual words and their basic senses in the L2 (MacArthur, 2010; Wang et al., 2020).

TPR was used as a supplementary technique to complement other methods, considering that not all metaphors allow for physical imitation (Casasanto & Gijssels, 2015; Gibbs, 2021). Besides, it is important to note that some teenage learners may feel self-conscious about participating in physical activities (see Piquer-Píriz & Martín-Gilete, forthcoming). Hence, in this study, the EFL instructor acted out the metaphors and/or used physical objects to promote learners' metaphor awareness. To enhance comprehension of how happiness can be viewed in terms of verticality in English, for instance, the instructor drew an upward movement smile in the air with her hands to depict the CM HAPPY IS UP and a downward movement smile to illustrate the CM SAD IS DOWN.

3.2.1.5. Metalanguage

Metaphor-mediated instruction certainly requires the use of metalanguage in the L2 classroom. Therefore, it was crucial to make informed decisions about *what* linguistic terms to use, *how* to present them, and *when* to introduce them (Low, 2008; MacArthur, 2010).

In my study, I employed CL linguistic terminology related to metaphoricity and followed the recommendations of Boers & Lindstromberg (2006, 2008b) and MacArthur (2010) by presenting key concepts, such as “metaphor”, “concret” vs. “abstract” concepts, and “basic” vs. “extended” meanings, in a learner-friendly way (see Figures 17–20 in Appendix L). I introduced the metalanguage at the beginning of the teaching treatment, avoiding complex explanations or specialised terminology typically used by metaphor scholars (see Section 3.2.2 of this chapter for more details).

3.2.1.6. Timing

Previous research has highlighted the impact of class schedules on learning outcomes (Condon, 2008; Piquer-Píriz & Martín-Gilete, forthcoming). To address this issue, I carefully scheduled CL-inspired activities during specific moments of the treatment sessions. Since participants attended their English lessons late in the evening, I allocated the cognitive effort required for deep processing of metaphors to the lesson’s beginning when learners’ attention levels were generally higher. Furthermore, I developed CL-inspired activities with teaching and learning routines that fostered metaphor awareness in short-term slots to prevent fatigue and sustain attention for both the instructor and L2 learners (cf. Pan, 2019).

3.2.1.7. Learning Environment

Learner-led approaches (e.g., Boers & Demecheleer, 2001), teacher-led methods (e.g., Boers et al., 2004b), teacher-student(s) interaction (e.g., Littlemore, 2004), and peer(s)-to-peer(s) interaction were deemed essential types of learning situations for raising learners’ metaphor awareness in the L2 classroom.

This study took into account the classroom context to design CL-inspired activities. The small class size enabled the instructor to effectively monitor learners' cognitive engagement in a student-centred learning environment that fostered active participation throughout the teaching sessions. Participants engaged in individual, pair, small-group, and whole-group activities, all within their familiar L2 classroom environment.

Littlemore and Low (2006a) point out that "experimenting with figurative language is a potentially embarrassing and face-threatening activity" (p. 201). Therefore, before reporting their exploration of the metaphorical potential of the target items back to the class as a whole, students initially worked in pairs or small groups. Combining these work strategies facilitated learners' deep processing and helped learners refine their reasoning and interpretation through group negotiation and discussion involving both students and the EFL teacher (MacArthur & Littlemore, 2008). This means that both teacher and peers contributed to raising metaphor awareness, emphasising the collaborative effort in the learning process.

3.2.1.8. Materials and Resources

To enhance metaphor awareness in CL-inspired activities, I incorporated widely recognised proposals for CL-oriented pedagogies (e.g., Boers & Lindstromberg, 2009; Lazar, 2003; Rudzka-Ostyn, 2003; Lindstromberg & Boers, 2008). Figures 21 and 22 in Appendix L show how collocations with *career*, *ambition*, *experience*, and *job* were presented in Unit 8 using CL-style techniques from these proposals.

Furthermore, I consulted reference works like *Cambridge Idioms Dictionary* (2006) and *Oxford Idioms Dictionary for Learners of English* (2006) to learn about the origins of the selected language. I used these materials to provide etymological explanations whenever possible to promote an understanding of the motivated nature of language, especially in the case of idioms, from a diachronic perspective (Boers et al., 2004b). I consulted the *MetaNet Metaphor Wiki*[19] to identify the dominant CM for each topic and used *Microsoft PowerPoint*

[19] The *MetaNet Metaphor Wiki* is an open-access repository of metaphors and frames located at: https://metaphor.icsi.berkeley.edu/pub/en/index.php/MetaNet_Metaphor_Wiki.

software (2021) to present these context-specific CMs. Besides, I employed the online graphic design tool *Canva*[20] to create visual aids that illustrated the use of metaphorical language.

To foster growing autonomy in engaging with metaphors over the learning process, L2 learners were encouraged to consult the MM on their smartphones during the CL-inspired activities. This resource aided them in obtaining contextual information to uncover the sense of metaphorically used words (MacArthur, 2010). Similar to Boers et al.'s (2023) study on conventional multiword expressions, this research went beyond mere awareness raising by guiding L2 learners to use resources such as online dictionaries, thus increasing their confidence in identifying metaphors. However, due to their developing competencies, learners did not use electronic corpora to check the applicability of the vocabulary they planned to use in their assignments (cf. MacArthur & Littlemore, 2008).

3.2.2. Implementation of CL-Inspired Activities Adapted to a Real TEFL Context

The teaching treatment spanned over 15 weeks and involved 32 sessions conducted in English using metaphor-mediated instruction. The CL-informed treatment consisted of two distinct phases, each with a specific objective. These phases are described in detail below.

3.2.2.1. Phase 1 of Teaching Treatment

The initial treatment period aimed to introduce L2 learners to metaphor and CL-oriented pedagogical methods, in line with Low's (2017, p. 254) recommendations for essential training in small-scale studies. Based on one of the lesson plans designed by Clandfield (2003) on the CM LIFE IS A JOURNEY, learners were taught that metaphors in language serve more than just decorative purposes. This approach followed Boers' (2004, p. 211) proposal for enhancing metaphor awareness, which includes: (a) recognition of metaphor as a common ingredient of everyday language; (b) recognition of the metaphoric themes (conceptual metaphors or source domains) behind many figurative expressions; (c) recognition of the non-

[20] The website for *Canva* is https://www.canva.com/.

arbitrary nature of many figurative expressions; (d) recognition of possible cross-cultural differences in metaphoric themes; and (d) recognition of cross-linguistic variety in the linguistic instantiations of those metaphoric themes.

Over four weeks, I implemented four CL-inspired activities in a total of eight sessions, with each activity being taught across two full sessions. The initial phase was conducted at a slower pace than regular classroom activities to allow participants to gradually become familiar with the ubiquitous nature of metaphor in language and its culture-relatedness to avoid or mitigate transfer effects leading to problems in communication (Boers et al., 2004a; Littlemore, 2003, 2010; MacArthur, 2016c). Participants were also taught the necessary metalanguage for the second treatment period in a learner-friendly way, as described earlier. This was particularly important since, due to the teaching in their Spanish language and literature classes in secondary education, participants had traditionally been made aware of metaphors only as a rhetorical device.

Table 5 below provides an outline of the CL-inspired activities that were implemented during the first phase of the teaching treatment. Figures 23–28 in Appendix L provide examples of the materials used to implement these activities during Phase 1 of teaching treatment.

Table 5

Outline of Teaching Treatment – Phase 1: Introduction to Metaphor

CL-inspired activities	Timing	Sessions	Phases	Type	CL-oriented methods	CM	Target lexis
Activity 1: "The journey of learning English"	Week 1	2	Presentation	Teacher-led	Pictorial elucidation; TPR; guessing strategies; verbal explanation.	LIFE IS A JOURNEY	Traveller; destination; route; hurdle; guide; progress; landmark; crossroad; provision.
Activity 2: "Life is a journey"	Week 2	2	Practice (oral and written)	Student-led	Guessing strategies; conceptual grouping.		Go nowhere; go downhill; be lost; spin your wheels; reach a dead-end; be

Activity 3: "The journey of my life"	Week 3	2	Production (oral)	Student-led	-	over the hill; be at a crossroads; miss the boat; cross the bridge when it comes to it; arrive; go through; move on; pass away; start off on the wrong foot; to take an expected direction.
Activity 4: "Is it worth learning a foreign language?"	Week 4	2	Follow-up (essay)	Student-led	-	Hurdles; crossroads; progress.

3.2.2.2. Phase 2 of Teaching Treatment

During the second treatment period, I incorporated CL-style instruction into the existing EFL syllabus. Over two months, L2 learners engaged in 24 CL-inspired activities — one per session — and completed eight follow-up homework activities, covering four topics from the textbook. It is important to note that CL-inspired activities developed in class were designed to complement, not replace, the communicative tasks within each lesson. Accordingly, I adapted the syllabus to include six lessons per topic, delivered over a two-week period.

During the first week, I used *Reading* and *Listening* texts to raise metaphor awareness and help them understand the basic and extended senses of selected words. Metaphor-mediated instruction combined teacher-led explanations and students' own motivations to promote a deeper understanding of metaphors (MacArthur, 2010). However, the need for lexical guidance and corrective feedback remained substantial during the treatment (Boers, 2011; Low, 1988; Skoufaki, 2008). L2 learners received feedback on whether or not the metaphorical language they had produced was conventional in English. Figure 29 in Appendix L illustrates how the CM TIME IS MONEY was presented to L2 learners, aimed at providing support for dealing with the metaphorical uses of verbs such as *spend*, *waste, save or invest* within the context of the *Reading* and *Listening* comprehension tasks of Unit 10, "Spend, Spend, Spend?", which explores the topic of *shopping and leisure*.

The second week aimed to enhance learners' awareness of the potential exploitation of metaphorical language in their speech and writing. To illustrate, I implemented a CL-inspired activity in the *Production* phase to provide students with more opportunities to use the idioms learnt in Unit 11, "Medical Matters", on the topic of *health and fitness* (see Figure 30 in Appendix L). Like Phase 1 of teaching treatment, the classroom activities that focused on the motivated nature of language were further reinforced by assigning follow-up homework written tasks (see Figure 31 in Appendix L). It is important to note that learners were allowed to use dictionaries to promote independent use and sensitivity to metaphor (Boers, 2023).

The CL-inspired activities implemented across the four units taught during Phase 2 of treatment are presented in Tables 6–9 below. In addition to these planned activities, I adopted an incidental approach as part of metaphor-mediated instruction. This more incidental fashion involved drawing attention to metaphorical language that arose unexpectedly in the communicative activities and was deemed essential for successful B2 training (Boers, 2011).

Table 6

Outline of Teaching Treatment – Phase 2: Unit 8 "Dream of the Stars" (Careers and Aspirations)

Timing	Lesson	CL-inspired activities	Phases	Type	CL-oriented methods	CM	Target lexis
Week 1	1	Activity 1: "Raising metaphor awareness: Success"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	-	Star; go out in a blaze of glory; hit the big time.
	2	Activity 2: "Comparing stars"	Practice (oral and written)	Student-led	Guessing strategies; verbal explanation.	-	
	3	Activity 3: "YouTube millionaire celebrities"	Production (oral and written)	Student-led	-	-	
Week 2	4	Activity 4: "Career is a building"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	CAREER IS A BUILDING	Collocations with *career*, *ambition*, *job*, and *experience*: make, pursue, achieve, fulfil, realise, gain, get, find, apply for, look for, build, offer, launch.
	5	Activity 5: "Building careers"	Practice (oral and written)	Student-led	Guessing strategies; conceptual grouping; verbal explanation.		
	6	Activity 6: "Talking about the dream of the stars"	Production (oral)	Student-led	-		
Homework		Activity 7: "Pros and cons of being famous"	Follow-up (essay)	Student-led	-	-	-
		Activity 8: "TV talent contests"	Follow-up (article)	Student-led	-	-	-

Table 7

Outline of Teaching Treatment – Phase 2: Unit 9 "Secrets of the Mind" (Happiness and the Mind)

Timing	Lesson	CL-inspired activities	Phases	Type	CL-oriented methods	CM	Target lexis
Week 1	1	Activity 1: "Raising metaphor awareness: Personality"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	-	Key; come to terms with; walking on air.
	2	Activity 2: "Listening to the secrets of the mind"	Practice (oral and written)	Student-led	Guessing strategies; verbal explanation.	-	
	3	Activity 3: "The ingredients of happiness"	Production (oral and written)	Student-led	-	-	
Week 2	4	Activity 4: "Happiness is verticality"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	HAPPINESS IS VERTICALITY	Someone's spirits rise; fall into depression; cheer up; feel as high as a kite; be on cloud nine; be in the depths of despair; boost someone's spirits; feel down; be in seventh heaven; be over the moon; someone's heart sinks; lift someone's spirits; be in high spirits; be in low spirits; be on top of the world.
	5	Activity 5: "The ups and downs of happiness"	Practice (oral and written)	Student-led	Guessing strategies; conceptual grouping; verbal explanation.		
	6	Activity 6: "Measuring happiness"	Production (oral)	Student-led	-		
Homework		Activity 7: "Pursuit of happiness"	Follow-up (essay)	Student-led	-	-	-
		Activity 8: "Childhood memories"	Follow-up (article)	Student-led	-	-	-

Table 8

Outline of Teaching Treatment – Phase 2: Unit 10 "Spend, Spend, Spend?" (Shopping and Leisure)

Timing	Lesson	CL-inspired activities	Phases	Type	CL-oriented methods	CM	Target lexis
Week 1	1	Activity 1: "Raising metaphor awareness: Shopping"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	-	Tight; be at a knockdown price; spend money like water.
	2	Activity 2: "The flow of tightness"	Practice (oral and written)	Student-led	Guessing strategies; verbal explanation.	-	
	3	Activity 3: "Knocking down the prices"	Production (oral and written)	Student-led	-	-	
Week 2	4	Activity 4: "Time is money"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	TIME IS MONEY	Spend; waste; run out; use up; buy; afford; spare; worth; valuable; save; invest; precious; account for.
	5	Activity 5: "It's time to use money"	Practice (oral and written)	Student-led	Guessing strategies; conceptual grouping; verbal explanation.		
	6	Activity 6: "Making a profit of your time"	Production (oral)	Student-led	-		
Homework		Activity 7: "Shopping locally vs. shopping online"	Follow-up (essay)	Student-led	-	-	-
		Activity 8: "How to enjoy yourself"	Follow-up (review)	Student-led	-	-	-

Table 9

Outline of Teaching Treatment – Phase 2: Unit 11 "Medical Matters" (Health and Fitness)

Timing	Lesson	CL-inspired activities	Phases	Type	CL-oriented methods	CM	Target lexis
Week 1	1	Activity 1: "Raising metaphor awareness: Health"	Presentation	Teacher-led	Guessing strategies; verbal explanation; TPR; pictorial elucidation.	HEALTH IS VERTICALITY	Fit; face problems; feel under the weather.
	2	Activity 2: "A visit to the doctor"	Practice (oral and written)	Student-led	Guessing strategies; verbal explanation.		
	3	Activity 3: "Can I have a look?"	Production (oral and written)	Student-led	-		
Week 2	4	Activity 4: "Idioms: Health & fitness"	Presentation	Teacher-led	Pictorial elucidation; guessing strategies; verbal explanation.	-	Be rushed off your feet; be off-colour; rub shoulders with someone; be thrown in at the deep end; get to the bottom of something; be/feel under the weather; be taken aback.
	5	Activity 5: "What do you mean?"	Practice (oral and written)	Student-led	Guessing strategies; conceptual grouping; verbal explanation.	-	
	6	Activity 6: "What's the matter?"	Production (oral)	Student-led	-	-	
Homework		Activity 7: "Healthy habits"	Follow-up (essay)	Student-led	-	-	-
		Activity 8: "Staying fit"	Follow-up (article)	Student-led	-	-	-

Throughout the treatment sessions, learners were encouraged to use metaphors in the L2 classroom, and positive reinforcement was used to promote their metaphor awareness. The target vocabulary was displayed on a notice board for a two-week period, and a point system was implemented to motivate and track learners' use of the target lexis. Under this system, learners were awarded one point for using the basic meaning of the target word and two points for using the extended meaning in context.

Metaphor awareness was also promoted by displaying the five CMs taught on the wall. These were arranged to illustrate the analogy of "learning English as a journey" toward reaching B2 level (see Figure 32 in Appendix L). Furthermore, the tree analogy was used to foster cognitive engagement and assist students in comprehending morphology contents critical for B2 level preparation. By using this analogy, learners were supported to understand linguistic terminology for the derivation and inflection of English words. Specifically, the basic word forms were associated with the *roots*, while parts of speech (nouns, verbs, adjectives, and adverbs) were compared to the *branches,* and the prefixes and suffixes were linked to the *leaves* (see Figure 33 in Appendix L). Additionally, learners were introduced to timelines to illustrate the reference to time made by grammar tenses, i.e., past, present, or future (e.g., see Alonso-Aparicio & Llopis-García, 2019).

Detailed field notes were systematically made through teacher observation. After each session, the EFL instructor completed a form that recorded information on the efficiency and implementation of CL-oriented activities, as well as student attendance, reflections, and comments. This systematic record of CL-oriented activities provided feedback from the EFL teacher's perspective on learner performance and potential learning gains. Audio-recordings were made of the sessions to supplement the teacher's observation, although they were not transcribed for this study. Student evaluation of CL-oriented activities did not occur regularly, as the aim was to integrate CL-oriented approaches into the L2 classroom as seamlessly as possible. However, for data triangulation purposes, student evaluation of metaphor-mediated instruction was conducted at the end of the study.

3.2.3. Concluding Remarks

This section (3.2) has presented an overview of the teaching treatment method and procedure employed in metaphor-mediated instruction. In summary, the key points are:

- The teaching treatment aimed to promote the English vocabulary growth of L2 learners by providing support to understand the semantic potential of new words and develop autonomy when dealing with conventional metaphors at B2 level.
- The experimental group was made aware of metaphor through a CL-style approach integrated into regular classroom activities from the B2 textbook. In contrast, the control group did not receive metaphor-mediated instruction. They engaged in construction-based vocabulary activities from the same mainstream material, following the standard communicative approach but with no opportunities for in-depth learning of vocabulary.
- CL-inspired pedagogical activities were developed based on specific design criteria and were tailored to the EFL syllabus and the real instructed L2 setting. Specifically:
    - The PPP approach was employed to facilitate meaning-based learning and provide opportunities for language use in communicative contexts.
    - Distributed learning of metaphors was enhanced by the topic-based textbook, which focused on metaphorical themes related to B2 level lexical items.
    - Metaphor-mediated instruction focused on general metaphorical language from the oral and written input texts, which was identified using the MIPVU method. The metaphorical language selected aligned with the EFL syllabus, instructed L2 setting, and CL-style techniques.
    - CL-oriented methods were integrated into regular B2 activities, using various techniques to adapt to different types of metaphors and learning styles.
    - CL linguistic terminology related to metaphoricity was introduced in the CL-inspired activities in a learner-friendly way.
    - CL-inspired activities were designed in short-term slots to sustain attention and prevent fatigue and scheduled at the beginning of sessions.

  - The student-centred learning environment encouraged active participation and effective monitoring of deep cognitive processing by the L2 learners.
  - Various materials and resources were used to design the activities, including proposals for CL-oriented pedagogies, reference works, and software.
- A total of 28 CL-inspired pedagogical activities were implemented in 32 sessions over a period of three months across two different CL-informed treatment phases:
  - Phase 1 aimed to introduce L2 learners to metaphor and CL-oriented pedagogical methods.
  - Phase 2 aimed to introduce metaphor awareness into the EFL syllabus (four units), promoting learners' ability to notice metaphors in L2 language use and helping them understand the underlying motivation of metaphorical language.

3.3. ANALYSING METAPHOR IN LEARNER DISCOURSE: METHODOLOGICAL CONSIDERATIONS

This section provides an overview of the method employed to analyse the use of metaphor in learner discourse, structured as follows. Section 3.3.1 explains the procedure adopted for metaphor identification and the reasoning behind the analytical decisions made during the analysis of participants' oral and written outputs. Section 3.3.2 reports on the reliability of the empirical procedure used for metaphor identification. Section 3.3.3 describes additional data treatment, and Section 3.3.4 provides some concluding remarks.

3.3.1. Identification of Uses of Metaphor

As anticipated in Chapter 2, metaphor identification involves analysing linguistic expressions to determine their metaphorical usage. In this study, the MIPVU procedure (Steen et al., 2010), a widely accepted and theoretically valid approach based on the MIP procedure

developed by the Pragglejaz Group (2007), was employed to systematically analyse metaphors in spoken and written learner discourse.[21]

To accommodate the adopted lexico-grammatical approach and the heterogeneity of L2 language use, this study slightly modified MIPVU (L2-MIPVU). The adapted version retained the basic procedure to enable comparison while addressing the specific analytical challenges posed by the research. The identification of uses of metaphor consisted of three analytical phases to ensure the consistent application of MIPVU. First, a systematic approach was developed for annotating data in the pre-analysis stage. Subsequently, the MIPVU procedure was applied step-by-step in the analysis stage, including troubleshooting methods to address language-specific issues experienced by L2 learners. Lastly, a final cleaning-up process was conducted in the post-analysis stage.

3.3.1.1. Pre-Analysis Stage: Setting up the Data Annotation System

In the initial phase of the analysis, practical tools were established for data annotation. Manual decisions were organised based on MIPVU guidelines for metaphor analysis to simplify the metaphor identification process. Further streamlining of the process was achieved by performing a series of actions during the pre-analysis stage, including (a) registering the data, (b) the Part-of-Speech (PoS) tagging process, and (c) cleaning up the data for accuracy along with preliminary tokenisation.

[21] The study likewise employed MIPVU (Steen et al., 2010) to identify metaphors in the input sources, specifically the 27 texts (containing 7,885 LUs) collected from the textbook (7,595 LUs) and testing measures (290 LUs). This analysis served two purposes: selecting the target metaphorical language used for metaphor-mediated instruction in the teaching treatment (see Section 3.2.1 of this chapter) and for MIPVU training purposes (see Section 3.3.2 of this chapter). It is important to note that input data was not subjected to any further processing for the purposes of this study. However, to provide a clearer understanding of the metaphor identification procedure, examples from the input analysis will be used in conjunction with those from the output analysis.

3.3.1.1.1. Data Registration Setting

To analyse metaphors in this study, *Microsoft Excel* software (2021) was used for data entry and coding. This analytical tool was selected due to its user-friendly data visualisation format and ease of data manipulation, facilitating additional quantitative analyses.

The 160 collected output texts were represented in a vertical column (column A) using an *Excel* spreadsheet, where each LU was identified by a reference and placed in a different cell. The adjacent columns (columns B–R) were used to enter supplementary research data, morphological information about the LUs, records of learner errors and unconventional constructions, and decisions about metaphor status. For an example of how the data was organised and the annotation format employed in this study, see Figures 34 and 35 in Appendix M, which also includes a description of the purpose of each column during the MIPVU procedure.

3.3.1.1.2. PoS-Tagging Process

To ensure the accurate identification of LUs, a complete semantic demarcation was performed before the analysis. The PoS tagging followed MIPVU suggestions and used the *Constituent Likelihood Automatic Word-tagging System* (CLAWS), an automatic annotation tool provided by the *Wmatrix5* software[22] (Rayson, 2008).

As part of the PoS-tagging process, all words were assigned their corresponding PoS and listed vertically as separate LUs. The PoS-tagged version of the 160 texts was downloaded, and the accompanying PoS information and lemmas were automatically entered into the corresponding columns of the *Excel* spreadsheet: “Lexical Unit” (column C), “Lemmas” (Column D), and “Part of Speech” (column F).

To ensure the reliability of the PoS tagging, a further check of LUs was conducted based on this tool’s reported 96–97% accuracy rating (Rayson, 2008). Although the PoS tagging of the 160 texts was generally effective, additional manual checks were carried out to

[22] *Wmatrix* can be accessed through the following link: http://ucrel.lancs.ac.uk/wmatrix/.

identify and correct any potential inconsistencies that may have arisen due to the L2 discourse nature of texts, as described in the following subsection.

3.3.1.1.3. Cleaning-up Process and Preliminary Tokenisation

To minimise manual decision-making in analysing the LUs, a series of cleaning-up actions were performed along with preliminary tokenisation before determining metaphor status. The cleaning-up process was designed to exclude cases that did not meet the criteria for separate LUs, as suggested by Steen et al. (2010, pp. 167–172). The cleaning-up process involved several steps, as follows.

First, I removed section breaks and punctuation marks from the *Excel* cells in the dataset. Next, unlike the MIPVU method, I replaced symbols and numbers with letters. For instance, "£" was replaced by "pound" or "4" by "four". Additionally, I excluded 40 cases of Saxon genitive ('s) from the data entries but preserved their orthographic representation next to the corresponding LUs to enhance the understanding of the texts. Notably, the Saxon genitive cases were not considered part of the total LUs analysed.

As part of the *preliminary* tokenisation (as explained in detail in Section 3.3.1.2.2 below), each cell was checked to ensure it contained one orthographic word. If multiple words were present, extra cells were inserted to separate them. I also separated orthographical contractions into individual elements and placed them in different cells. For example, *don't* was treated as two separate LUs: "do" and "n't".

Finally, the accuracy of the PoS tagging and lemmas provided by *Wmatrix5* (Rayson, 2008) was verified by proofreading the 160 texts, amounting to a total of 31,077 LUs (oral discourse = 16,347 LUs; written discourse = 14,730 LUs). Proofreading was used as the first approach to text comprehension and familiarity with the discourse context. Modifications were made where needed based on the MIPVU criteria and the adapted version used in this study.

3.3.1.2. Analysis Stage: Operationalisation of the Metaphor Identification Procedure

This section provides an overview of how to operationalise the MIPVU procedure.[23] It begins by explaining the external sources of reference used for metaphor analysis following MIPVU, with a focus on the rationale for consulting different dictionaries to identify metaphors. Next, the MIPVU protocol is illustrated with examples to show how the procedure was applied during the analysis phase. In turn, this section describes the troubleshooting methods introduced to address the L2 language-specific issues that required consideration when applying the MIPVU protocols and tools in the present research.

3.3.1.2.1. Resources for Metaphor Analysis

As a non-native speaker of English, similar to the original MIPVU (Steen et al., 2010) researchers, identifying metaphors in English discourse can be challenging. My potential difficulty lies in not having the necessary knowledge provided by the dictionary entries. Therefore, I consulted several corpus-based dictionaries of contemporary English[24] to establish a solid foundation for identifying metaphorical language.[25] In addition, using multiple external sources of reference for metaphor analysis served various purposes when applying the MIPVU method, as explained below.

The primary reference resource used to identify the meanings of the LUs was the *Macmillan English Dictionary* online (MM; 2023),[26] a dictionary based on a relatively recent corpus that provides a description of current English. MM is noteworthy for its unique treatment of metaphorical meanings, which are identified with “Metaphor Boxes” containing notes on

[23] For a detailed explanation of the MIPVU protocol, see Steen et al. (2010, pp. 25–42).
[24] To determine basic and contextual meanings, this study selected British English when dictionaries provided different senses for the same word entry based on the variety of language used.
[25] The information about the dictionary used in the analysis was entered in “Column L” of the *Excel* spreadsheet.
[26] The MM was accessible at https://www.macmillandictionary.com/ during the investigation. However, it is important to acknowledge that Macmillan Education Ltd announced the closure of the *Macmillan English Dictionary* and blog websites on 30th June 2023. As a result, for reference purposes, please consult the printed version of the dictionary (e.g., Rundell, 2007).

metaphor awareness. This suggests that the authors of MM considered the use of metaphors in English when writing the definitions (Pragglejaz Group, 2007, p. 16).

The *Longman Dictionary of Contemporary English* online (LM; 2022) was employed to supplement MM in cases where these were insufficiently defined or did not include contemporary English senses. For instance, while MM defines *invest* as a monosemous verb meaning "to use your money with the aim of making a profit from it, for example by buying property or buying shares in a company" (MM1), it does not cover the sense of "to use a lot of time, effort etc or spend money to make something succeed" (LM3). Notably, the metaphorical senses identified in my analysis were fully conventional and were recorded in a dictionary.

In addition, LM was consulted to determine the meaning of multiword LUs, except for compounds. This was necessary because MM tends to conflate phrasal verbs and prepositional verbs into a single verb category. For example, MM identifies *belong to* as a phrasal verb, but it is actually a combination of a verb (*belong*) and a bound preposition (*to*), i.e., a prepositional verb.

The *Cambridge Dictionary* online (CAMD; 2022) was used in the rare cases where a LU was not attested in MM or LM, or to verify contemporary senses of a word that were not recorded in either MM or LM. For instance, the adjective *uncreative* and the noun *member* in the sense of "a leg or arm" were only found in CAMD.

MIPVU limits metaphor identification to metaphorical meanings that are relevant to contemporary language users. The *Oxford English Dictionary* online (OED; 2023) was consulted whenever it was challenging to make decisions on the metaphorical status due to an incomplete understanding of the LU. For example, in the case of the verb *cry*, as shown in (18), it was important to consult an etymological dictionary to understand the verb's historical development to determine its basic sense.

(18) <PREB2.ST0308> […] yes but the the teachers give us a lot of homework to do it and in my case I'm very stressed and I **cry** a lot for help because I don't have time to do it […] </PREB2.ST0308> (PREB2.SOUTPUT.S4.ST0308)[27]

In this case, both the first and second entries in MM for the verb *cry* ("to have tears coming from your eyes, especially because you are sad" and "to shout something", respectively), describe human-oriented senses. However, consulting the OED revealed that the second entry was historically older and thus considered the more basic meaning. The meaning of *cry*, referring to "utter[ing] the voice loudly and with exclamatory effort" (OED3) or to "[…] implore[ing], in a loud and […] excited voice" (OED1), was extended to the sense of "shed tears", even when no sound is uttered. This meaning was originally expressed by the verb *weep* but was replaced by *cry* via the notion of uttering a loud, vehement sound. Therefore, identifying the second entry in MM as a more basic meaning was significant because in (18), *cry* was being used in its basic sense and thus non-metaphorically used.

3.3.1.2.2. MIPVU Protocol and Troubleshooting Methods

The MIPVU procedure is a systematic approach that involves explicitly identifying potentially metaphorically used words as metaphors in discourse, using protocols across four consecutive phases. These phases include (a) identifying LUs in the text; (b) establishing the contextual meanings of the identified LUs; (c) determining the more basic meaning of the LUs; and (d) deciding whether the identified LUs are being used metaphorically or not.

Identifying metaphorical language in non-native-like English using the MIPVU procedure has been acknowledged as a challenging task (e.g., see MacArthur, 2019; Nacey, 2013). The original MIPVU procedure does not provide specific guidelines for L2 processing and language production research observations. This approach assumes that the "language user is the idealized native speaker of English as represented in the description of English by

[27] Examples extracted from learner English discourse are presented in this PhD dissertation in their original form, including any anomalies in their production.

the dictionary […] contain[ing] a complete and culturally sanctioned representation of the knowledge about the English lexicon" (Steen et al., 2010, p. 7).

This study applies a slightly adapted version of MIPVU to address the distinctive features of the linguistic data identified, which affected the four stages of the procedure. The relatively homogenous group of L2 English speakers sharing the same L1 and enrolled in the same proficiency language level courses allowed for principled decisions on how to deal with the L2 language-specific issues. In what follows, I describe the consistent criteria adopted and justify the conservative modifications introduced across the four stages of MIPVU: Phase 1 (identifying lexical units in the text); Phase 2 (establishing contextual meanings); Phase 3 (deciding about more basic meanings); Phase 4 (deciding about uses of metaphors).

Phase 1. Identifying Lexical Units in the Text

To begin the analysis, I thoroughly read the entire text to understand its overall meaning. Next, I conducted a word-by-word examination to identify the LUs present in the text. This approach proved effective in analysing texts of varying lengths in the present research. I systematically read and comprehended the context for each text before making analytical decisions, which was particularly important when analysing texts of different lengths but similar topics. I read each text carefully to avoid any biased understanding of its meaning.

Next, I will detail the special treatment of certain LUs (Section A) and the rationale behind discarding others in metaphor analysis (Section B).

A. Special Treatment of Lexical Units

Determining the basic unit of analysis is a crucial step in the MIPVU procedure since it affects the total count of LUs in the text and, therefore, the quantitative analysis of metaphor density.[28] MIPVU generally considers an orthographic word (a group of letters with spaces on

[28] The classification of the different types of LUs was recorded in "Column E" of the *Excel* spreadsheet.

either side) as an independent LU that can potentially be metaphorically used in context (see Steen et al., 2010, pp. 26–27).

As explained in detail below, the MIPVU method departs from treating word strings as a single LU in certain cases where multiword units, such as compounds, some proper nouns, phrasal verbs, and polywords, pose challenges for word-level analysis. In my study, I adapted the MIPVU guidelines on demarcating LUs to address L2 learner constraints in language use. These adjustments focused specifically on compounds, phrasal verbs, polywords, grammatical phrases, and numerals, as described below.

Compounds. In the MIPVU method, compound words that are listed in the dictionary as a single entry and spelled as one word or with two hyphenated words are treated as one LU. For example, the compound noun *sundress* in (19) and the compound adjective *off-colour* in (20) were both considered one LU.

(19) Yet even with all these distractions, I can't stop thinking about that pretty **sundress** in the window of the market. (TB10.WINPUT.R5)

(20) I've been feeling a bit **off-colour** for some time now […] (TB10.OINPUT.L3)

However, if a compound word is a novel formation and not documented in the dictionary, it is analysed as separate LUs. For example, *pay-per-click* in (21) was a novel formation consisting of three independent LUs that were not found as a single entry in the dictionary and were therefore treated as separate LUs in the analysis.

(21) The YouTube payment system works on a **pay-per-click** basis (TB8.WINPUT.R1)

(22) Oh me too who wouldn't be after all those **mid-year** tests we've been doing. (TB9.OINPUT.L1)

Similarly, hyphenated compound words that follow a productive morphological rule but are not considered conventionalised compounds in the dictionary, such as *mid-year* in (22), are also treated as two independent LUs.

Compound words that are spelled as two separate words are treated as a single LU in MIPVU if they are listed as a compound in the dictionary and pronounced with the stress on

the first item. For instance, *shopping centre* /ˈʃɒp.ɪŋ ˌsen.tər/ in (23) was considered one LU, whereas *mobile phone* /ˌməʊ.baɪl ˈfəʊn/ in (24) was treated as two independent LUs.

(23) Kerry particularly enjoys the **shopping centre**'s feeling of luxury. (TB10.OINPUT.L4)

(24) You overhear a boy calling his friend on his **mobile phone**. (TB9.OINPUT.L1)

It is important to note that suprasegmental features, such as stress placement, may not be considered relevant in L2 discourse or written discourse analysis (MacArthur, 2019). However, MIPVU was applied rigorously for the sake of meaningful comparison.

In the present study, the production of compound words by L2 speakers of English was generally not problematic when determining LUs using the MIPVU technique. However, spelling mistakes[29] occasionally did affect the application of MIPVU, as illustrated in (25), where the term *sport centre* was used instead of the correct form *sports centre*.

(25) <POSTB2.ST0304> yes we always go to the to the **sport centre** and we practise football or bask- or basketball at the weekend. </POSTB2.ST0304> (POSTB2.SOUTPUT.S4.ST0304)

According to the MIPVU method, the term *sports centre* /ˈspɔːts ˌsen.tər/ can be classified as a compound noun because it appears as a separate entry in the dictionary ("a building where you can go to do different sports and other activities" [MM1]) and is pronounced as a compound with the stress on the first syllable. However, in example (25), the production of *sport centre* was inaccurate as the final "s" was omitted, making it impossible to treat this as a compound word. Nonetheless, the approach to identifying compounds was not modified, as such special cases were rare in the present research. Instead, the component parts of such examples as *sport centre* were treated as two separate LUs.

Proper Nouns. MIPVU generally treats each word in a proper noun as a separate LU, as seen in the analysis of *The University of Chicago* in example (26). However, certain proper nouns bestowed on public entities (e.g., *St Andrews* in [27]) or titles (e.g., *Nobel Prize* in [28])

[29] Treatment of learner errors is fully explained further in this chapter.

are exceptions to this criterion. If these proper nouns were listed in the dictionary, they were treated as a single LU.

(26) On arrival at **The University of Chicago** fifty years ago I was disappointed to find that academic psychologists were trying to understand human behaviour by studying rats in a laboratory. (TB9.WINPUT.R5)

(27) He was discovered by someone from the theatre while performing in plays in Edinburg and **St Andrews** […] (TB8.WINPUT.R7)

(28) I felt incredibly lucky that I had the opportunity to work in laboratories where I rubbed shoulders with **Nobel Prize** winners. (TB11.WINPUT.R6)

It is important to point out that the metaphorical use of proper nouns is possible when the meaning of a proper noun has become conventionally associated with a common noun (Wee, 2006). However, such use is limited due to the culture-specific nature of these expressions (Lakoff, 1993). In my analysis, a total of 96 LUs were identified as proper nouns. Notably, *The Green Room* in (29) and (30) was the only instance related to metaphor.

(29) The contestants were asked to wait in **the Green Room** for the show to begin. (TB8.OINPUT.L2)

(30) Anyway, later, what he told me was that before the show he stood around with the other participants in somewhere called **'The Green Room'**, where they chatted to each other and were given something to eat and drink, and they got to know each other a bit. (TB8.OINPUT.L2)

In the analysis, the term *The Green Room* from the input text "TB8.OINPUT.L2" was identified as one LU based on its dictionary definition from MM1, which describes it as "a room where performers can relax in a theatre or television studio". The term originally referred to a backstage room in a theater where actors would relax and wait for their cues to go on stage, but over time it has come to be used more generally to refer to any room where people could relax and socialise away from public view. In examples (29) and (30), the term *The Green Room* is used metaphorically to refer to a space of anticipation before an important event or meeting. However, it was not tagged as metaphorically used in the analysis because the term

lacked a more basic meaning listed in the dictionary that it could be contrasted with to be understood in comparison with and be regarded as metaphorically in use.

Phrasal Verbs. Phrasal verbs are multiword verbal expressions consisting of more than one LU. However, MIPVU considers phrasal verbs that comprise a verb and an adverbial particle (e.g., *look up*) or those consisting of a verb followed by an adverbial particle and a prepositional particle (e.g., *look down on*) as a single LU. MIPVU argues that breaking down multiword verbal units into their constituent words can result in a loss of meaning, particularly when phrasal verbs are used metaphorically. For instance, the phrasal verb *turn down* can be used metaphorically to denote refusal in (31) instead of its non-metaphorical meaning of reducing the amount of sound. Therefore, MIPVU advocates analysing phrasal verbs as single LUs to capture their full meaning.

(31) If I am offered a job down in London, there's no way I'd ever be able to **turn** it **down**. (TB8.WINPUT.R7)

In my study, I introduced a criterion for decomposing phrasal verbs into two or three LUs to address L2 learners' difficulties in language use. Unlike native speakers of English, L2 learners often process phrasal verbs as full decomposable expressions or novel compounds rather than fixed chunks of language (MacArthur & Littlemore, 2011), for example, by inserting non-standard particles or adverbs (e.g., see Alejo-González, 2010).

Based on this observation, previous research on identifying metaphor use in English learner discourse (e.g., Littlemore et al., 2014; MacArthur, 2019) opted for breaking phrasal verbs into their component words, i.e., verb and adverb. In the present study, special attention was paid to identifying phrasal verbs as separate LUs to address these difficulties. The phrasal verb *strip away* in (32) exemplifies this departure from MIPVU. In my analysis, the lexical verb *strip* and the adverb *away* were considered as separate LUs rather than checking the meaning of the phrasal verb *strip away* itself.

(32) People will give you all sorts of advice about your acting which helps you to **strip away** your bad habits. (TB8.WINPUT.R7)

It is important to acknowledge that the approach of treating phrasal verbs as separate LUs affects the number of metaphors in the analysis when a phrasal verb has only one meaning. For instance, applying MIPVU strictly would identify *stood around* as not metaphorically used in example (33). The only meaning listed for the phrasal verb "stand around" in LM ("to stand somewhere and not do anything") covers both the contextual sense and the more basic sense. However, the approach used to identify phrasal verbs as two different LUs resulted in *stood* and *around* being tagged as metaphorically used for each case.

(33) Anyway, later, what he told me was that before the show he **stood around** with the other participants in somewhere called The Green Room. (TB8.OINPUT.L2)

The potential for decomposability also impacted prepositional verbs, i.e., verbs that consist of a verb and a bound preposition. For instance, consider how the preposition *at* is linked to the verb *look* in (34).

(34) Dad goes off to **look at** sports equipment or electronics while mum hunts for clothes. (TB10.OINPUT.L4)

The resources used for metaphor analysis in this study categorise prepositional verbs as "phrasal verbs," implying that these are considered as one LU in the dictionaries. However, the approach taken in this research treated the two elements of prepositional verbs as separate LUs based on their individual meanings.

Polywords. MIPVU generally applies a semantic decomposability criterion to analyse polywords, which are multiword expressions that form a lexical or grammatical unit and distinct from those discussed earlier. The analysis of the multiword expression *on the other hand* in example (35) serves as an illustration of this approach. However, some polywords, such as those listed in the *List of Multiwords and Associated Tags* (2000) developed by the *British National Corpus* (BNC)[30] and included in the *CLAWS PoS-tagging* system (Rayson, 2008), are considered as fixed LUs and cannot be decomposed. Hence, they are treated as a single

[30] The BNC list of polywords can be found at http://www.natcorp.ox.ac.uk/docs/multiwd.htm.

LU. For instance, in my analysis, *according to*, *such as*, and *per cent* in (36) were deemed non-decomposable since they appear on the BNC list.

(35) **On the other hand**, I'm someone who prefers doing things on his own like playing computer games […] (TB9.OINPUT.L1)

(36) **According to** a recent report, some products **such as** toys are as much as sixty **per cent** cheaper online. (TB10.WINPUT.R2)

Polywords, just like phrasal verbs, are processed differently by L2 speakers of English compared to L1 speakers. For instance, while the polywords *because of* and *instead of* were used correctly in example (37), the L2 learner used *on the top of* instead of *on top of*.

(37) **On the top of** that, being rushed of the feet will not be a problem **because of** writing on a computer **instead of** by hand. (POSTB2.WOUTPUT.W1.ST0103)

The addition of the article *the* to the polyword *on top of* raises concerns about the decomposability of fixed multiword expressions by L2 learners. Unlike phrasal verbs, identifying multiword LUs as fixed or not can be problematic when using additional analytical tests alongside the main metaphor identification procedure (see MacArthur, 2019, pp. 295–297). In my analysis, the BNC list of polywords was used solely as a reference in conjunction with MIPVU to identify multiword units as single LUs through MIPVU. However, when the constituent elements of polywords were used incorrectly, they were treated as separate LUs and submitted to metaphor identification, just like the fixed multiword units that are not listed in the BNC list of polywords.

Multiword Unit Verbs. Multiword unit verbs were considered as non-decomposable LUs. Therefore, verbal phrases such as the future tense construction "going to" in (38) and the multiword modal verb "have to" in (39) were treated as a single LU in the analysis.

(38) So I'm **going to** tell you about my dad's ten minutes of fame. (TB8.OINPUT.L2)

(39) As you grow older, you **have to** find new challenges which are more appropriate to your age. (TB9.WINPUT.R5)

Infinitive verbs with the marker "to" present some special cases that required attention. In this study, they were treated as one LU when they indicated the purpose or reason for doing something, as illustrated in (40), or as a complement to the main verb in (41), to a noun in (42), or adjective in (43).

(40) Every month advertisers have started focusing on the most popular video makers **to take** advantage of their loyal public. (TB8.WINPUT.R1)

(41) No one *wants* **to see** their high street shops disappear but why would people *want* **to shop** locally? (TB10.WINPUT.R2)

(42) There are a lot of opportunities up here, so it makes *sense* **to stay** closer to home. (TB8.WINPUT.R7)

(43) She didn't go because she was worried that she would be too *nervous* **to answer** any questions (TB8.OINPUT.L2)

Numerals. In my study, numeral determiners were consistently counted as a single LU, including ordinal and cardinal numbers. For example, the multiword numeral expression *one hundred million* was considered as an undecomposable LU in (44).

(44) With over **one hundred million** visitors to YouTube every month, advertisers have started focusing on the most popular video makers to take advantage of their loyal public (TB8.WINPUT.R1).

B. Discarded Lexical Units for Metaphor Analysis

The first stage of the MIPVU procedure involves deciding which linguistic forms should be included or excluded for the metaphor identification process. The topic-based approach adopted in this study led to the exclusion of concrete LUs.[31]

Identifying metaphors in L2 production presented additional challenges, such as learner errors and non-native-like phraseology, which made it difficult to determine which

[31] These decisions were recorded in "Column K" of the *Excel* spreadsheet.

occurrences were suitable for metaphor analysis.[32] Next, I will provide a detailed account of how I addressed all these challenges during the analysis while applying the MIPVU procedure.

Automatically Discarded Words for Metaphor Analysis. MIPVU uses a process of exclusion to identify metaphors, whereby LUs are discarded for metaphor analysis (DFMA) if they cannot be analysed. This may occur when words are truncated, as exemplified by *foot-* in example (45).

(45) <POSTB2.ST0403> I don't know in a **foot-** something with football or in a covered space no? don't you think? </POSTB2.ST0403> (POSTB2.SOUTPUT.S4.ST0403)

Incomplete utterances or lack of context, which are common in oral discourse, can pose difficulties for researchers in inferring the intended word from the context. Thus, problematic cases such as the use of *secondary* instead of the complete phrase "secondary education" in (46) are generally excluded from metaphor analysis.

(46) Firstly, in schools nowadays you have to study **secondary** obligatory while in the past you did not need it. (PREB2.WOUTPUT.W1.ST0106)

In this study, specific grammatical functions and non-semantically full words (cf. Deignan, 2005) were excluded from the analysis as they were arguably of lesser importance in the topic-based approach to discourse, as explained in detail below.

Multiword expressions listed in the BNC list of polywords, such as *even if*, *each other,* and *at least* in example (47), were excluded from metaphor analysis due to their grammatical nature, despite their potential for metaphorical use in discourse.

(47) I mean **even if** we don't get on with all of **each other**'s friends **at least** we put up with them and don't show it (TB9.OINPUT.L1)

Grammatical-function words such as the negation adverb "not", the verb "to be", the preposition "by" in passive voice constructions and all elements in the verb group of auxiliary and modal verbs were excluded, except for the lexical verb or head of the v-ing form.

[32] This information was recorded in the analysis from "Column G" to "Column J" of the *Excel* spreadsheet.

Delexicalised verbs such as do*, have*, *get*, *make*, or *take* have been argued to have lesser semantic content than lexical verbs (Cameron, 2003; Deignan, 2005). For example, *have* in the phrase “have a try” could be replaced by a more semantically rich single verb like “try” to refer to attempting to do something. However, this study included delexicalised verbs in metaphor identification to ensure a meaningful comparison with previous research.

The present research focused on open-class words, including nouns, verbs, adjectives, and adverbs, and excluded closed-class words, such as determiners, pronouns, and conjunctions. An exception was made for prepositions, as they can collocate with open-class words and were deemed to have the potential for metaphorical use.

“Bound” prepositions, unlike “free” prepositions, are those that depend on the lexical element before them rather than the prepositional object. For example, in the phrase *devoted his life to studying happiness* (48), the preposition “to” depends on the verbal form “devoted” rather than the prepositional object “studying happiness”. Similarly, certain nouns, adjectives, and adverbs can also depend on specific prepositions. For instance, the noun *impact* can collocate with *on* (49), the adjective *glad* with *about* (50), and the adverb *closer* with *to* (51).

(48) Mihaly Csikszentmihalyi has **devoted** his life **to** studying happiness. (TB9.WINPUT.R5)

(49) There are a variety of factors which could have an **impact on** this: better diets, our minds are stimulated more, or just getting more practice in tests of this type. (TB9.OINPUT.L1)

(50) Speaking personally, I am very **glad about** the actual education. (PREB2.WOUTPUT.W1.ST0103)

(51) There are a lot of opportunities up here so it makes sense to stay **closer to** home. (TB8.WINPUT.R7)

To ensure meaningful comparisons with previous research of a similar nature, both “free” and “bound” prepositions were included in the analysis. However, the prepositions “of” and “for” were discarded from metaphor analysis due to the difficulty in pinpointing their basic meanings, as explained in more detail later in this chapter.

The analysis excluded repetitions of LUs, such as *used to* in (52), or self-corrections, like *actually* in (53). Metaphorical use was only identified on the last occurrence of the same LU, while intended repetitions that clarified meaning were analysed. For instance, the phrase *my brother* in (52) was analysed.

(52) <PREB2.ST0103> well my mum **used to** *used to* help me a lot in the past but now she has to help **my brother** *my little brother* so she doesn't have a lot of time for me […] </PREB2.ST0103> (PREB2.SOUTPUT.S4.ST0103)

(53) <POSTB2.ST0101> it's a good idea now also for children that are **actually** *currently* they are always in front of a screen […] </POSTB2.ST0101> (POSTB2.SOUTPUT.S4.ST0101)

Other linguistic forms and special cases that were not relevant to the study were also omitted, along with the LUs discarded by MIPVU. In this study, acronyms (54), abbreviations (55), interjections (56), and foreign words (57) were automatically excluded from metaphor identification, as demonstrated by the examples in the dataset.

(54) You hear a psychologist in the **UK** talking about intelligence. (TB9.OINPUT.L1)

(55) Most people want a busy shopping street in their town with butchers, bookshops, boutiques, **cafs**, and restaurants […] (TB10.WINPUT.R2)

(56) **Well**, now Lufthansa the German airline has made the calculation that on eighty per cent of its flights there is in fact a doctor amongst the passengers. (TB11.WINPUT.R3)

(57) <PREB2.ST0310> yeah but maybe in the afternoon if you stay at home **osea** sorry if you stay at school you won't have time for your own. </PREB2.ST0310> (PREB2.SOUTPUT.S4.ST0310)

Learner Errors and Non-Native-Like Phraseology. Identifying metaphorical uses in the language produced by L2 learners can be challenging due to the unconventional forms that are often present in such a heterogeneous data set. Littlemore et al. (2014) investigated metaphor use in L2 writing across proficiency levels and used two marking criteria to explore L1 influence in learner errors. They applied a "strict criterion" that considered non-native-like

phraseology incorrect and a “generous criterion” that counted it as correct (p. 124). However, they did not consider learner errors in applying the MIPVU procedure.

Similarly, Hoang and Boers (2018) developed a post-hoc “grammatically target-like” approach to determine the conventionality of L2 metaphorical uses in their study on EFL undergraduates’ *Writing* proficiency. They analysed errors contained in the identified metaphors to test whether they conformed to the conventions of English grammar, focusing on specific cases of non-target-like language use: incorrect spelling, incorrect PoS, inflectional errors, missing or wrong function words, and errors of valency and colligation (see Hoang & Boers, 2018, p. 4).

In contrast, Nacey (2020) established consistent criteria for treating learner errors and addressed them during the metaphor identification process. In her longitudinal study on the development of metaphorical production of L2 learners of English across grade levels, Nacey (2020) adopted low-level error criteria, that is, “errors that were of no consequence for either metaphor identification or word count were not adjusted” (p. 180). For example, the LU *princesse* was not altered because “its metaphorical status may be determined through consultation of MM’s entry for the English noun *princess*, the only obvious intended target” (Nacey, 2020, p. 181). Likewise, the occurrence of *her self* (written as two LUs) was corrected to *herself* (one LU) because the error affected the overall number of LUs and, therefore, metaphorical density.

In my study, L2 learners’ achievement at B2 level was essential for analysing the outcomes. To identify learner errors in metaphor identification, I implemented strict criteria that considered how the LU was produced rather than relying solely on the analyst’s interpretation of the intended written or spoken word. Although these high-level error criteria may have resulted in fewer potential metaphors, they ensured consistent treatment of learner errors, with a particular focus on lexis, such as spelling mistakes and unconventional constructions. Despite initially seeming more complex than previous approaches to error tagging in metaphor identification (cf. Nacey, 2020), this approach proved straightforward throughout the analysis, as will be demonstrated later in this chapter.

Spelling Mistakes: Errors made by L2 learners resulting from spelling mistakes in the written outputs or mistakes in pronunciation in the oral outputs were not corrected in standard orthographic form if they were not recognised as words in the dictionary. Consequently, LUs with misspelled words were excluded from metaphor analysis, even if the L2 learners knew the correct spelling. For instance, in example (58), the word *completaly* was misspelled and, therefore, not included in my analysis.

(58) It is **completaly** known that if you have enough money to move to a big city you will have more chances to find a good job. (PREB2.WOUTPUT.W1.ST0105)

Spelling mistakes encompassed not only errors in terms of standard spelling but also cases of omission (e.g., *posibilities* instead of *possibilities* in [59]), addition (e.g., *puppils* instead of *pupils* in [60]), possible L1 influence (e.g., *mayority* [*mayoría* in Spanish] instead of *majority* in [61]), and phonological replication (e.g., *revaise* instead of *revise /rɪˈvaɪz/* in [62]).

(59) First and foremost, it is a common belief that young people have more **posibilities** to study […] (PREB2.WOUTPUT.W1.ST0206)

(60) It is also vitally important to take the subjects into consideration because subjects that **puppils** are learning nowadays […] (POSTB2.WOUTPUT.W1.ST0401)

(61) For the great **mayority** of people, there is enough technology to change schools. (POSTB2.WOUTPUT.W1.ST0103)

(62) <PREB2.ST0110> […] because they they have to **revaise** the things that they did at school but as you said not a lot. </PREB2.ST0110> (PREB2.SOUTPUT.S4.ST0110)

By contrast, if a misspelled or mispronounced LU was already codified in the dictionary, e.g., *prize* instead of *price*, or *come /kʌm/* instead of *can /kæn/*, my approach accepted the produced word for metaphor identification. For instance, in example (63), *prize* was analysed even though the L2 learner may have known the correct spelling of *price* and intended to convey "the amount of money that you have to pay in order to buy something" (MM1) in this context. Similarly, although the participant may have intended to use the modal verb *can /kæn/* to express opportunity in example (64), my approach consulted the dictionary entry for the verb *come /kʌm/* during the analysis to identify metaphor.

(63) […] people can travel by many different means of transport and sometimes for a cheap **prize**. (PREB2.WOUTPUT.W1.ST0309)

(64) <POSTB2.ST0410> […] but it's it's okay because on the city they **come** meet more more people than in the countryside. </POSTB2.ST0410> (POSTB2.SOUTPUT.S4.ST0410)

It is important to acknowledge that the analytical decisions made in the analysis may have impacted the number of identified LUs in the study. For example, the conjunction *though* in (65)*,* instead of the past participle form *thought,* was automatically excluded from the analysis. Conversely, in example (66), the participant's omission of the infinitive marker in the modal verb "need to" led to the inclusion of the lexical verb *need* for metaphor analysis. Beyond the effect of this approach on the total number of LUs identified, using consistent criteria to address learner errors provided a reliable method for avoiding disagreement and uncertainty in the analytical decisions taken throughout the study.

(65) It is often **though** that people who were born many years ago were able to study […] (PREB2.WOUTPUT.W1.ST0208)

(66) <PREB2.ST0106> well when I **need** do some drawings for some homework yes but not really because I I I know how to do things on my own. </PREB2.ST0106> (PREB2.SOUTPUT.S4.ST0106)

Grammar Errors: The study adopted a lenient approach towards spelling mistakes caused by grammar mistakes, such as number form, verbal tense, or word class. This approach assumed that such spelling mistakes would not hinder the comprehension of the meaning in context or identifying LUs in the dictionary. For instance, in example (67), the study accepted the singular third-person form of the verb *pass* for metaphor analysis, even though the correct form "passes" was not used. However, this error in number form did not alter the meaning of "if time or a period of time passes, it happens and comes to an end" (MM6).

(67) […] each era has like its customs and as time **pass**, things are getting more different. (POSTB2.WOUTPUT.W1.ST0110)

Similarly, the study included the verb *give* in (68) for metaphor identification. Even though the correct past simple form *gave* was not used, this error in verbal tense did not impede the comprehension of the intended meaning of "to let someone have something as a present, or to provide something for someone" (LM1) in this context.

(68) <POSTB2.ST0410> […] when I was a child my parents when I got a good mark or something that my parents **give** me a little toy or just a lollipop […] </POSTB2.ST0410> (PREB2.SOUTPUT.S4.ST0403)

The study also included the adjective form *easy* in example (69), although the correct adverb form *easily* was not used. Despite this error in word class, the intended word could still be identifiable for metaphor identification purposes, as it conveyed the meaning of "without difficulty or effort" (MM1) in the given context.

(69) Therefore, travel gives the opportunity to learn more things **easy** and faster. (PREB2.WOUTPUT.W1.ST0104)

Unconventional Constructions: The study did not rectify incorrect constructions resulting from non-native-like phraseology used by L2 learners. Such constructions were only accepted for metaphor identification if the LUs were codified in the dictionary.

This study excluded coinages, i.e., neologisms, from the metaphor analysis as they were not listed in the reference source. Therefore, unconventional words such as *antecessors* (*antecesores* in Spanish) instead of *predecessors* or *ancestors* in (70) or *conscient* (*consciente* in Spanish) instead of *aware* in (71) were not included in my analysis.

(70) […] young people nowadays have more opportunities and a better education than their **antecessors**. (PREB2.WOUTPUT.W1.ST0105)

(71) <POSTB2.ST0110> […] and you see that nature is so beautiful maybe people can be **conscient** that we are like destroying the planet (.) </POSTB2.ST0110> (POSTB2.SOUTPUT.S4.ST0110)

This study included English cognates in the metaphor analysis, considering that such words with the same origin or related meaning were listed in the dictionary. For instance, in

examples (72) and (73), the terms *idioms* and *ambient* were used instead of the English equivalents, *languages* (*idiomas* in Spanish) and *environment* (*ambiente* in Spanish), respectively. The LUs *idioms* and *ambient* were considered for metaphor identification as they were recorded in the dictionary.

(72) In the old times, students were not obligated to learn different **idioms** so travelling becomes a difficult thing to do. (PREB2.WOUTPUT.W1.ST0109)

(73) <POSTB2.ST0109> they can stay fit in a natural and outside **ambient** and having fun not in the indoors […] </POSTB2.ST0109> (POSTB2.SOUTPUT.S4.ST0109)

As part of my analysis, I considered unidiomatic word combinations resulting from the misuse of words. For instance, in example (74), the word *qualify* illustrates incorrect word choice. It appears that the L2 learner intended to convey the idea of "assess" as "to carefully consider a situation, person, or problem in order to make a judgment" (MM1). Despite *qualify* being an incorrect word choice in this context, it was still considered for analysis because it was listed as a LU in the dictionary.

(74) For instance, if there are not exams they will need to develop different **forms** to **qualify**. (POSTB2.WOUTPUT.W1.ST0209)

The noun *forms* in (74) also illustrates non-native-like phraseology. The context suggests that *forms* was used as a synonym for *way*, meaning "a method for doing something" (MM1). In Spanish, the nouns *forma* (form) and *manera* (*way*) are interchangeable when referring to a method or manner of doing something. However, in English, this meaning is only expressed using the word "way". It seems that the L2 learner translated the sense of "method" or "manner" from Spanish into English using *form* instead of *way*. The similarity between the Spanish *forma* and the English *form* may have influenced the L2 speaker's choice, possibly due to possible L1 influence or transfer/calque. Non-native-like phraseology was included in the analysis if the LU was recorded in the dictionary, such as in the case of the noun *form*.

Phase 2. Establishing Contextual Meanings

The second step of the MIPVU procedure involves determining the contextual meaning for each LU identified in the text.[33] The contextual sense refers to the specific meaning a LU holds in the discourse. For example, the adjective *comfortable* is defined by six sense entries in MM, ranging in meaning from "feeling pleasant" (MM1) to "won/winning easily" (MM6).

(75) It cannot be denied that their homes are much more prepared and **comfortable** than they were in the past. (PREB2.WOUTPUT.W1.ST0402)

In the context of example (75), the adjective *comfortable* indicates "pleasant to use/wear" referring to furniture ("a comfortable piece of furniture feels pleasant to sit or lie on" [MM2]). However, none of the other senses recorded in the dictionary for this adjective were used, e.g., "having enough money" ("rich enough to pay for everything you need" [MM3]).

Establishing contextual senses can be a daunting task when identifying errors in L2 learner discourse (see Nacey, 2013). In my study, determining contextual meanings was somewhat straightforward since most LUs were used conventionally, and their senses were listed in dictionaries. The troubleshooting methods used to address learner errors and non-native-like phraseology provided consistent criteria to establish judgments in cases of spelling mistakes, grammar errors, and unconventional constructions arising from cognates and word misuse. By exclusively applying the MIPVU procedure to words recorded in the dictionary, I could establish contextual meanings based on the words produced. In the following, I explain some troubleshooting methods for establishing contextual meanings in my study.

Spelling Mistakes. When establishing the contextual meaning of misspelled or mispronounced words listed in the dictionary, I analysed their usage within the discourse and considered their basic meaning. For example, I examined the usage of the verb form *talking* instead of *taking* in (76).

[33] This was recorded in "Column M" of the *Excel* spreadsheet.

(76) **Talking** all the above into consideration, I strongly believe that young people in the past had less opportunities than young people nowadays […] (PREB2.WOUTPUT.W1.ST0310)

My approach did not consider the potentially assumed meaning of *take* ("to react to someone or something or consider them in a particular way" [LM12]) in the collocational phrase *taking all the above into consideration*. Instead, I focused solely on the basic meaning of *talk* ("to use words to communicate" [MM1]) to establish its contextual meaning.

It is worth noting that the systematic approach used to determine the contextual meaning of misspelled or mispronounced words listed in the dictionary had an impact on only 28 LUs (0.09%) in the entire dataset.

Grammar Errors. In cases where spelling mistakes resulted from grammar errors, I established contextual meanings based on the actual contextual sense of the words produced by L2 learners. For example, in (77), despite the L2 learner using the incorrect spelling of the irregular plural noun form *lives*, I was able to identify the intended noun *life* in the dictionary, facilitating the comprehension of its contextual meaning. Consequently, I established the contextual sense of the word *lifes* ("your particular way of living and the experiences that you have" [MM2]) based on the entry meanings recorded for *life* in the dictionary.

(77) […] teachers will be really different from the nowadays teachers just because of the different types of **lifes** and changes in education. (POSTB2.WOUTPUT.W1.ST0309)

This study followed a similar approach in dealing with inaccurate grammatical uses that impeded meaning comprehension. For instance, in example (78), the L2 learner's incorrect use of the noun form *attentions* hindered comprehension. The use of *attentions* refers to the third entry of the noun "attention" in MM, which describes "special care, help, or treatment for someone or something". However, the L2 learner used it as a countable noun, which only applies to the fifth entry related to "behaviour that shows someone that you love them or that you take a lot of interest in them".

(78) <PREB2.ST0104> […] normally my parents didn't help me because my sister is younger so probably the **attentions** more for my sister than for me that I need to do things on my own. (PREB2.SOUTPUT.S4.ST0104) </PREB2.ST0104>

This grammar error impeded establishing the contextual meaning of *attentions*. Thus, the study relied on the basic sense of *attention* ("the interest or thought that you give to something you are listening to or watching" [MM1]) to establish its contextual meaning.

Unconventional Constructions. Given the incorrect use of cognates in L2 discourse, this study employed the same troubleshooting method for establishing the contextual meaning of such unconventional constructions. For instance, in (79), the noun *diary* (*diario* in Spanish) was incorrectly used out of context in the form of an adjective. To determine the contextual meaning, the study relied on the basic sense of the LU.

(79) There will be subjects related with **diary** life. (POSTB2.WOUTPUT.W1.ST0409)

The Spanish term *diario* can function as both a noun (*diary* or *journal* in English)[34] and an adjective (*daily* in English). In example (79), the L2 learner used *diary* instead of the correct English adjective *daily*, possibly due to the similarity between the two words. To establish the contextual meaning of *diary* in this example, the study consulted its basic sense in the dictionary: "a book or digital document that has spaces for each day of the year, where you can make a note of things you plan to do" (MM2).[35] Therefore, the analyst's interpretation of *diary* as "done or happening every day" (MM1) was not considered in the analysis. In fact, *diary* in example (79) was regarded as metaphorically used in the analysis based on the comparison between daily life and being written down on someone's agenda.

To address unidiomatic word combinations caused by word misuse, I adopted the same approach used for cognates to determine their contextual meanings. For example, in (80), the preposition *during* was incorrectly used instead of *over*. Thus, I consulted the

[34] Note that the Spanish term *diario* can be translated into English as "diary" for American English or "journal" in British English.
[35] Following the analytical approach adopted in this study, the English word *diary* was interpreted in the sense of British English, meaning agenda or planner.

dictionary to determine the basic sense of *during* (“at some point in a period of time” [LM2]) and established it as the appropriate contextual meaning.

(80) Nevertheless, other people think that they are not going to improve **during** the years (POSTB2.WOUTPUT.W1.ST0407)

Non-native-like phraseology arising from unidiomatic word combinations, although meaningful, was subjected to a different procedure to determine their contextual meanings. While the MIPVU procedure was also used on the produced word, the actual contextual sense of the LU used was considered instead of its basic meaning. This analytical decision was based on the observation that comprehension of contextual meaning appeared to remain unaffected by the unconventional use of the produced word. For example, in (81), the noun *work* was used instead of the correct noun form *job*.

(81) […] today there are more opportunities to find a **work**. (PREB2.WOUTPUT.W1.ST0405)

The Spanish noun *trabajo* can mean both “activity that involves physical or mental effort” (MM1) and “work that you do regularly to earn money” (MM1). In English, these senses are distinguished by *work* and *job*, respectively. Some L2 learners may mistakenly use *work* instead of *job* in the phrase “find a job”. This study relied on the second dictionary entry of *work* in MM, which defines it as “a job that someone is paid to do” to determine its contextual meaning. Despite the unconventional word choice, the context made the meaning understandable.

Non-native-like phraseology can also result from unidiomatic word combinations that do not match their contextual meanings as recorded in dictionaries. An example of such an unconventional use is *put* in (82), where the meaning was not listed in the dictionary.

(82) </PREB2.ST0306> I agree with you because in the school teachers **put** a lot of homework and children have to have to do it but of course they have to be with their friends and family […] </PREB2.ST0306> (PREB2.SOUTPUT.S4.ST0306)

The verb *put* in this context does not match any of the nine senses listed in MM, ranging in meaning from “move something to position” (MM1) to “throw heavy metal ball” (MM9). In

English, the verbs *give* ("to tell someone to do a job or piece of work" [MM11]) or *set* ("to give students work to do as part of a course of study" [MM7a]) are more appropriate collocates when assigning students a piece of work that must be completed as part of their studies. The verb *put* (*poner* in Spanish) in "put a lot of homework" appears to be a possible L1 influence or transfer/calque from the Spanish phrase *poner muchos deberes*.

Although the contextual meaning of "put homework" is easily understood, it was impossible to confirm the comparison between the contextual and basic meanings in the dictionary. To address this issue, the study followed Nacey's (2013) approach to deciding about the metaphoricity of unconventional uses of LUs produced by L2 learners, which involves considering such novel uses as possibly metaphorical: "in the absence of any definitively determinable contextual sense for comparison, MIPVU allows one to retain such instantiations as possibly metaphor-related without a full commitment to claims of metaphoricity" (Nacey, 2013, p. 97).

Phase 3. Deciding about More Basic Meanings

The third step of the MIPVU procedure involves determining whether each LU in the text has a more basic sense in other contexts.[36] According to the MIPVU guidelines, a basic meaning is "a more concrete, specific, and human-oriented sense in contemporary language use" (Steen et al., 2010, p. 35). Thus, a meaning cannot be considered basic if it is not listed in a contemporary user's dictionary entry. However, a basic meaning is not necessarily the LU's most frequent, salient, and historically older sense. MIPVU does not systematically confirm the etymology of each LU in the procedure. Generally, basic meanings are physical, i.e., easier to visualise, hear, feel, smell, or taste, and more explicit, related to bodily action.

For example, consider the noun *step* in (83). After carefully examining all the meaning entries, the first entry of *step* in MM ("a short movement made by putting one foot in front of the other") was identified as the basic meaning. This sense is more concrete and directly

[36] This information was recorded in the "Column N" of the *Excel* spreadsheet.

related to bodily action, making it easier to perceive than the contextual meaning ("one of the stages in a process, or one of the levels on a scale" [MM6]).

(83) 22-year-old Scott is already a **step** ahead of his classmates, as he has been offered a year-long contract with the National Theatre of Scotland's under-26 theatre group. (TB8.WINPUT.R7)

In some cases, all senses of a LU can be regarded as "equally basic", indicating that there is no sufficient distinction between any of the dictionary entries for the contemporary user (Nacey et al., 2019b, p. 48). One such example is the noun *tests* in (84), where the contextual meaning of *test* as a "check of knowledge" ("a set of written or spoken questions used for finding out how much someone knows about a subject" [MM1]) is not sufficiently distinct from the other meaning entries: "check of body" (MM2); "check of machine" (MM3), "difficult situation" (MM4), and "test match in sports" (MM5), all of which are regarded as *equally basic*.

(84) It's hard to prove that our intelligence is actually increasing, even if young people nowadays tend to get higher marks in intelligence **tests** than they did 50 or 60 years ago. (TB9.OINPUT.L1)

In my study, recording decisions of basic meanings in a database proved to be a practical and time-saving method that ensured consistent decisions about basic meanings. For this, I created an *Excel* spreadsheet of basic senses alongside the metaphor identification spreadsheet in the same *Excel* file. As I encountered a new LU during the analysis, I added it to the database, which included information such as term, PoS, dictionary consulted, entry number, and definition (see Figure 36 in Appendix M). It should be noted that separate entries were created for homographs, such as nouns and verbs that shared the same form (e.g., change). However, although basic senses were quickly accessed at any point during the analysis, I carefully considered the basic sense recorded in the database when making decisions about metaphoricity for every LU under analysis. Next, I explain the troubleshooting methods used to decide the basic meaning of prepositions and multiword unit verbs.

Prepositions. Establishing basic senses can be challenging, especially with prepositions. Unlike "content" words (e.g., nouns, verbs, adjectives, and adverbs) which often have concrete and precise senses, "grammatical" words like prepositions can be more elusive.

Generally, prepositions denoting spatial relations are typically easier to determine. For example, the preposition *in* (85) indicates "used with the name of a container, place, or area to say where someone or something is" (LM1), while the preposition *up* (86) refers to "towards a higher place or position" (LM1), and the preposition *on* (86) suggests "touching a surface or being supported by a surface" (LM1a).

(85) When they build them **in** the country, everyone has to get there by car […] (TB10.OINPUT.L4)

(86) It was just so exhilarating to get **up on** stage and perform in front of an audience. (TB8.WINPUT.R7)

Some prepositions such as "of" and "for" present highly abstract meanings, which makes it difficult to establish a distinction between their basic and contextual meanings (Krennmayr, 2017). Therefore, these prepositions were consistently discarded for metaphor analysis in the present study.

No other grammatical words posed similar problems in determining more basic meanings since this study excluded closed-class words (e.g., determiners, conjunctions, pronouns, or auxiliary verbs) from metaphor identification, as explained earlier in this chapter.

Multiword Unit Verbs. Establishing the basic meanings of LUs in L2 discourse required analytical decisions based on adopted semantic demarcation troubleshooting methods. This study faced challenges when dealing with phrasal or prepositional verbs. For example, to establish the basic meaning of the prepositional verb *consisted of* in (87), the dictionary entries of its two constituent parts were consulted independently in the analysis. However, the proposition *of* was excluded from the analysis.

(87) This usually **consisted of** writing an essay on a topic related to my studies. (TB11.WINPUT.R6)

Additionally, the lexical verb *consist* lacked a clear definition in MM and LM. Hence, the multiword verb *consist of* was looked up in the dictionary to establish its basic meaning ("to be formed from two or more things or people" [LM1]) and facilitate metaphor identification. Although this appears to contradict pp. 49–50, it is worth noting that such cases where the dictionaries did not provide an independent definition for the lexical verb of multiword unit verbs were infrequent (only 0.12% of cases).

Phase 4. Deciding about Uses of Metaphor

In the final stage of the MIPVU procedure, a three-category mark-up system is used to determine the distinctness and similarity of LUs: Metaphor-Related Word (MRW; Section A); non-Metaphor-Related Word (non-MRW; Section B); and When in Doubt, Leave it in (WiDLii; Section C).[37] The succeeding subsections detail this classification, including the troubleshooting methods (Section D) that I employed to decide about uses of metaphor in analysing L2 discourse.

A. Metaphor-Related Words

A LU is considered a "Metaphor-Related Word" (MRW) when is potentially motivated by some form of similarity, indicating a non-literal comparison. The MIPVU procedure categorises metaphors into three types: "indirect", "direct", and "implicit" metaphor.[38]

Indirect MRWs. "Indirect" MRWs refer to LUs where the basic and contextual meanings are sufficiently distinct and are understood via indirect comparison, i.e., the contextual meaning contrasts with the basic meaning. An example is the noun *lockdown* in (88), which conveys metaphorical meaning in context.

(88) Basically, now that we are in a **lockdown,** we have realised that with technology everything is possible. (POSTB2.WOUTPUT.W1.ST0105)

[37] In my analysis, this information was recorded in "Column O" of the *Excel* spreadsheet.
[38] This classification was recorded in "Column P" of the *Excel* spreadsheet.

The contextual meaning of the term *lockdown* is defined as “a time when large numbers of people are ordered to stay at home either most or all of the time” (MM3), while its more basic meaning refers to “an occasion or time when prisoners are locked in their cells” (MM1). These senses are recorded as separate entries in the dictionary, indicating their distinctness. However, they are related through comparison, as being forced to stay at home and not being allowed to leave can be understood by comparison with being prisoners who are confined to their cells. Therefore, the contrast and indirect comparison of the contextual and basic meanings involve metaphoricity in *lockdown*. This example illustrates how we view “confinement” in terms of “prison” based on some form of similarity, reflecting the metaphorical mapping of FREEDOM IS OPEN AIR to think and/or communicate about imprisonment in English.

One of the main well-known differences between MIP and MIPVU is how they treat morphological boundaries when identifying metaphors. The Pragglejaz Group (2007) considers LUs with different dictionary entries, identical base forms, and distinct parts of speech as a single LU. For example, the noun and verb forms of *structure* are viewed as the same lexeme. However, MIPVU treats different parts of speech of the same lexeme as distinct LUs with different word class boundaries. Hence, both forms are considered homonyms, indicating that they are two LUs with the same spelling but distinct meanings.

The MIP approach considers the monosemous verb form of *structure* (“to plan or organize something” [MM1]) in (89) as “MRW” because its meaning contrasts and can be understood in comparison with the basic sense recorded for the noun form of *structure* (“something large such as a building or a bridge that is built from different parts” [MM3]), as illustrated in (90).

(89) Some people share the view that schools are badly **structured**. (POSTB2.WOUTPUT.W1.ST0404)

(90) Some new subjects will appear like the building of 3D **structures**. (POSTB2.WOUTPUT.W1.ST0406)

According to MIPVU, the verb form of *structure* in (89) is considered non-metaphorical because there is no basic meaning in the dictionary entry that it can contrast and be

understood in comparison with to be regarded as metaphorical in use. Thus, this study did not annotate the potential metaphorical use of *structure* (verb) following MIPVU.

Unlike the MIP procedure, MIPVU takes into account syntactic boundaries and suggests comparing LUs that belong to the same word category and identical grammatical classes. This means distinguishing between transitive and intransitive uses of the same verbal LU or countable and uncountable uses of identical nominal LU.

(91) […] it was when he **starred** on a TV show a few years ago and it happened like this. (TB8.OINPUT.L2)

For instance, the intransitive verb form of *star* (verb) in example (91) cannot be contrasted with the transitive verb form in the third entry of MM. While there may be a relationship between the basic ("to put a sign shaped like a star next to something on a list in order to mark it as special or important" [MM3]) and contextual meanings ("to be the main actor or performer in a film, play, television programme etc" [MM1]), MIPVU does not presume a metaphorical relationship between both verb forms of *star*.

Semino (2019) has pointed out that applying MIPVU word class and grammatical uses boundaries can be problematic, as lexicographers' "assumptions about the mental structure of the lexicon and about language" may affect analytical decisions (Semino, 2019, p. 317). Considering dictionaries as the result of individual decisions, inconsistencies may arise (see also MacArthur, 2015). However, this study strictly applies MIPVU's criteria for determining metaphoricity by taking into account PoS and grammar information. Using the dictionary consistently was essential for examining L2 heterogeneity in language use and ensuring the accuracy of my analysis.

Direct Metaphors. A LU is classified as a "direct" metaphor, such as a simile, when the contextual and basic senses of the LU do not explicitly contrast with each other, even though an underlying metaphorical comparison is evident at the conceptual level. The comparison is often expressed through *direct* language use, which may or may not be signalled by metaphor flags (MFlag) indicating that an underlying cross-domain mapping may be at play (cf. Goatly,

1997). Markers of "direct" metaphor, such as *like*, *such*, *as*, or *as if*, function to connect the two concepts being compared.

In example (92), the lexical marker *like* indicates a direct comparison between *study* and *game*, alerting of the non-literal meaning of *game* in this context ("an activity that you do for fun that has rules, and that you can win or lose" [MM1]). While there is no explicit contrast between the contextual and basic meanings, the use of the MFlag signals that the underlying metaphorical mapping FULFILLING RESPONSIBILITY IS PLAYING A GAME is at play.

(92) </PREB2.ST0208> […] most children have to know that **study** is an obligation and it's not **like** a **game** that if you do it right you are going to give a prize or you are going to go to any good place like a theme park for do it right […] </PREB2.ST0208> (PREB2.SOUTPUT.S4.ST0208)

(93) […] **technology** should be mentioned because it **is** the **base** of the future. (POSTB2.WOUTPUT.W1.ST0307)

Example (93) is an A is B metaphor, which draws a direct comparison between *technology* and *base* without any explicit signal of a metaphor marker. Instead, the contextual sense of the LU remains the same as its more basic sense ("the bottom part, edge, or surface of something" [MM1]). Although no metaphoricity is found at the linguistic level, the target domain *source of development* is structured through the source domain *foundation*, indicating a form of cross-domain mapping.

Implicit MRWs. "Implicit" MRWs do not show a contrast between contextual meaning and basic meanings, similar to "direct" MRWs. "Implicit" MRWs trigger metaphorical comparisons between two elements through substitution or ellipsis. This occurs when pro-forms such as pronouns (e.g., *it* or *one*) or determiners (e.g., *this* or *that*) refer to LUs that were previously used metaphorically in the text. An example of this is found in (94), where the pronoun *it* refers back to the phrase "taste of performing". The metaphoricity is not explicitly marked because the pronoun *it* itself is not metaphorically used. Instead, the pronoun refers to the previously used LU *taste* in the text, which was identified as an "MRW".

(94) Kim got her first **taste** of performing as a ballet dancer when she was just two and has been hooked on performing ever since. Although her ballet days are now behind her, she admits that **it** gave her a great introduction. (TB8.WINPUT.R7)

It is important to note that "implicit" metaphors were not identified in the analysis. These metaphors are closed-class words and were thus excluded from metaphor identification as they were deemed of no particular interest for the present research.

B. Non-Metaphor-Related Words

A LU is considered a "non-Metaphor-Related Word" (non-MRW) when the basic sense of the word is not sufficiently distinct from the contextual sense, which can occur when one of its senses is more specific or more general. In example (95), the verb *see* can be understood in context as "to watch something such as a film or television programme" (MM1b). However, this contextual sense is not sufficiently distinct from its basic meaning, defined as "to notice someone or something using your eyes" (MM1).

(95) Anyway, he did the show which was recorded and all of us were **longing** to **see** him in it. (TB8.OINPUT.L2)

Sometimes, the contextual sense of an LU aligns with its literal meaning, indicating that the contextual sense is the same as the basic one. In (95), the adjective *longing* can be tagged "non-MRW", as it means "a strong feeling of wanting someone or something" (MM1).

It is important to note that identifying a LU as "non-MRW" does not necessarily mean that it cannot express some other type of figurative meaning. For example, a LU can express metonymy, where the contextual meaning of the LU is not related by similarity but by contiguity, e.g., stand-for or part-for-whole relationship, as *money* in (96).

(96) This **suggests** that there is a minimum amount of **money** we need to earn to make us happy […] (TB9.WINPUT.R5)

(97) If you want you can travel by plane or by train but in the past there were not the **amazing** possibilities that we have. (PREB2.WOUTPUT.W1.ST0202)

A LU may also express hyperbole, where exaggeration is used for emphasis, as *amazing* in (97). Additionally, possible personification (PP)[39] can be employed to attribute human qualities to non-human entities, as demonstrated by *suggests* in example (96).

C. When in Doubt, Leave it in

The MIPVU procedure identifies a LU as "When in Doubt, Leave it in" (WiDLii) when its contextual meaning cannot be clearly classified as either "MRW" or "non-MRW". For instance, in example (98), the L2 learner used the verb *afford*, which has both a basic sense ("if you can afford something, you have enough money to be able to pay for it" [MM1]) and a metaphorical sense ("if you can afford something you can do it without having to worry about it causing problems for you" [MM2]).

(98) However, it not should be forgotten that young people in the past had a lot of wars or economical problems so they could not **afford** to take up new activities because they were solving another problems. (PREB2.WOUTPUT.W1.ST0310)

The use of *afford* in this example can be seen as a "bridge metaphor" (Nacey, 2013, p. 97) or a "conflated case" (MacArthur et al., 2015, p. 193) since it represents a borderline case where the LU can be viewed as both "MRW" and "non-MRW" in context.

In other cases, ambiguity for analysis arose due to their intertwined nature of metaphor and metonymy. In example (99), the noun *studies* presented a challenging case for analysis. Although the contextual meaning of *studies* ("the work that you do while you are at a college or university" [MM2]) does not correspond to any dictionary entry, it is related to the sense of "someone's experience of learning or being taught" (MM1b), which can be found in the term *education*. The unconventional use of *studies* in this context could be seen as an extension of the basic sense of *study* ("a subject that people study at a college or university" [MM3]). Thus,

[39] MIPVU includes an additional tag for PP in metaphor identification. However, the present research did not investigate the use of personification. In my analysis, PP cases were marked as "non-MRW" since there is no contrast between the contextual and basic meaning.

the LU *studies* can be considered “MRW”, indicating that knowledge is viewed in terms of the number of courses taken.

(99) Thirdly, to have a job is needed a high level of **studies**. (PREB2.WOUTPUT.W1.ST0305)

However, the contextual meaning of *studies* may also be metonymic, as it refers to the work done as part of the learning process rather than the concept of education itself. This made it difficult to confidently identify *studies* as “MRW” because the contextual meaning is not clearly metaphorical. The relationship between the amount of knowledge and the subjects studied in the learning process may also be considered contiguous, creating a form of indirect meaning. Consequently, the LU *studies* was labelled as “WiDLii” in the analysis.

D. Troubleshooting Methods for Deciding about Uses of Metaphor

Overall, using MIPVU to identify metaphorical meanings in my study was straightforward. However, there were certain instances where it posed challenges, especially in cases of learner errors and non-native-like phraseology. To address these challenges, I relied on the analytical process previously described to establish contextual meanings based on the word produced.

For general grammar errors resulting from spelling mistakes, I followed the standard MIPVU criteria to determine whether the contextual meaning was sufficiently different from the basic sense and related by some form of similarity. However, misspelled or mispronounced words that were codified in the dictionary, grammar errors that impeded meaning comprehension, and unconventional constructions such as cognates were directly classified as “non-MRW” in my analysis. The decision to equate the basic sense with the contextual sense in identifying such LUs meant that the two meanings were identical and lacked metaphorical use. As a result, they were tagged as “non-MRW”. However, deciding whether to treat non-native-like phraseology as potentially MRW or simply errors proved more complex when dealing with unidiomatic word combinations. Several factors required careful consideration, which I will explain in the following subsections.

Learner Errors. As an example of misspelled or mispronounced words that were codified in the dictionary, consider the case of using *contract* instead of *contrast* in (100). Despite being listed in the dictionary, I established the basic sense of *contract* as the contextual meaning, which refers to "a written legal agreement between two people or businesses that says what each must do for the other or give to the other" (MM1).

(100) In **contract**, more young people have opportunities so […] (PREB2.WOUTPUT.W1.ST0104)

Consequently, *contract* was identified as "non-MRW", and the potential metaphorical use of *contrast*, hence, was not analysed. Comparing the LU senses was not feasible because the contextual meaning was "a difference between people, ideas, situations, things etc that are being compared" (MM1), whereas its basic meaning was "the degree of difference between the light and dark parts of a television picture, X-ray, photocopy etc" (MM3). Although this approach may have limited the number of potential MRWs in my study, it represented a more precise analysis of L2 learners' proficiency at B2 level.

Non-Native-Like Phraseology. The same approach was taken to identify metaphorical meanings with unconventional constructions. For example, consider the case of using the cognate *facilities* instead of *opportunity* in (101) or *sounds* instead of *noise* in (102). In my analysis, both LUs were treated as errors and, hence, automatically tagged as "non-MRW".

(101) Nowadays we have got more **facilities** to travel than in the past and […] (PREB2.WOUTPUT.W1.ST0101)

(102) <POSTB2.ST0308> […] you can be more relaxed and I don't know in the city you have a lot of a lot of **sounds**. </POSTB2.ST0308> (POSTB2.SOUTPUT.S4.ST0308)

When a LU was used unconventionally in discourse, but its intended meaning was successfully conveyed, it was not considered an error and thus not directly tagged as "non-MRW". In such cases, the contextual meaning of the actual LU produced was used for metaphor identification since its intended meaning in the discourse was found in one of the dictionary entries, as mentioned earlier. In example (103), the verb *lead* in the phrase "lead a

business" can be considered grammatically incorrect as the correct form should be the -ing form *leading*, which works as a noun in the text. However, in this study, the verb *lead* was accepted for metaphor identification as the grammar error did not hinder the identification of the LU in the dictionary.

(103) […] new technologies helps young people to have opportunities such as work as an influencer or **lead** a business easily. (PREB2.WOUTPUT.W1.ST0102)

Furthermore, the use of *lead* instead of *run* in this context could also be seen as an incorrect word choice. However, this unconventional construction did not seem to have impeded the comprehension of "control" or "organisation", which are recorded in the second entry of *run* in MM: "to control and organize something such as a business, organization, or event". Following the MIPVU procedure, the contextual meaning of *lead* ("to be in charge of an organization, country, or team, or a group of people who are trying to do something" [LM3]) was contrasted with its basic sense ("to take someone somewhere by going in front of them while they follow, or by pulling them gently" [LM1]) for metaphor analysis. Although the incorrect word choice of *lead* could be seen as a simple error, the criteria established to address meaningful uses of unconventional English for metaphor identification were applied. Hence, *lead* was considered potentially motivated by a non-literal comparison in context and was thus tagged as "MRW".

To identify potentially metaphorical unconventional phraseology in non-native English discourse, this study followed MacArthur's (2019) approach and classified them as "UNC".[40] However, the study did not conduct an in-depth analysis to explore the reasons behind these unconventional uses, such as whether they resulted from possible L1 influence or transfer/calque, creative use, rote learning, potential mistranslation of meaning, or some other source. This study treated unconventional uses as "metaphor alternatives". Future research could explore the learner discourse from this perspective.

[40] This information was recorded in "Column Q" of the *Excel* spreadsheet.

3.3.1.3. Post-Analysis Stage: Cleaning-up Process

To ensure consistency and accuracy, a final cleaning-up phase was conducted to apply default criteria to the 31,077 LUs. Prior to analysing the outcomes, the dataset was carefully reviewed for systematic errors and corrected as necessary. The post-analysis cleaning-up actions addressed two main issues: incomplete data entry and technical oversights. Next, I describe the process used to check for errors and clean up the dataset.

To ensure the completeness of data entries for metaphor identification, a double-check was conducted on the *Excel* spreadsheet to confirm that all data entries were complete in the relevant columns. A few rare instances of empty cells were identified, particularly related to the decision to accept or reject LUs for metaphor identification. In addition, some missed cases were identified when recording the dictionary used to determine the presence of metaphoricity or marking further decisions about metaphorically-used metaphors, such as “indirect” or “direct”. Any empty cells were filled in following the MIPVU protocol. Special attention was given to identifying technical oversights that may have affected the overall findings. All cells related to metaphoricity were thoroughly checked for possible errors.

All identified errors were corrected during the post-analysis phase. However, it is important to note that the cleaning-up process is not entirely foolproof. The decisions made across the four stages of the MIPVU procedure were carried out manually and thus fallible. Error rate reports were not included. Instead, this study opted to test the reliability of the methodology through alternative research approaches, as presented in the next section.

3.3.2. Testing Reliability

To ensure reliable metaphor identification in this study, three linguists conducted two different reliability tests. The raters included both non-native speakers of English whose L1 is Spanish (Analyst 1 and Analyst 3) and native speakers of English (Analyst 2). Analyst 1 is the author of the present study, and Analyst 2 and Analyst 3 are experts in metaphor in discourse and have undergone extensive training in metaphor identification. These independent raters

are part of a research group that has closely collaborated on two nationally-funded research projects on metaphor identification, with Analyst 1 having participated in one of them.

Analyst 1 received a four-month training period in MIPVU consisting of three one-day training sessions facilitated by the two external researchers in the pre-analysis stage. The training aimed to ensure consistent implementation of the metaphor identification process and involved two rounds of metaphor identification analysis using selected discourse, consisting of oral and written input texts. The sample data comprised three texts selected from the textbook Unit 8, "Dream of the stars" (Brook-Hart, 2014, pp. 84–93), with a total of 1,749 words (*M* = 470.17, *SD* = 361.61).

In the following subsections, I will describe the first reliability test conducted during the analysis stage in order to report on the inter-rater agreement for metaphor identification. Finally, I will explain the second reliability test, which was performed by Analyst 1 during the post-analysis stage to evaluate the stability of metaphor identification across time.

3.3.2.1. Reliability Testing 1: Inter-Rater Reliability Testing

The analytical criteria slightly adjusted from MIPVU underwent an inter-rater reliability test to ensure consistency. The first reliability test involved three independent analysts (Analyst 1, Analyst 2, and Analyst 3). In what follows, I will describe the methodology used for conducting the test and, subsequently, report and discuss the results of reliability testing 1.

3.3.2.1.1. Method

To assess the reliability of the slightly adapted MIPVU procedure, six texts were randomly selected from the pre-test and post-test text files and rated by the three independent analysts in reliability testing 1. The sample texts consisted of learner discourse, comprising 1,077 words (*M* = 178.45, *SD* = 21.11). Notably, the first reliability test did not include spoken data since the transcription process had not been completed at that research stage.

Steen et al. (2010) recommended two statistical measures for assessing inter-rater reliability: Cochran's *Q* (e.g., Dunn, 1989) and the Kappa (κ, e.g., Howell, 2010). Cochran's *Q*

measures the overall difference between individual researchers, while the κ measures agreement across individual cases or ratings by accounting for chance agreement between analysts (Steen et al., pp. 150–151). In this study, I chose to use the κ alone as it provides a more comprehensive measurement of inter-rater reliability. Specifically, Fleiss' κ (McHugh, 2012) was used for this investigation since it is suitable for calculating reliability between three or more analysts. The 95% Confidence Interval (CI) for the κ value was also calculated to provide insight into the degree of agreement between different analysts when applying the MIPVU procedure in a general context.

Analyst 2 and Analyst 3 submitted their annotated *Excel* spreadsheet to Analyst 1, who was responsible for conducting the statistical analysis. The data from the individual spreadsheets were merged to create a single database, and the frequencies were calculated based on LUs. The κ measures were then computed using the *R* software environment (2022). However, the reliability test encountered some challenges as the MIPVU procedure does not conform to a two-fold classification for determining LUs. Consequently, I chose to exclude the demarcation of LUs in inter-rater reliability testing and instead focused solely on decisions regarding metaphoricity, as suggested by Steen et al. (2010).

3.3.2.1.2. Results

Table 10 below presents the results of the three-way inter-rater reliability test on the six sample texts. The analysis revealed complete agreement among the three analysts on 96.70% of the cases (997 LUs), whether they were metaphorical or not. Only 3.3% of cases (34 LUs) showed discrepancies in interpretation among the analysts, primarily due to coding differences (25 LUs) and procedural misunderstandings (9 LUs). These results are consistent with Steen et al.'s (2010) study on metaphor in discourse (see pp.153–161).

Table 10

Results of the Inter-Rater Reliability Test for Six Sample Texts across Three Independent Analysts

File ID	LUs	Percentage unanimous					Fleiss' κ	95% CI
		Non-MRW	MRW	DFMA	WiDLii	Total		
IRRT.01	173	43.93 (*n* = 76)	10.98 (*n* = 19)	42.20 (*n* = 73)	0.58 (*n* = 1)	97.69 (*n* = 169)	0.96	[0.92, 1.00]
IRRT.02	164	44.51 (*n* = 73)	5.49 (*n* = 9)	47.56 (*n* = 78)	0.00	97.56 (*n* = 160)	0.89	[0.81, 0.98]
IRRT.03	189	24.87 (*n* = 47)	17.99 (*n* = 34)	52.91 (*n* = 100)	0.00	95.77 (*n* = 181)	0.95	[0.91, 0.99]
IRRT.04	146	35.62 (*n* = 52)	7.53 (*n* = 11)	51.37 (*n* = 75)	0.00	94.52 (*n* = 138)	0.87	[0.79, 0.96]
IRRT.05	166	37.95 (*n* = 63)	12.05 (*n* = 20)	48.80 (*n* = 81)	0.00	98.80 (*n* = 164)	0.98	[0.95, 1.01]
IRRT.06	193	31.09 (*n* = 60)	16.58 (*n* = 32)	48.19 (*n* = 93)	0.00	95.85 (*n* = 185)	0.91	[0.86, 0.96]
Total	1,031	35.98 (*n* = 371)	12.12 (*n* = 125)	48.50 (*n* = 500)	0.10 (*n* = 1)	96.70 (*n* = 997)	0.94	[0.92, 0.96]

The statistical analysis conducted on the data showed a high level of inter-case agreement (Fleiss' κ = 0.94, 95% CI = [0.92, 0.96]), indicating that the ratings were not the result of chance alone. These findings provide robust evidence that modifying certain aspects of MIPVU did not lead to biased outcomes in the present study.

3.3.2.2. Reliability Testing 2: Stability Testing

Due to time constraints, conducting a second round of inter-rater statistical analysis or reanalysing the entire dataset, as suggested by the Pragglejaz Group (2007), was not feasible. Instead, I opted for an alternative approach proposed by Nacey (2013) to test the consistency of the findings across time. This involved "repeating the identification process at a later date on a selection of texts, to give an indication of stability" (Nacey, 2013, pp. 115–116).

A second reliability test was conducted to assess the stability of the author's identification of metaphors throughout the analysis. In turn, this test aimed to confirm the consistency of the inter-rater reliability test results and rule out any chance of accidental findings. Next, I will describe the methodology employed for carrying out the test and then report on the outcomes of reliability testing 2.

3.3.2.2.1. Method

The second coding round was conducted two months after completing the first pass of the entire dataset. Six texts were randomly selected, which represented various types of discourse and consisted of three spoken and three written sample texts comprising 2,128 words (*M* = 301.52, *SD* = 228.94). These six texts accounted for about 5.5% of the entire dataset, including two input texts from the textbook (ST.01 and ST.02) that were identified at the beginning of the analysis, and four output texts produced in the pre-test (ST.03 and ST.04) and post-test (ST.05 and ST.06), which were identified halfway and toward the end of the first pass, respectively, "to compensate for potential inconsistencies" (Nacey, 2013, p. 116).

To assess the reliability of the analysis, Cohen's κ (Landis & Koch, 1977) was employed. This measure is commonly used to evaluate the degree of agreement between two raters, or in this case, between the same rater at two different time points. In this study, Cohen's κ and the 95% CI were computed to evaluate the degree of agreement between my analyses of the six selected texts in the two coding rounds.

3.3.2.2.2. Results

Table 11 below presents the results of metaphor identification from the first and second coding rounds. The analysis revealed that out of the 2,160 LUs marked in the second reliability testing, only 5 MRW scores differed between the first pass (329) and the second pass (334). Only 14 LUs were identified differently, accounting for about 1% of the selected sample texts. Six of these 14 LUs were initially considered "non-MRW", with four being reclassified as "MRW" and two as "DFMA". Three LUs that were initially marked as "MRW" were later changed to "non-MRW", while two LUs originally coded as "WiDLii" were reclassified as "MRW". The remaining three LUs were first marked as "DFMA", with two being changed to "MRW" and one to "non-MRW".

The discrepancies between the two passes were mainly attributed to mechanical errors in registration (4 LUs), coding discrepancies (9 LUs), or oversights in learner errors (1 LU). However, these discrepancies were not substantial enough to undermine the internal

consistency of the analysis. The metaphor analysis revision allowed for post hoc corrections to ensure consistency across the entire dataset. The entire dataset was searched again for instances where the above discrepancies occurred.

The reliability test performed on the dependent samples of the two coding rounds for both MRW and non-MRW cases showed a high degree of agreement between the first pass and second pass (Cohen's κ = 0.98, 95% CI = [0.97, 0.99]). These findings suggest that the quality of the slightly adapted MIPVU procedure did not decrease, and consistent results were obtained from the start to the end of the annotation process.

Table 11

Results of the Stability Reliability Test for Metaphor Identification: First Pass and Second Pass

File ID	LUs	First pass						Second pass						Cohen's κ	95% CI
		MRW		Non-MRW	WiDLii	DFMA	MFlag	MRW		Non-MRW	WiDLii	DFMA	MFlag		
		Indirect	Direct					Indirect	Direct						
ST.01	784	133	0	306	1	344	0	135	0	303	1	345	0	0.98	[0.97, 1.00]
ST.02	128	11	0	54	0	63	0	10	0	55	0	63	0	0.94	[0.83, 1.05]
ST.03	234	42	0	92	0	100	0	41	0	93	0	100	0	0.95	[0.89, 1.01]
ST.04	411	31	5	100	0	271	4	34	5	100	0	268	4	0.98	[0.94, 1.02]
ST.05	257	54	1	91	0	111	0	54	1	90	0	112	0	1.00	[1.00, 1.00]
ST.06	346	51	1	85	2	206	1	53	1	85	0	206	1	1.00	[1.00, 1.00]
Total	2,160	322	7	728	3	1,095	5	327	7	726	1	1,094	5	0.98	[0.97, 0.99]

3.3.3. Data Treatment

To address the three main research questions presented in this PhD dissertation (as explained in Section 2.5 of Chapter 2), both qualitative and quantitative approaches were adopted. Further details about these methods are presented below.

3.3.3.1. Quantitative Analysis

To answer RQ1, *Microsoft Excel* (2021) was used to automatically retrieve all metaphors in the examined texts, classify decisions on metaphoricity, and categorise types of metaphorical language into open-class (i.e., nouns, verbs, adjectives, and adverbs) and closed-class (i.e., prepositions) metaphors. To determine both the amount and the proportion of metaphors, the analysis focused on LU tokens instead of LU types. Metaphor density was

calculated by dividing the number of MRWs by the total number of LUs, and the resulting score was multiplied by 100.

MRW types were also extracted using Microsoft *Excel* (2021), providing a detailed account of the open- and closed-class metaphors found in the learner discourse. To calculate the relative frequencies (RF) of MRW types within each word class, each absolute frequency (AF) was divided by the total number of MRW types in each category and multiplied by 100.

3.3.3.1.1. Statistical Procedures

Statistical analyses were conducted using the *R* software environment (2022) for statistical computing and graphics. Non-parametric tests were chosen to ensure the robustness of the results, given that the sample sizes were not greater than 30.

The Shapiro-Wilk test for normality was used to assess the distribution of the variables. A test statistic greater than .05 indicates normal distribution, while a test statistic less than .05 implies non-normal distribution. As outliers can have a significant impact on the results, especially with a small sample size of only 40 participants, I employed a box plot to identify any extreme values in the data distribution.

To conduct inferential analyses of RQ1 and RQ2, I employed the independent Wilcoxon Rank Sum test for unpaired samples to compare the results of the control and experimental groups in the pre-test and post-test. A significant difference was considered to exist between the groups when the *p* value was less than .05, which indicates strong evidence against the null hypothesis. Additionally, I used the dependent Wilcoxon Signed Rank test for paired samples to examine the differences in the evolution of metaphor use and B2 level performance within each group (pre-test vs. post-test). The absolute values were converted into relative increases (RI) using the following formula:

$$\frac{\text{Post-test score - Pre-test score}}{\text{Pre-test score}}$$

It is important to acknowledge that statistical analyses carried out on a sample of 20 participants may not be entirely representative of the general population. However, these

analyses can still provide valuable insights into potential trends. To complement the comparative statistical analyses of RQ1 and RQ2, Cohen's *d* coefficient and the 95% CI were calculated to determine statistically significant differences between groups. Cohen's *d* measures the effect size of the difference between the two groups, indicating the magnitude of the difference beyond statistical significance. A Cohen's *d* of 0.20 is considered a small effect size, while a Cohen's *d* of 0.50 is considered a medium effect size, and a Cohen's *d* of 0.80 or higher is considered a large effect size (Cohen, 1988). The 95% CI provides a range of values within which the true population parameter is likely to lie with a probability of 95%.

To answer RQ3, I employed Spearman's ρ correlation coefficient to examine the relationship between metaphor use and achievement at B2 level. This statistical analysis method allowed for determining whether there was a significant correlation between the two variables and measuring the strength of the relationship. A correlation coefficient of 0 indicates no correlation, while a coefficient of +1 or -1 indicates a perfect positive or negative correlation, respectively. The strength of the correlation can be classified based on the magnitude of the correlation coefficient. If *r* is equal to or greater than .700 or smaller than -.700, the correlation is strong. If *r* is between .300 and .700 or between -.300 and -.700, it is moderate. If *r* is between 0 and .300 or 0 and -.300, the correlation is weak.

3.3.3.2. Qualitative Analysis

To complement the quantitative results of RQ1, a qualitative analysis was conducted to better understand participants' depth of vocabulary knowledge in their oral and written discourse. Drawing on the importance of topic similarity in preparing for high-stakes ESOL exams, the analysis aimed to explore the topic-related uses of metaphors targeted at B2 level by both groups of participants before and after the longitudinal study. To achieve this, examples from their oral and written outputs were manually examined and compared the B2-level uses of open-class metaphors related to the given topics between both groups. The *Cambridge Dictionary* online (CAMD) was used to identify proficiency levels in the vocabulary under examination, as this dictionary assigns CEFR levels to words.

3.3.4. Concluding Remarks

This section (3.3) has described the method used to analyse the results of metaphor use in the learner discourse of the present study. These are the key points:

- The study used a slightly adapted version of MIPVU (Steen et al., 2010) to identify metaphor use in the oral and written outputs of L2 learners of English (31,077 LUs).
- To ensure replicability, the conservative deviations from the standard MIPVU method have been described in detail, including methodological issues and analytical decisions that impacted the four analytical phases of MIPVU.
    - Phase 1: Adjustments to the MIPVU guidelines involved special demarcation and exclusion of concrete LUs in accounting for the topic-based approach adopted and L2 learners' constraints in language use.
    - Phase 2: Consistent criteria were applied to address spelling mistakes, grammar errors resulting from spelling mistakes, and non-native-like phraseology when establishing contextual meanings of LUs.
    - Phase 3: Troubleshooting methods were employed to decide the basic meaning of prepositions and multiword unit verbs.
    - Phase 4: Troubleshooting methods were used to determine the metaphor status of misspelled or mispronounced LUs and unidiomatic word combinations.
- Reliability testing of the adapted MIPVU procedure used in the study demonstrated consistent application of analytical criteria by different analysts and the principal researcher across time.
- The data treatment encompassed both quantitative and qualitative analyses. Non-parametric tests were used for statistical analyses to ensure the validity of the results, as the sample sizes were less than 30.

CHAPTER 4

RESULTS

Chapter 4 summarises the results obtained from analysing the impact of metaphor-mediated instruction on L2 learners' metaphor use and compares these findings with their CEFR B2 level of English proficiency, as measured by the *B2 First for Schools Cambridge English* qualification. This chapter is divided into three main sections, each devoted to answering a specific research question.

Section 4.1 examines the extent to which exposure to CL-oriented approaches affects metaphor use in the oral and written production of L2 learners of English aiming at reaching a B2 level (RQ1). In section 4.2, the overall B2 achievement of participants and their individual performance in the productive skills of *Speaking* and *Writing* are compared before and after the teaching treatment (RQ2). Section 4.3 explores whether a relationship can be established between metaphor use in the L2 and learners' performance on the *B2 First for Schools* examination, which is a widely recognised measure of their B2 proficiency level (RQ3).

4.1. RQ1: METAPHOR PERFORMANCE AT B2 LEVEL

To what extent does incorporating CL-oriented approaches to distributed learning of metaphor in the L2 classroom affect metaphor use in the oral and written production of L2 learners of English studying at B2 level?

This section examines the differences in metaphor use between L2 learners who received metaphor-mediated instruction (experimental group) and those not exposed to the application of CL-oriented teaching methods (control group) for B2 training. The answer to RQ1 will be addressed by comparing the oral and written productions of participants from both groups at two levels of analysis (RQ1a and RQ1b respectively) in the pre-test and delayed

post-test. Specifically, this study explores the *amount* of metaphorical language (measured in metaphor density) and the *type* of metaphors used (open-class vs. closed-class metaphors).

4.1.1. RQ1a: Use of Metaphorical Language in Oral and Written Discourse

What is the effect of metaphor-mediated instruction on the amount (measured in metaphor density) of metaphorical language use in oral and written learner discourse at B2 level?

This section presents the results of the analysis of metaphor density in the oral and written discourse produced by the participants from the control and experimental groups in the pre-test and post-testing measures of the *Speaking* and *Writing* tests aimed at B2 level.

To answer RQ1a, this section is divided into three parts. The first and second subsections (pre-test and post-test, respectively) describe the raw scores and density distribution of participants' metaphor use in oral and written discourse, followed by the inferential analyses of the differences between the metaphor density displayed in L2 learners' oral and written outputs in each testing measure. In the third subsection, the focus shifts to the evolution of metaphor density within each group over time, with a particular emphasis on comparing the pre-testing and the post-testing phases from a longitudinal perspective. Additionally, a comparative analysis of the use of metaphor between the two groups is performed to identify similarities and differences in their respective progress.

4.1.1.1. Metaphor Density: Pre-Test

The results of the analysis of metaphor density in the pre-test for the oral and written discourse produced by the control and experimental groups are shown in Table 12. A more detailed description can be found in Tables 79–82 in Appendix N.

Table 12

Metaphor Use in Learner Oral and Written Discourse (Control vs. Experimental Groups): Pre-Test

Discourse	Groups	LUs	Non-MRWs	MRWs	Metaphor density (%)			
					M	*Min.*	*Max.*	*SD*
Oral	Control	3,965	1,157	453	11.58	7.60	20.00	3.12
	Experimental	4,724	1,456	501	10.88	7.11	18.64	2.50
Written	Control	3,463	1,375	547	15.69	10.86	20.20	2.70
	Experimental	3,627	1,449	604	16.63	12.97	20.79	2.29

In the pre-test, the control and experimental groups exhibited similar rates of metaphor density in both their oral and written discourse. It was observed that both groups produced a significantly higher frequency of metaphors in their written essays compared to their spoken productions (W[156.50]) = 5.152, $p < .001$). To delve deeper into these findings, additional detailed information is provided below to shed lights on these patterns.

The *Speaking* pre-test analyses indicated that the control group had a higher mean score for metaphor use (M = 11.58%, SD = 3.12) in spoken productions compared to the experimental group (M = 10.88%, SD = 2.50) before the teaching treatment.

The *Writing* pre-test analyses revealed that the experimental group exhibited a higher rate of metaphor density (M = 16.63%, SD = 2.29) in their written essays than the control group (M = 15.69%, SD = 2.70) before the instructional intervention.

The Wilcoxon tests indicated no statistically significant differences ($p > .05$) in metaphor density between the control and experimental groups in the pre-test's oral or written discourse. For a more detailed description, see Table 13 below.

Table 13

Metaphor Density in Participant Groups' Oral and Written Discourse: Pre-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	11.58	10.88	0.502	225	.512	0.25	[-0.39, 0.89]
Written	15.69	16.63	0.777	156.50	.245	0.38	[-1.02, 0.27]

These initial findings provide a solid foundation for comparing and evaluating metaphor density produced by the control and experimental groups. By establishing a comparable

starting point in terms of metaphorical usage in the pre-test, the impact of the instructional intervention can be effectively assessed in the post-test and draw meaningful conclusions about its effectiveness in enhancing metaphorical language in the L2 classroom.

4.1.1.2. Metaphor Density: Post-Test

Table 14 displays the findings of the analysis of metaphor density in the post-test for the oral and written discourse produced by the control and experimental groups. For a more comprehensive description of these results, see Tables 83–86 in Appendix N.

Table 14

Metaphor Use in Learner Oral and Written Discourse (Control vs. Experimental Groups): Post-Test

Discourse	Groups	LUs	Non-MRWs	MRWs	Metaphor density (%)			
					M	*Min.*	*Max.*	*SD*
Oral	Control	3,197	1,002	310	9.63	5.23	14.02	2.95
	Experimental	4,461	1,459	502	11.07	6.47	19.24	3.08
Written	Control	3,759	1,308	619	16.40	12.18	21.30	2.53
	Experimental	3,881	1,229	748	19.31	13.24	26.67	3.16

In the post-test, differences in metaphor use between the control and experimental groups became apparent in both oral and written discourse. A significantly greater prevalence of metaphors was observed in the written essays of both groups than in their spoken productions ($W[78]) = 7.607$, $p < .001$). Findings are further explored below to provide detailed information that clarify these trends.

Table 15

Metaphor Density in Participant Groups' Oral and Written Discourse: Post-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	9.63	11.07	1.453	156	.242	0.48	[-1.13, 0.17]
Written	16.40	19.31	2.821	92.50	.004	1.02	[-1.70, -0.34]

The *Speaking* post-test analyses revealed that the experimental group showed a higher mean score ($p > .05$) of metaphor density in their spoken productions ($M = 11.07\%$, *SD*

= 3.08) than the control group (*M* = 9.63%, *SD* = 2.95) at the end of the study. For more detailed information regarding the statistical analyses, see Table 15 above.

The *Writing* post-test analyses indicated that the experimental group showed a statistically significantly higher average rate of metaphor density (*W*[92.50]) = 2.821, *p* = .004) in their written discourse (*M* = 19.31%, *SD* = 3.16) after receiving metaphor-mediated instruction, compared to the control group who used fewer metaphors in their essays (*M* = 16.40%, *SD* = 2.53) after not being exposed to CL-oriented methods. This statistically significant difference was supported by a large effect size (*d* = 1.02, 95% CI = [-1.70, -0.34]).

4.1.1.3. Evolution of Metaphor Density: Pre-Test vs. Post-Test

Table 16 summarises the variations in metaphor use in the oral and written discourse of the control and experimental groups, comparing the pre-test and post-test results for metaphor density. See Tables 87–90 in Appendix N for a detailed breakdown of the findings.

Table 16

Metaphor Density Variations in Participant Groups' Oral and Written Discourse: Pre-Test vs. Post-Test

Discourse	Groups	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				Absolute Increase (AI)	Relative Increase (RI)	Variation (%)
Oral	Control	11.58	9.63	-1.96	0.14	-14
	Experimental	10.88	11.07	0.18	0.06	6
Written	Control	15.69	16.40	0.71	0.07	7
	Experimental	16.63	19.31	2.68	0.18	18

After the instructional intervention, the experimental group demonstrated a favourable progression in their use of metaphors in both oral and written discourse. In contrast, the control group showed both positive and negative variations across both types of discourse. In what follows, a comprehensive analysis of these observed variations is provided.

The results of the *Speaking* analyses revealed that, after being made aware of metaphors in the L2 classroom, the experimental group showed a substantial improvement. Specifically, there was a 6% increase (*p* > .05) in metaphor density in their oral production

(rising from 10.88% to 11.07%). In contrast, the control group's metaphor density in oral discourse statistically significantly decreased by 14% (*W*[162] = 1.821, *p* = .033), declining from 11.58% to 9.63%. This negative variation was supported by a medium effect size (*d* = 0.64, 95% CI = [-0.01, 1.30]). For more details regarding the statistical analyses, see Table 17.

Table 17

Metaphor Density Progress in Learner Oral and Written Discourse of Each Group: A Comparative Analysis of Pre-Test vs. Post-Test

Discourse	Group	MRW (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
		Pre-test	Post-test					
Oral	Control	11.58	9.63	1.821	162	.033	0.64	[-0.01, 1.30]
	Experimental	10.88	11.07	0.076	103	.956	0.07	[-0.71, 0.57]
Written	Control	15.69	16.40	0.864	76	.294	0.27	[-0.91, 0.37]
	Experimental	16.63	19.31	2.625	35	.007	0.97	[-1.65, -0.29]

Concerning the *Writing* analyses, the findings indicated that the teaching treatment had a profound impact on the experimental group's evolution rate in metaphor density with an statistically significant increase of 18% (*W*[35] = 2.625, *p* = .007) in metaphor density, observed in the written essays. This positive variation was accompanied by a large effect size (*d* = 0.97, 95% CI = [-1.65, -0.29]). In contrast, the control group only exhibited a 7% increase (*p* > .05) after not being taught with CL-oriented methods. For further details, see Table 17.

Table 18

Evolution of Metaphor Density in Learner Oral and Written Discourse: A Comparison of Control and Experimental Groups

Discourse	MRW % Variations (AI)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	-1.96	0.18	1.907	146	.149	0.57	[-1.23, 0.08]
Written	0.71	2.68	2.018	146	.149	0.58	[-1.23, 0.07]

The statistical analyses revealed that the differences between both groups reached no significant differences ($p > .05$) in the evolution of metaphor use between the control and experimental groups in both their spoken productions and their written essays after the teaching treatment. For further information, see Table 18 above.

4.1.1.4. Concluding Remarks of RQ1a

This section has explored metaphor density in the oral and written discourse of participants from the control and experimental groups before and after being exposed to CL-oriented teaching methods. The key findings of RQ1a include the following:

1. In the **pre-test**:
    - There was a significantly greater overall occurrence of metaphor for both groups in written than in oral discourse before the teaching treatment.
    - No statistically significant differences in metaphor use between the control and experimental groups in either oral or written discourse were found. This reliable baseline provides a foundation for comparing and evaluating both groups.
2. In the **post-test**:
    - The written essays from both groups exhibited a significantly higher prevalence of metaphors compared to their spoken productions after the longitudinal study.
    - The experimental group exhibited a statistically significant higher use of metaphor in written discourse than the control group. This statistically significant difference was underscored by a large effect size, emphasising a substantial and meaningful impact of the CL-oriented teaching intervention.
    - The metaphor density in the experimental group was also higher than that of the control group in oral discourse but this difference between the two groups did not reach statistical significance at the end of the study.

3. Comparison of control and experimental groups' **evolution**:
    - The experimental group exhibited consistent and positive progress in employing metaphors across both types of discourse. Significantly, the experimental group demonstrated an increase in metaphor use in their written production from the pre-test to the post-test. This statistically significant difference was accompanied by a large effect size, which underlines the potential benefits of the CL-oriented instructional intervention on developing L2 learners' ability to employ metaphors, particularly supporting their written skills.
    - The control group demonstrated an imbalanced evolution in both types of discourse. Notably, the control group exhibited a statistically significant decrease in metaphor usage in their spoken productions, with a medium effect size providing support. This decline highlights a detrimental effect on their ability to employ metaphor, despite any positive progress made in written discourse.
    - No statistically significant differences were found when comparing both groups' metaphor density evolution rates in their spoken productions or written essays. This finding suggests that factors other than the instructional intervention might have also influenced the development of metaphor use in both groups.

4.1.2. RQ1b: Type of Metaphorical Language Used in Oral and Written Discourse

What is the effect of metaphor-mediated instruction on the type (open-class vs. closed-class metaphors) of metaphorical language use in oral and written learner discourse at B2 level?

This section presents the findings of the analysis of the use of open-class metaphors (i.e., nouns, verbs, adjectives, and adverbs) vs. closed-class metaphors (i.e., prepositions) in the oral and written discourse of participants from the control and experimental groups, as measured in the pre-test and post-test.

To address RQ1b, this section is divided into three parts. The first two subsections describe participants' use of open-class and closed-class metaphors in each testing measure,

along with the distribution of the results and the inferential analyses of the differences in metaphor density between the two groups. The third subsection reports the evolution in open- vs. closed-class metaphor use within each group from the pre-test to the post-test and compares the metaphor density rates for each group based on the type of metaphor employed.

4.1.2.1. Type of Metaphors Used (Open-Class vs. Closed-Class Metaphors): Pre-Test

This section summarises the findings on the analysis of the use of open-class (Tables 21–24) and closed-class metaphors (Table 25) in the oral and written pre-tests by the control and experimental groups. See Tables 91–94 and 99–102 in Appendix O for a detailed breakdown of these results.

In the pre-test, both the control and experimental groups exhibited significant higher rates of closed-class metaphors than open-class metaphors in both oral discourse and written discourse ($p < .001$). For further details, see Table 19. Notably, prepositions were the only function words included in the analysis, and they were metaphorically used more frequently than the lexical words examined, such as nouns, verbs, adjectives, and adverbs.

Table 19

A Comparison of Open-Class vs. Closed-Class Metaphor Use in Learner English Discourse: Pre-Test

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Open-class	Closed-class					
Oral							
Control	17.57	52.03	32.716	4	<.001	2.58	[-3.42, -1.72]
Experimental	15.40	53.58	37.958	0	<.001	2.98	[-3.88, -2.06]
Written							
Control	17.84	56.05	37.682	0	<.001	4.14	[-5.25, -3.02]
Experimental	19.58	55.57	35.315	0	<.001	4.19	[-5.31, -3.06]

Table 20

A Comparison of Open-Class vs. Closed-Class Metaphor Use (Oral vs. Written Discourse): Pre-Test

Type of metaphor	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Oral discourse	Written discourse					
Open-class							
Control	17.57	17.84	0.654	176.50	.534	0.05	[-0.67, 0.57]
Experimental	15.40	19.58	4.008	74	<.001	1.20	[-1.87, -0.52]
Closed-class							
Control	52.03	56.05	4.396	169	.409	0.26	[-0.88, 0.36]
Experimental	53.58	55.57	2.383	183	.655	0.13	[-0.75, 0.49]

The analysis revealed that at the pre-test administration, both types of metaphors were observed in higher densities in written discourse. The experimental group showed a clear preference for open-class metaphors in their written essays, using them to a significantly greater extent ($p < .001$) than in their spoken productions. On the other hand, both groups used closed-class metaphors at more similar rates in both their oral and written outputs ($p > .05$). For further details regarding the statistical analyses, see Table 20 above.

To delve deeper into these findings, the following subsections provide additional detailed information in order to shed lights on the patterns observed concerning types of metaphors and groups of participants.

4.1.2.1.1. Use of Open-Class Metaphors

The pre-test analyses indicated that both groups had comparable mean scores for the density of open-class metaphors in their oral and written discourse, as shown in Table 21. Additionally, the control and experimental groups exhibited similar patterns in terms of MRW-type rates of open-class metaphors in both their spoken productions and written essays. These findings are further examined in detail below.

Table 21

Open-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Pre-Test

Discourse	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Oral	Control	1,854	1,094	319	84	17.57	9.90	35.94	6.09
	Experimental	2,271	1,383	347	87	15.40	8.13	26.19	3.65
Written	Control	2,046	1,328	368	111	17.84	10.00	25.00	4.38
	Experimental	2,151	1,389	421	122	19.58	14.29	26.14	3.29

The *Speaking* pre-test analyses showed that the control group exhibited a higher mean score (*M* = 17.57%, *SD* = 6.09) for open-class metaphors in their spoken productions (see Table 21) than the experimental group (*M* = 15.40%, *SD* = 3.65). However, the Wilcoxon test indicated no significant differences ($p > .05$) in the use of open-class metaphors between both groups. For a comprehensive breakdown of the statistical analyses, see Table 22 below.

Table 22

Open-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Pre-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	17.57	15.40	1.776	250	.183	0.43	[-0.20, 1.06]
Written	17.84	19.58	1.604	147	.156	0.45	[-1.07, 0.18]

However, the control and experimental groups showed different uses of open-class metaphors in terms of word class in their pre-test oral discourse (see Table 23). Although both groups share some similarities in metaphor density, particularly in using MRW nouns (24.40% vs. 23.07%), there were noticeable differences. The control group tended to use MRW adjectives more frequently (*M* = 35.90%, *SD* = 26.55) compared to other open-class words, whereas the experimental group showed greater use of MRW nouns (*M* = 23.07%, *SD* = 7.19).

Table 23

Open-Class Metaphor Distribution by Word Class in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

Word class	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Noun	Control	520	391	128	24	24.40	8.33	35.48	8.34
	Experimental	617	473	143	28	23.07	8.33	42.86	7.19
Verb	Control	883	394	153	38	18.04	8.77	40.91	8.01
	Experimental	1,075	500	156	37	15.17	5.97	35.00	5.97
Adjective	Control	98	69	29	15	35.90	0.00	100.00	26.55
	Experimental	139	113	26	16	15.03	0.00	55.56	16.94
Adverb	Control	353	240	9	8	2.81	0.00	11.11	3.72
	Experimental	440	297	22	7	5.12	0.00	18.18	6.09

The analysis showed that both groups exhibited comparable MRW-type rates, as shown in Table 23. Appendices Q–T provide a detailed account of the open-class metaphors used in the oral pre-test. The appendices include MRW nouns (Table 115 in Appendix Q), MRW verbs (Table 119 in Appendix R), MRW adjectives (Table 123 in Appendix S), and MRW adverbs (Table 127 in Appendix T), along with their corresponding relative frequency (RF) values within each group.

The *Writing* pre-test analyses indicated that the experimental group had a higher mean score of 19.58% (*SD* = 3.29) for open-class metaphors in their written essays (see Table 21) compared to the control group (*M* = 17.84% *SD* = 4.38). However, this difference was found no significant ($p > .05$), as indicated by the inferential analyses presented in Table 22 above.

According to Table 24, the control and experimental groups exhibited similar patterns for the distribution of word classes in terms of open-class metaphor uses in written discourse. MRW verbs were the most frequently used open-class metaphor for the control (*M* = 26.10%, *SD* = 7.92) and experimental groups (*M* = 28.12%, *SD* = 8.57) in their post-test written essays. The *Writing* pre-test analysis also revealed that, except for MRW adjectives (15.77 vs. 14.56), the experimental group had a higher average score for metaphor density across all word classes than the control group.

Table 24

Open-Class Metaphor Distribution by Word Class in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

Word class	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Noun	Control	759	635	121	43	15.54	4.00	32.50	6.37
	Experimental	745	614	130	44	17.45	5.71	29.41	6.15
Verb	Control	719	241	186	40	26.10	14.29	44.44	7.92
	Experimental	771	263	216	44	28.12	10.34	44.12	8.57
Adjective	Control	253	216	37	22	15.77	0.00	57.14	13.85
	Experimental	287	240	47	23	14.56	0.00	43.75	13.55
Adverb	Control	315	236	24	8	7.44	0.00	16.67	5.57
	Experimental	348	272	28	13	8.18	0.00	17.65	5.55

The analysis indicated that both groups had similar MRW-type rates, as shown in Table 24. For a detailed account of the open-class metaphors used in the written pre-test, see Appendices Q–T. These include MRW nouns (Table 117 in Appendix Q), MRW verbs (Table 121 in Appendix R), MRW adjectives (Table 125 in Appendix S), and MRW adverbs (Table 129 in Appendix T), along with their corresponding RF values within each group.

4.1.2.1.2. Use of Closed-Class Metaphors

The pre-test analyses indicated that both the control and experimental group participants consistently used closed-class metaphors in their oral and written discourse before the teaching treatment, as shown in Table 25. A more detailed examination of these findings is presented below.

Table 25

Closed-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Pre-Test

Discourse	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Oral	Control	269	63	134	8	52.03	25.00	83.33	17.89
	Experimental	303	73	154	11	53.58	29.17	100.00	17.74
Written	Control	322	47	179	14	56.05	37.50	75.00	12.28
	Experimental	327	60	183	12	55.57	33.33	82.35	11.69

The *Speaking* pre-test indicated that the control group (*M* = 52.03%, *SD* = 17.89) and the experimental group (*M* = 53.58%, *SD* = 17.74) showed similar mean scores ($p > .05$) for the density of closed-class metaphors in their spoken productions. More detailed information regarding the statistical analyses is presented in Table 26 below. Furthermore, Table 25 shows that both groups displayed comparable rates of MRW types (8 vs. 11) in oral discourse.

Table 26

Closed-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Pre-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	52.03	53.58	0.489	195	.903	0.09	[-0.71, 0.53]
Written	56.05	55.57	0.433	206	.882	0.04	[-0.58, 0.66]

The *Writing* pre-test revealed that both groups also performed similarly ($p > .05$) in the use of closed-class metaphors in their written essays, with a ratio of 56.05% (*SD* = 12.28) for the control group and a slower ratio of 55.57% (*SD* = 11.69) for the experimental group. Table 26 above provides an expanded breakdown of the statistical analyses. Additionally, Table 25 illustrates that the rates of MRW types were comparable between the control and experimental groups (14 vs. 12) in their written discourse.

Appendix P provides a more detailed account of the closed-class metaphors used in the oral pre-test (see Table 111) and written pre-test (see Table 113), along with their corresponding RF values within each group.

These findings establish a strong basis for comparing and assessing the frequency of open-class and closed-class metaphors produced by both the control and experimental groups. By establishing a similar level of metaphor usage in the pre-test, the impact of the teaching treatment can be effectively evaluated in the post-test. This enables to draw significant conclusions regarding its effectiveness in enhancing metaphorical language to support L2 learners in developing their vocabulary use at B2 level.

4.1.2.2. Type of Metaphors Used (Open-Class vs. Closed-Class Metaphors): Post-Test

This section presents the findings of the analysis of the use of open-class (Tables 29–32) and closed-class metaphors (Table 33) in the oral and written post-tests by the participants from the control and experimental groups. For a more detailed description of the use of the two types of metaphors, see Tables 95–98 and 103–106 in Appendix O.

In the post-test, the control and experimental groups had significantly higher rates of closed-class metaphors than open-class metaphors in both oral discourse and written discourse ($p < .001$). For comprehensive information on the statistical analyses, see Table 27 below. However, the analysis revealed differences between the two groups in their use of open-class and closed-class metaphors, as will be elaborated upon subsequently.

Table 27

A Comparison of Open-Class vs. Closed-Class Metaphor Use in Learner English Discourse: Post-Test

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Open-class	Closed-class					
Oral							
Control	15.44	31.09	14.772	68	<.001	1.46	[-2.15, -0.75]
Experimental	17.99	30.00	12.606	45	<.001	1.73	[-2.45, -0.99]
Written							
Control	18.70	59.06	38.829	0	<.001	3.63	[-4.64, -2.60]
Experimental	22.94	64.02	40.991	0	<.001	3.75	[-4.78, -2.70]

Table 28

A Comparison of Open-Class vs. Closed-Class Metaphor Use (Oral vs. Written Discourse): Post-Test

Type of metaphor	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Oral discourse	Written discourse					
Open-class							
Control	15.44	18.70	2.979	119	.029	0.70	[-1.33, -0.05]
Experimental	17.99	22.94	5.718	88	.003	0.95	[-1.60, -0.29]
Closed-class							
Control	31.09	59.06	27.500	36	<.001	1.90	[-2.64, -1.14]
Experimental	30.00	64.02	34.512	10.50	<.001	2.86	[-3.75, -1.96]

Analyses also showed that at the post-test administration, both groups demonstrated a significantly greater density of open-class metaphors ($p < .05$) and closed-class metaphors ($p < .001$) in written discourse than oral discourse. For further details, see Table 28 above.

To gain a deeper understanding of these findings, the following subsections offer more information in order to shed light into these patterns regarding types of metaphor and groups of participants.

4.1.2.2.1. Use of Open-Class Metaphors

The post-test analyses indicated that the experimental group showed a higher density of open-class metaphors in their oral and written discourse compared to the control group at the end of the study, as displayed in Table 29. A more in-depth and comprehensive examination of these findings is carried out below.

Table 29

Open-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Post-Test

Discourse	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Oral	Control	1,499	903	238	68	15.44	2.99	25.29	5.58
	Experimental	2,144	1,291	393	103	17.99	8.45	30.36	5.97
Written	Control	2,301	1,269	430	118	18.70	11.48	23.08	3.53
	Experimental	2,357	1,191	537	170	22.94	11.94	32.23	4.35

The *Speaking* post-test analyses revealed that the experimental group achieved a higher mean score (M = 17.99%, SD = 5.97) for open-class metaphor use in their spoken productions (see Table 29) compared to the control group (M = 15.44%, SD = 5.58). However, no statistically significant differences ($p > .05$) were found between both groups. For more detailed information regarding the statistical analyses, see Table 30 below.

Table 30

Open-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Post-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	15.44	17.99	1.681	161	.297	0.44	[-1.06, 0.19]
Written	18.70	22.94	3.845	83.50	.002	1.07	[-1.73, -0.40]

Furthermore, the analysis indicated that MRW nouns were the most frequently used open-class metaphor by both the control (*M* = 21.53%, *SD* = 9.50) and experimental groups (*M* = 22.59%, *SD* = 10.33) in the post-test spoken productions. As shown in Table 31, there were differences in the mean scores for metaphor use between both groups, while the overall distribution of open-class metaphors was similar. It is crucial to highlight that the experimental group exhibited a higher average rate of metaphor density across all word classes than the control group in the oral post-test.

Table 31

Open-Class Metaphor Distribution by Word Class in Oral Discourse (Control vs. Experimental Groups): Post-Test

Word class	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Noun	Control	397	309	84	19	21.53	0.00	42.31	9.50
	Experimental	591	454	136	40	22.59	9.52	41.67	10.33
Verb	Control	706	295	116	33	15.94	4.55	26.67	6.86
	Experimental	959	387	178	39	18.50	0.00	26.67	6.07
Adjective	Control	100	85	16	9	12.45	0.00	42.86	14.33
	Experimental	170	138	32	16	14.00	0.00	50.00	14.67
Adverb	Control	296	215	22	8	6.36	0.00	20.00	7.34
	Experimental	424	312	47	11	11.43	0.00	42.86	9.85

As Table 31 reveals, the MRW-type rates of the control and experimental groups were similar across all word classes except for MRW nouns (19 vs. 40) and MRW adjectives (9 vs. 16). Appendices Q–T provide a detailed account of the open-class metaphors used in the oral post-test. The appendices cover MRW nouns (Table 116 in Appendix Q), MRW verbs (Table

120 in Appendix R), MRW adjectives (Table 124 in Appendix S), and MRW adverbs (Table 128 in Appendix T), along with their corresponding RF values within each group.

The *Writing* post-test analyses showed that the experimental group also had a higher mean score (*M* = 22.94%, *SD* = 4.35) for open-class metaphor use in their written essays (see Table 29) than the control group (*M* = 18.70%, *SD* = 3.53). The Wilcoxon test revealed that the experimental group used significantly more open-class metaphors in written discourse than the control group (*W*[83.50] = 3.845, *p* = .002) at the end of the study. This difference was further supported by a large effect size (*d* = 1.07, 95% CI = [-1.76, -0.39]).

Table 32 shows that the control (*M* = 27.33%, *SD* = 7.46) and experimental groups (*M* = 31.90%, *SD* = 7.80) used MRW nouns most frequently in the post-test written essays. The analysis revealed a balanced overall distribution of open-class metaphor use between the two groups. Furthermore, the experimental group exhibited higher average rates of metaphor density in all word classes in their written discourse than the control group in the post-test.

Table 32

Open-Class Metaphor Distribution by Word Class in Written Discourse (Control vs. Experimental Groups): Post-Test

Word class	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRWs (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Noun	Control	792	576	216	45	27.33	16.22	40.00	7.46
	Experimental	782	533	249	68	31.90	18.42	48.00	7.80
Verb	Control	963	264	155	45	16.12	7.02	27.08	4.96
	Experimental	1035	257	205	66	19.80	9.80	28.57	4.88
Adjective	Control	230	191	38	21	17.09	0.00	44.44	13.93
	Experimental	231	176	55	27	24.10	0.00	52.94	13.28
Adverb	Control	316	238	21	9	6.41	0.00	22.22	5.56
	Experimental	309	225	28	12	9.09	0.00	22.22	5.93

Both the control and experimental groups had similar rates of MRW types, except for MRW nouns (45 vs. 68) and MRW verbs (45 vs. 66), as presented in Table 32. For a more detailed account of the open-class metaphors used in the written post-test, see Appendices Q–T. These include MRW nouns (Table 118 in Appendix Q), MRW verbs (Table 122 in

Appendix R), MRW adjectives (Table 126 in Appendix S), and MRW adverbs (Table 130 in Appendix T), along with their corresponding RF values within each group.

4.1.2.2.2. Use of Closed-Class Metaphors

The post-test analyses revealed that both the control and experimental groups exhibited similar trends in their use of closed-class metaphors in oral discourse at the end of the study, as shown in Table 33. However, notable differences were observed in the context of written production. To delve deeper into these findings, a more detailed examination follows.

The *Speaking* post-test analyses showed that the control group had a similar mean score of closed-class metaphors (*M* = 31.09%, *SD* = 14.10) in their spoken productions compared to the experimental group (*M* = 30%, *SD* = 7.80). No statistically significant differences ($p > .05$) were observed between both groups of participants. For a more thorough breakdown of the statistical analyses, see Table 34 below. Additionally, Table 33 indicates that both groups had similar rates of MRW types (7 vs. 9) in their spoken outputs.

Table 33

Closed-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Post-Test

Discourse	Groups	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	MRW (types)	Metaphor density (%)			
						M	*Min.*	*Max.*	*SD*
Oral	Control	229	99	72	7	31.09	11.76	60.00	14.10
	Experimental	358	168	109	9	30.00	16.67	47.06	7.80
Written	Control	317	39	189	12	59.06	37.50	87.50	15.33
	Experimental	335	38	211	15	64.02	33.33	88.24	14.87

Table 34

Closed-Class Metaphor Use in Oral and Written Discourse (Control vs. Experimental Groups): Post-Test Comparative Analysis

Discourse	Metaphor density (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	31.09	30.00	0.405	205.50	.892	0.10	[-0.53, 0.72]
Written	59.06	64.02	5.721	156.50	.245	0.33	[-0.95, 0.30]

The *Writing* post-test analyses indicated that the experimental group showed a higher mean score (*M* = 64.02%, *SD* = 14.87) in the production of written essays compared to the control group (*M* = 59.06%, *SD* = 15.33). However, this difference was not found to be statistically significant ($p > .05$), as indicated in Table 34 above. Furthermore, Table 33 shows that both groups had comparable rates of MRW types (12 vs. 15) in their written outputs after the teaching treatment.

For a more comprehensive understanding of the closed-class metaphors used in the oral post-test (see Table 112) and written post-test (see Table 114), along with their corresponding RF values within each group, see Appendix P.

4.1.2.3. Open-Class vs. Closed-Class Metaphor Growth: Pre-Test vs. Post-Test

This section summarises the variations in the use of open-class (Table 35) and closed-class metaphors (Table 38) in oral and written discourse by the control and experimental groups, comparing the pre-test and post-test results. For a more detailed description of these findings by participants within each group, see Tables 107–110 in Appendix O.

After the teaching treatment, both the control and experimental groups showed differences in their use of open-class metaphors in oral and written discourse. Conversely, similar trends in closed-class metaphor density were found between the two groups across both types of discourse. To enhance the depth of understanding, the following subsections present additional detailed information that elucidates these patterns.

4.1.2.3.1. Evolution of Open-Class Metaphor Usage

As Table 35 indicates, the experimental group exhibited a similar increase in their use of open-class metaphors in both oral and written discourse after receiving the teaching treatment. In contrast, the control group showed positive and negative progress in their use of open-class metaphors across different types of discourse. It is important to highlight that the specific progression of individual open-class metaphors, namely, nouns, verbs, adjectives, and adverbs, was not the main focus of this study and therefore not explicitly assessed.

However, this aspect presents an avenue for further research to explore and measure such evolution in more detail. With that in mind, the overall results are reported as follows.

Table 35

Variations in Open-Class Metaphor Use in Learner Oral and Written Discourse: Pre-Test vs. Post-Test

Discourse	Groups	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				Absolute Increase (AI)	Relative Increase (RI)	Variation (%)
Oral	Control	17.57	15.44	-2.13	-0.05	-5%
	Experimental	15.40	17.99	2.59	0.20	20%
Written	Control	17.84	18.70	0.86	0.11	11%
	Experimental	19.58	22.94	3.36	0.19	19%

Regarding oral discourse, the experimental group exhibited an increase of 20% (rising from 15.28% to 18.33%) in their use of open-class metaphors following exposure to CL-oriented methods. In contrast, the control group showed an 5% decrease (declining from 17.21% to 15.88%) in their use of open-class metaphors after not being exposed to the same teaching approach. However, the variations in the use of open-class metaphors in the spoken productions of each group did not reach statistical significance ($p > .05$). For a more comprehensive description of the statistical analyses, see Table 36 below.

Table 36

Variations of Open-Class Metaphor Use in Learner Oral and Written Discourse of Each Group: A Comparative Analysis of Pre-Test vs. Post-Test

Discourse	Group	MRW (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
		Pre-test	Post-test					
Oral	Control	17.57	15.44	2.034	130	.368	0.36	[-0.26, 0.99]
	Experimental	15.40	17.99	2.593	63	.123	0.52	[-1.15, 0.11]
Written	Control	17.84	18.70	0.908	90	.596	0.22	[-0.84, 0.41]
	Experimental	19.58	22.94	3.455	32	.005	0.87	[-1.52, -0.22]

As for written discourse, both groups showed positive evolution in their use of open-class metaphors at the end of the study. However, the experimental group exhibited a larger

growth rate of 19% (rising from 19.57% to 22.78%) compared to the control group's 11% increase (p > .05), rising from 17.99% to 18.69%. The Wilcoxon test revealed that the experimental group exhibited a significant increase (W[32] = 3.455, p = .005) in their use of open-class metaphors in written discourse, with a large effect size (d = 0.87, 95% CI = [-1.52, -0.22]), after receiving metaphor-focused instruction. For further details, see Table 36 above.

The statistical analyses revealed significant differences between the two groups in terms of their evolution in using open-class metaphors in oral discourse (W[127] = 4.165, p = .049) after the CL-oriented instructional intervention. This finding was further supported by a medium effect size (d = 0.70, 95% CI = [-1.34, 0.06]). However, no significant differences (p > .05) were observed between both groups' progress in written discourse. For more detailed information on the statistical analyses conducted, see Table 37 below.

Table 37

Evolution of Open-Class Metaphor Use in Learner Oral and Written Discourse: A Comparison of Control and Experimental Groups

Discourse	MRW % Variations (AI)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	-2.13	2.59	4.165	127	.049	0.70	[-1.34, -0.06]
Written	0.86	3.36	2.617	144	.134	0.51	[-1.13, 0.13]

4.1.2.3.2. Evolution of Closed-Class Metaphor Usage

Table 38 indicates that the control and the experimental groups showed a negative trend in using closed-class metaphors in oral discourse. In contrast, a positive trend was found in written discourse for both groups. Further exploration of these findings is presented below.

Table 38

Variations in Closed-Class Metaphor Use in Learner Oral and Written Discourse: Pre-Test vs. Post-Test

Discourse	Groups	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				Absolute Increase (AI)	Relative Increase (RI)	Variation (%)
Oral	Control	52.03	31.09	-20.94	-0.30	-30%
	Experimental	53.58	30.00	-23.58	-0.37	-37%
Written	Control	56.05	59.06	3.01	0.10	10%
	Experimental	55.57	64.02	8.45	0.21	21%

Concerning oral discourse, the experimental group showed a higher decline in the use of closed-class metaphors, exhibiting a 37% decrease (declining from 50.83% to 30.45%) in their spoken productions compared to the control group's 30% decrease (declining from 49.81% to 31.44%) after the teaching treatment. The statistical analyses revealed significant decreases, accompanied by large effect sizes, in the use of closed-class metaphors for both the control (W[184] = 21.671, p = .003, d = 1.30, 95% CI= [0.61, 1.98]) and experimental groups W[198] = 21.681, p < .001, d = 1.72, 95% CI = [0.98, 2.44]) in their spoken productions.

Table 39

Variations of Closed-Class Metaphor Use in Learner Oral and Written Discourse of Each Group: A Comparative Analysis of Pre-Test vs. Post-Test

Discourse	Group	MRW (%)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
		Pre-test	Post-test					
Oral	Control	52.03	31.09	21.671	184	.003	1.30	[0.61, 1.98]
	Experimental	53.58	30.00	21.681	198	<.001	1.72	[0.98, 2.44]
Written	Control	56.05	59.06	2.869	77	.481	0.22	[-0.84, 0.41]
	Experimental	55.57	64.02	6.992	50	.040	0.63	[1.26, 0.01]

As for written discourse, the experimental group showed a greater increase of 21% (from 55.96% to 62.99%) compared to the control group's 10% increase (p > .05), rising from 55.59% to 59.62%, after the teaching treatment. As Table 39 shows, the experimental group

significantly increased (W[50] = 6.992, p = .040) in their use of closed-class metaphors in written discourse, supported by a medium effect size (d = 0.63, 95% CI = [1.26, 0.01]).

According to Table 40, the Wilcoxon analyses revealed no significant differences (p < .05) in the evolution of using closed-class metaphor between the control and experimental groups. This was observed in both their spoken productions and in their written essays.

Table 40

Evolution of Closed-Class Metaphor Use in Learner Oral and Written Discourse: A Comparison of Control and Experimental Groups

Discourse	MRW % Variations (AI)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Oral	-20.94	-23.58	1.851	208	.839	0.11	[-0.51, 0.73]
Written	3.01	8.45	4.551	172	.461	0.30	[-0.92, 0.33]

4.1.2.4. Concluding Remarks of RQ1b

This section has compared the use of open-class vs. closed-class metaphors in the oral and written discourse of control and experimental groups before and after receiving metaphor-mediated instruction. The key findings of RQ1b are as follows:

1. In the **pre-test**:
    - Both groups showed a significantly greater frequency of closed-class compared to open-class metaphors, with this trend being more pronounced in written discourse.
    - No statistically significant differences were found between groups in their use of open-class and closed-class metaphors in either oral or written discourse. This finding establishes a baseline for further comparisons and analyses.
2. In the **post-test**:
    - Closed-class metaphors were significantly more prevalent compared to open-class metaphors, especially in the written essays produced by the participants.
    - The experimental group exhibited a significantly higher overall use of open-class metaphors in written discourse than the control group, with a large effect size.

Additionally, the metaphor density of open-class metaphors in the experimental group was higher than that of the control group in oral discourse, although this difference did not reach statistical significance.

- No statistically significant differences were found between the two groups in their usage of closed-class metaphors in both oral and written discourse. However, while similar usage was observed in oral discourse, the experimental group exhibited a higher frequency of closed-class metaphors in written discourse.

3. Control and experimental groups' **evolution** comparison:
    - The experimental group demonstrated significant evolution in their use of open- and closed-class metaphors in written discourse, with large and medium effect sizes, respectively. While the experimental group showed a greater increase in their use of open-class metaphors in oral discourse compared to the control group, who had a decrease in usage, the differences in the use of open-class metaphors in the spoken productions of each group did not reach statistical significance.
    - Both the control and experimental groups exhibited significant decreases with large effect sizes in their use of closed-class metaphors in oral discourse.
    - Significant differences were identified in the evolution of open-class metaphors in oral discourse between the two groups, accompanied by a medium effect size. However, no significant differences were observed when comparing both groups' evolution of using open-class metaphors in written discourse or closed-class metaphors in both oral and written production.

4.2. RQ2: ENGLISH LANGUAGE PEFORMANCE AT B2 LEVEL

How does raising metaphor awareness in topic-based instruction affect L2 learners' achievement at B2 level, as measured by the B2 First for Schools Cambridge English qualification, compared to not using CL-oriented methodological techniques?

This section examines the impact of metaphor-mediated instruction on the English language proficiency of participants, specifically focusing on their overall B2 achievement and individual performance in the productive skills of the *Speaking* and *Writing* tests within the widely recognised *B2 First for Schools Cambridge English* examination.

To address RQ2, this section is structured into three parts. The first two subsections present the results from the pre-test and post-test, respectively, followed by the inferential analyses of the differences between the two groups in relation to their overall B2 performance, as well as their individual performance in the productive skills of the *Speaking* and *Writing* tests. The third subsection analyses the evolution of each group from the pre-test to the post-test and compares their rates for achievement at the upper-intermediate level.

4.2.1. Achievement of L2 Learners at B2 Level: Pre-Test

This section presents the results of the pre-test analysis of the B2 language proficiency of the control and experimental groups. This includes findings on their overall achievement (Table 41), as well as their performance in the *Speaking* (Table 42) and *Writing* (Table 44) tests.

In the pre-test, both the control and experimental groups showed similar levels of performance in terms of their overall B2 achievement. Likewise, their proficiency in the *Speaking* and *Writing* skills demonstrated comparable patterns. Both groups of participants globally reached B2 level in their oral skills, while their overall performance and written skills were below this level before the teaching treatment. The following subsections will provide a comprehensive analysis of these observed trends.

4.2.1.1. Overall Proficiency at B2 Level

The pre-test analysis indicated that neither the control nor the experimental group globally achieved a minimum score of 160 in their overall B2 performance, which is necessary to reach B2 level (UCLES, 2015). Table 41 presents the pre-test mean scores for each group, revealing similar results between the control (M = 145.75, SD = 10.03) and experimental

groups (*M* = 145.55, *SD* = 16.70), with no significant differences observed (*W*[233] = 3.000, *p* = .379, *d* = 0.01, 95% CI [-0.63, 0.65]).

Table 41

Overall Performance at B2 Level (Control vs. Experimental Groups): Pre-Test

Groups	*Cambridge English Scale* Scores								
	Overall B2 level score				Rd.	UoE	Wr.	Ls.	Sp.
	M	*Min.*	*Max.*	*SD*					
Control	145.75	128	164	10.03	155.55	97.90	148.05	129.15	163.45
Experimental	145.55	125	185	16.70	140.75	102.15	142.65	123.40	166.35

While the experimental group scored slightly lower in the *Reading*, *Writing*, and *Listening* tests, they performed slightly better in the *Use of English* and *Speaking* tests before the instructional intervention. For a detailed breakdown of the pre-test results of the overall B2 performance by participants within each group, see Tables 131 and 132 in Appendix U.

4.2.1.2. Speaking Proficiency at B2 Level

The *Speaking* pre-test analysis revealed that both groups achieved a Grade C (160–172), indicating that they reached a global B2 level proficiency in *Speaking,* specifically focusing on the overall score of the *Speaking* test (UCLES, 2015). Table 42 shows that the experimental group scored higher, with a mean of 166.35 (*SD* = 9.21), compared to the control group's mean score of 163.45 (*SD* = 5.97). Additionally, both the control and experimental groups also exhibited consistent performance across the assessment descriptors used to evaluate oral production in the pre-test. Tables 133 and 134 in Appendix U provide a comprehensive description of the pre-test *Speaking* results by participants within each group.

Table 42

Speaking Performance at B2 Level (Control vs. Experimental Groups): Pre-Test

Groups	*Cambridge English Scale* Scores								
	Overall *Speaking* Score				Gr. & vocab.	Disc. mgmt.	Pron.	Inter. comm.	Global mark
	M	*Min.*	*Max.*	*SD*					
Control	163.45	151	176	5.97	3.10	3.25	3.28	3.23	3.18
Experimental	166.35	151	184	9.21	3.13	3.55	3.43	3.48	3.30

The Wilcoxon tests indicated no statistically significant differences ($p > .05$) between the control and experimental groups in terms of their *Speaking* performance prior to the longitudinal study. Similarly, the primary focus of this study, vocabulary production, did not yield any statistically significant differences ($p > .05$) between both groups before the teaching treatment. However, the analyses revealed that the experimental group showed a statistically significantly superior performance ($W[125] = 0.500$, $p = .030$) in specific assessment descriptors such as "Discourse Management", supported by a medium effect size ($d = 0.74$, 95% CI [-1.40, -0.08]). For a more detailed breakdown of these results, see Table 43 below.

Table 43

Speaking Achievement at B2 Level (Control vs. Experimental Groups): Pre-Test Comparative Analysis

Speaking Performance	B2 proficiency score		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	163.45	166.35	2.000	163	.321	0.37	[-1.02, 0.27]
Gr. & vocab.	3.10	3.13	0.000	199	.988	0.04	[-0.69, 0.60]
Disc. mgmt.	3.25	3.55	0.500	125	.030	0.74	[-1.40, -0.08]
Pron.	3.28	3.43	0.000	180	.565	0.31	[-0.95, 0.33]
Inter. comm.	3.23	3.48	0.000	138.50	.076	0.61	[-1.26, 0.05]
Global mark	3.18	3.30	0.000	174	.433	0.29	[-0.94, 0.35]

4.2.1.3. Writing Proficiency at B2 Level

The *Writing* pre-test analysis indicated that neither the control nor the experimental group globally achieved a minimum score of 160 to reach B2 level (UCLES, 2015). Table 44 illustrates that both the control and experimental groups scored similarly in their written production, specifically in *Writing* Part 1 (essay writing), before the teaching treatment. The experimental group obtained a higher mean score ($M = 138.40$, $SD = 23.51$) compared to the experimental group ($M = 136.35$, $SD = 8.47$). In addition, both groups performed similarly across the *Writing* assessment descriptors in the pre-test. For a detailed description of pre-test *Writing* scores by participants within each group, see Tables 135 and 136 in Appendix U.

Table 44

Writing Performance at B2 Level (Control vs. Experimental Groups): Pre-Test

Groups	*Cambridge English Scale* Scores							
	Writing Score (Part 1)				Content	Comm. achv.	Organisation	Language
	M	*Min.*	*Max.*	*SD*				
Control	136.35	130	160	8.47	2.13	2.15	2.33	2.08
Experimental	138.40	53	180	23.51	2.25	2.20	2.68	2.18

Table 45

Writing Achievement at B2 Level (Control vs. Experimental Groups): Pre-Test Comparative Analysis

Writing Performance	B2 proficiency score		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	136.35	138.40	7.000	141	.103	0.12	[-0.76, 0.52]
Content	2.13	2.25	0.000	176.5	.440	0.26	[-0.90, 0.38]
Comm. achv.	2.15	2.20	0.000	189	.717	0.11	[-0.75, 0.53]
Organisation	2.33	2.68	0.500	118	.020	0.67	[-1.33, -0.02]
Language	2.08	2.18	0.000	188	.708	0.20	[-0.84, 0.44]

The statistical analyses revealed that there were no significant differences ($p > .05$) between the control and experimental groups regarding their *Writing* performance and vocabulary production before the teaching treatment. However, the Wilcoxon test indicated that the experimental group showed a statistically significant advantage in the "Organisation" assessment descriptor ($W[118] = 0.500$, $p = .020$), with a medium effect size ($d = 0.67$, 95% CI [-1.33, -0.02]). For a more extensive breakdown of these findings, see Table 45 above.

These findings highlight the initial similarity in English language proficiency levels between both groups, as evidenced by their comparable overall performance and language production skills in both *Speaking* and *Writing*. Although the pre-test results reveal specific advantages of the experimental group over the control group in certain areas, these are unrelated to the focus of study, namely vocabulary differences. Therefore, findings provide a baseline for assessing the effects of the teaching treatment in the post-test phase.

4.2.2. Achievement of L2 Learners at B2 Level: Post-Test

This section presents the results of the post-test analysis of the B2 proficiency of the control and experimental groups. This section includes participants' overall achievement (Table 46), as well as their performance in the *Speaking* (Table 47) and *Writing* (Table 49) tests.

In the post-test, despite the fact that neither the control nor experimental group reached a global B2 level in their overall performance, both groups showed a global B2 level of English proficiency in *Speaking* and *Writing*. Although the two groups exhibited consistent overall performance, the experimental group performed better in *Speaking* skills, whereas the control did better in *Writing* skills at the post-test administration. To further explore these findings, the following subsections delve into additional detailed information.

4.2.2.1. Overall Proficiency at B2 Level

The post-test analysis revealed that neither the control nor experimental group globally reached B2 level of English proficiency in their overall performance, as their mean scores were below the minimum score of 160 (UCLES, 2015). Table 46 shows that the control group (M = 156.65, SD = 9.16) achieved a higher mean score than the experimental group (M = 154.20, SD = 16.28). However, no significant differences (W[238] = 3.000, p = .309, d = 0.19, 95% CI [-0.46, 0.83]) were observed between both groups' overall performance at the end of the study.

Table 46

Overall Performance at B2 Level (Control vs. Experimental Groups): Post-Test

Groups	*Cambridge English Scale* Scores								
	Overall B2 level score				Rd.	UoE	Wr.	Ls.	Sp.
	M	*Min.*	*Max.*	*SD*					
Control	156.65	137	172	9.16	146.05	134.05	172.90	137.95	165.30
Experimental	154.20	129	188	16.28	144.55	117.40	162.35	141.15	171.85

This difference in performance between groups was also observed in the individual post-test comprehension and production tests. Despite scoring lower in *Reading*, *Use of*

English, and *Writing*, the experimental group demonstrated superior *Listening* and *Speaking* performance after the CL-oriented instructional intervention. A comprehensive breakdown of the global post-test results of the overall B2 performance by participants within each group can be found in Tables 137 and 138 in Appendix U.

4.2.2.2. Speaking Proficiency at B2 Level

The *Speaking* post-test analysis showed that the control and experimental groups achieved a Grade C (160–172), indicating that both groups reached a global B2 level of English proficiency in *Speaking* (UCLES, 2015). Tables 139 and 140 in Appendix U present a more detailed description of the post-test *Speaking* results by participants within each group.

Table 47

Speaking Performance at B2 Level (Control vs. Experimental Groups): Post-Test

Groups	*Cambridge English Scale* Scores								
	Overall *Speaking* Score				Gr. & vocab.	Disc. mgmt.	Pron.	Inter. comm.	Global mark
	M	*Min.*	*Max.*	*SD*					
Control	165.30	157	177	5.52	3.03	3.30	3.38	3.45	3.35
Experimental	171.85	160	188	7.99	3.45	3.73	3.63	3.73	3.70

After the CL-oriented instructional intervention, the experimental group exhibited a significantly higher (W[94.50] = 5.000, p = .004) level of English proficiency (M = 171.85, SD = 7.99) than the control group (M = 165.30, SD = 5.52), as shown in Table 47. This statistically significant difference was supported by a large effect size (d = 0.95, 95% CI = [-1.63, -0.28]).

Table 48

Speaking Achievement at B2 Level (Control vs. Experimental Groups): Post-Test Comparative Analysis

Speaking performance	B2 proficiency score		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	165.30	171.85	5.000	94.50	.004	0.95	[-1.63, -0.28]
Gr. & vocab.	3.03	3.45	0.500	99.50	.003	1.01	[0.33, 1.69]
Disc. mgmt.	3.30	3.73	0.500	95	.003	1.08	[-1.76, -0.39]
Pron.	3.38	3.63	0.000	148.50	.140	0.53	[-1.18, 0.12]
Inter. comm.	3.45	3.73	0.500	126.50	.031	0.74	[-1.41, -0.08]
Global mark	3.35	3.70	0.500	110	.007	0.94	[-1.62, 0.27]

The experimental group outperformed the control group across all five *Speaking* descriptors after receiving B2 training using CL-oriented methods. The experimental group exhibited a statistically significant outperformance (W[99.50] = 0.500, p = .003) compared to the control group on the "Grammar and Vocabulary" assessment descriptor (3.03 vs. 3.45). This statistically significant difference was accompanied by a large effect size (d = 1.01, 95% CI = [0.33, 1.69]). It is important to note that the experimental group achieved statistically significant higher scores, in all descriptors within the *Speaking* test, except for "Pronunciation" (p > .05) at the post-test administration. For further details, see Table 48 above.

4.2.2.3. Writing Proficiency at B2 Level

The *Writing* post-test analysis revealed that the control and experimental groups achieved a Grade C (160–172), indicating both groups reached a global B2 level of English proficiency in *Writing*. For a more detailed breakdown of the post-test *Writing* results by participants within each group, see Tables 141 and 142 in Appendix U.

Table 49

Writing Performance at B2 Level (Control vs. Experimental Groups): Post-Test

Groups	*Cambridge English Scale* Scores							
	Writing Score (Part 1)				Content	Comm. achv.	Organisation	Language
	M	*Min.*	*Max.*	*SD*				
Control	167.50	157	180	6.00	3.50	3.40	3.58	3.13
Experimental	163.40	147	190	9.81	3.33	3.18	3.38	3.05

Table 49 shows that the teaching treatment had a more limited effect in supporting written skills, with the control group scoring significantly higher (W[277] = 5.000, p = .037) than the experimental group (M = 167.50, SD = 6.00 vs. M = 163.40, SD = 9.81). This statistically significant difference was supported by a medium effect size (d = 0.50, 95% CI = [-0.15, 1.15]).

Table 50

Writing Achievement at B2 Level (Control vs. Experimental Groups): Post-Test Comparative Analysis

Writing performance	B2 proficiency score		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	167.50	163.40	5.000	277	.037	0.50	[-0.15, 1.15]
Content	3.50	3.33	0.411	260	.086	0.38	[-0.27, 1.02]
Comm. achv.	3.40	3.18	0.500	277	.026	0.46	[-0.18, 1.11]
Organisation	3.58	3.38	0.500	271.50	.037	0.43	[-0.21, 1.08]
Language	3.13	3.05	0.000	236.50	.288	0.14	[-0.50, 0.78]

Additionally, the control group exhibited higher performance across all four *Writing* assessment descriptors. It is worth noting that there were similarities between both groups in the "Language" descriptor (3.13 vs. 3.05), and no significant differences were observed ($p > .05$). However, the control group exhibited a statistically significant performance advantage supported by medium effect sizes, compared to the experimental group in terms of "Communicative Achievement" (W[277] = 0.500, p = .026, d = 0.46, 95% CI = [-0.18, 1.11]) and "Organisation" (W[271.50] = 0.500, p = .037, d = 0.43, 95% CI = [-0.21, 1.08]). For a more comprehensive overview of the findings, see Table 50 above.

4.2.3. Evolution of L2 Learners' Achievement at B2 Level: Pre-Test vs. Post-Test

This section explores the evolution of the control and experimental groups' achievement at B2 level by comparing the pre-test and post-test results. The findings are summarised in Table 51, which includes both groups' global evolution rates in their overall achievement and their rates in *Speaking* and *Writing* proficiency.

Table 51

Variations in Participant Groups' Achievement at B2 Level: Pre-Test vs. Post-Test Performance

B2 Performance	Groups	Pre-test score	Post-test score	Evolution rates		
				Absolute Increase (AI)	Relative Increase (RI)	Variation (%)
Overall	Control	145.75	156.65	10.90	0.08	8
	Experimental	145.55	154.20	8.65	0.06	6
Speaking	Control	163.45	165.30	1.85	0.01	1
	Experimental	166.35	171.85	5.50	0.03	3
Writing	Control	136.35	167.50	31.15	0.23	23
	Experimental	138.40	163.40	25.00	0.18	18

After the teaching treatment, both the control and experimental groups showed favourable progress in their overall achievement, along with their *Speaking* and, notably, *Writing* proficiency at B2 level. While both groups showed similar evolution rates of overall and *Speaking* performance, the control group exhibited greater evolution in *Writing*. The following subsections will present a more comprehensive analysis of the results, enabling a deeper understanding of these findings.

4.2.3.1. Overall Proficiency at B2 Level

The results of the analysis revealed that both the control and experimental groups demonstrated a similar increase in their overall performance at the end of the study. As shown in Table 51 above, the control group showed a statistically significant growth rate of 8% (W[1.50] = 10.601, $p < .001$), rising from 145.75 to 156.65, with a large effect size (d = 1.14, 95% CI = [-1.82, -0.45]). Similarly, the experimental group exhibited a lower but still statistically significant growth rate of 6% (W[7] = 8.500, $p < .001$), rising from 145.55 to 154.20, with a medium effect size (d = 0.52, 95% CI = [-1.18, 0.13]). However, despite this considerable progress observed in both groups, they did not attain B2 level after the instructional intervention. For a more detailed breakdown of these results by participants within each group, see Tables 143 and 144 in Appendix U.

The inferential comparison of the evolution of the two groups during the study confirmed that there were no statistically significant differences (W[324] = 2.000, p = .364, d = 0.33, 95% CI = [-0.32, 0.97]) between the groups in terms of overall achievement at B2 level.

4.2.3.2. Speaking Proficiency at B2 Level

The findings indicated differences in the evolution of *Speaking* proficiency between the control and experimental groups. Table 52 shows that the teaching treatment had a stronger impact on the experimental group's *Speaking*, resulting in a statistically significant increase of 3% (W[22.50] = 6.500, p = .006) with a medium effect size (d = 0.64, 95% CI = [-1.29, 0.02]). Specifically, their proficiency level increased from 166.35 to 171.85. In contrast, the control group showed a 1% increase (p > .05), rising from 163.45 to 165.30. It is worth noting that, despite the varying rates of improvement, both groups successfully achieved B2 level proficiency in *Speaking* at the end of the study. For a more comprehensive examination of these findings by participants within each group, see Tables 145–156 in Appendix U.

Table 52

Speaking Performance Variations in Participant Groups' at B2 Level: Pre-Test vs. Post-Test

Speaking performance	Groups	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
Overall	Control	163.45	165.30	1.85	0.01	1
	Experimental	166.35	171.85	5.50	0.03	3
Gr. & vocab.	Control	3.10	3.03	-0.08	-0.02	-2
	Experimental	3.13	3.45	0.33	0.14	14
Disc. mgmt.	Control	3.25	3.30	0.05	0.02	2
	Experimental	3.55	3.73	0.18	0.06	6
Pron.	Control	3.28	3.38	0.10	0.10	10
	Experimental	3.43	3.63	0.20	0.07	7
Inter. comm.	Control	3.23	3.45	0.23	0.08	8
	Experimental	3.48	3.73	0.25	0.08	8
Global mark	Control	3.18	3.35	0.18	0.06	6
	Experimental	3.30	3.70	0.40	0.14	14

However, the impact of the instructional intervention on *Speaking* performance is evident in the remarkable progress of the experimental group across multiple *Speaking* assessment descriptors, as illustrated in Table 52 above. Specifically, in terms of the

"Grammar and Vocabulary" descriptor, the experimental group exhibited a statistically significant increase of 14% (*W*[15.50] = 0.500, *p* = .032), rising from 3.13 to 3.45, with a medium effect size (*d* = 0.52, 95% CI = [-1.17, 0.13]). Conversely, the control group decreased by 2% (*p* > .05), declining from 3.10 to 3.03 in the same assessment descriptor. A similar trend was observed in the "Global mark" assessment descriptor, where the experimental group demonstrated a statistically significant 14% improvement (*W*[22.50] = 0.500, *p* = .004) with a large effect size (*d* = 0.91, 95% CI = [-1.58, 0.23]), rising from 3.30 to 3.70. In contrast, the control group exhibited an increase of 6% (*p* > .05), rising from 3.18 to 3.35. Table 53 provides further details on the statistical analyses.

Table 53

Progress in Speaking Performance at B2 Level for Each Group: A Comparative Analysis of Pre-Test vs. Post-Test

Speaking performance	Groups	Pre-test	Post-test	diff.	*W* score	*p* value	Cohen's *d*	95% CI
Overall	Control	163.45	165.30	2.000	54.50	.183	0.32	[-0.97, 0.32]
	Experimental	166.35	171.85	6.500	22.50	.006	0.64	[-1.29, 0.02]
Gr. & vocab.	Control	3.10	3.03	0.000	42	.393	0.25	[-0.40, 0.89]
	Experimental	3.13	3.45	0.500	15.50	.032	0.52	[-1.17, 0.13]
Disc. mgmt.	Control	3.25	3.30	0.000	13.50	.530	0.15	[-0.79, 0.49]
	Experimental	3.55	3.73	0.500	15	.097	0.39	[-1.04, 0.26]
Pron.	Control	3.28	3.38	0.250	13.50	.276	0.29	[-0.93, 0.36]
	Experimental	3.43	3.63	0.500	9	.052	0.35	[-0.99, 0.30]
Inter. comm.	Control	3.23	3.45	0.500	5	.015	0.64	[-1.30, 0.02]
	Experimental	3.48	3.73	0.500	13	.014	0.58	[-1.24, 0.07]
Global mark	Control	3.18	3.35	0.500	10	.059	0.50	[-1.15, 0.15]
	Experimental	3.30	3.70	0.500	22.50	.004	0.91	[-1.58, -0.23]

The control group only exhibited statistically significant growth (8%) in the "Interactive Communication" assessment descriptor (*W*[5] = 0.500, *p* = .015) with a medium effect size (*d* = 0.64, 95% CI = [-1.30, 0.02]), rising from 3.23 to 3.45. However, the experimental group also showed a statistically significant increase of 8% (*W*[13] = 0.500, *p* = .014) with a medium effect size (*d* = 0.58, 95% CI = [-1.24, 0.07]) in this assessment descriptor, rising from 3.48 to 3.73. For further details regarding statistical analyses, see Table 53 above.

When comparing both group's evolution, the experimental group exhibited statistically significantly greater growth in the "Grammar and Vocabulary" assessment descriptor (*W*[112] = 0.500, *p* = .013), showing statistically significant progress with a large effect size (*d* = 0.81, 95% CI [-1.48, -0.14]). However, no statistically significant differences (*p* > .05) were observed in overall *Speaking* proficiency when compared to the control group after the CL-oriented instructional intervention. For more detailed information, see Table 54.

Table 54

Evolution of Speaking Performance at B2 Level: A Comparison of Control and Experimental Groups

Speaking performance	Variations (AI)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	1.85	5.50	4.000	135	.080	0.57	[-1.22, 0.08]
Gr. & vocab.	-0.08	0.33	0.500	112	.013	0.81	[-1.48, -0.14]
Disc. mgmt.	0.05	0.18	0.000	169.50	.374	0.33	[-0.97, 0.32]
Pron.	0.10	0.20	0.000	174	.450	0.25	[-0.89, 0.39]
Inter. comm.	0.23	0.25	0.000	187.50	.722	0.07	[-0.71, 0.57]
Global mark	0.18	0.40	0.500	134.50	.061	0.53	[-1.18, 0.12]

4.2.3.3. Writing Proficiency at B2 Level

The analyses revealed that both the control and experimental groups showed favourable progress in their *Writing* proficiency after the instructional intervention, resulting in achieving B2 level for both groups. Table 55 clearly demonstrate this, with the experimental group achieving a statistically significant increase of 18% (*W* [1] = 22.437, *p* < .001, *d* = 1.39, 95% CI = [-2.10, -0.67]), rising from 138.40 to 163.40.

Similarly, the control group also exhibited a statistical significant improvement of 23%(*W*[0] = 31.000, *p* < .001, *d* = 4.24, 95% CI = [-5.40, -3.09]), rising from 136.35 to 167.50. It is important to acknowledge that these improvements were supported by large effect sizes in both group of participants. For a detailed breakdown of individual evolution rates within each group, see Tables 157–166 in Appendix U.

Table 55

Writing Performance Variations in Participant Groups at B2 Level: Pre-Test vs. Post-Test

Writing performance	Groups	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
Overall	Control	136.35	167.50	31.15	0.23	23
	Experimental	138.40	163.40	25.00	0.18	18
Content	Control	2.13	3.50	1.38	0.67	67
	Experimental	2.25	3.33	1.08	0.56	56
Comm. achv.	Control	2.15	3.40	1.25	0.58	58
	Experimental	2.20	3.18	0.98	0.44	44
Organisation	Control	2.33	3.58	1.25	0.57	57
	Experimental	2.68	3.38	0.70	0.35	35
Language	Control	2.08	3.13	1.05	0.54	54
	Experimental	2.18	3.05	0.88	0.46	46

The progress in *Writing* showed positive trends across various assessment descriptors, as illustrated in Table 55. Regarding the “Language” descriptor, both the control and experimental groups showed similar, favourable progress in written discourse. The control group significantly increased by 54% (W[0] = 1.000, p < .001, d = 2.87, 95% CI = [-3.79, -1.96]), rising from 2.08 to 3.13. Similarly, the experimental group showed a slightly lower, yet significant, increase of 46% (W[0] = 1.000, p < .001, d = 1.41, 95% CI = [-2.13, -0.70]), rising from 2.18 to 3.05. These findings were supported by large effect sizes in both groups. Similar trends were observed across other *Writing* assessment descriptors, as indicated in Table 56, which provides additional information regarding the statistical analyses.

Table 56

Progress in Writing Performance at B2 Level for Each Group: A Comparative Analysis of Pre-Test vs. Post-Test

Writing performance	Groups	Pre-test	Post-test	diff.	*W* score	*p* value	Cohen's *d*	95% CI
Overall	Control	136.35	167.50	31.000	0	<.001	4.24	[-5.40, -3.09]
	Experimental	138.40	163.40	22.437	1	<.001	1.39	[-2.10, -0.67]
Content	Control	2.13	3.50	1.500	0	<.001	4.27	[-5.43, -3.11]
	Experimental	2.25	3.33	1.250	2.5	<.001	1.85	[-2.61, -1.08]
Comm. achv.	Control	2.15	3.40	1.250	0	<.001	3.69	[-4.75, -2.64]
	Experimental	2.20	3.18	1.000	0	<.001	1.68	[-2.42, -0.93]
Organisation	Control	2.33	3.58	1.250	0	<.001	3.73	[-4.79, -2.67]
	Experimental	2.68	3.38	1.000	0	.001	1.15	[-1.84, -0.46]
Language	Control	2.08	3.13	1.000	0	<.001	2.87	[-3.79, -1.96]
	Experimental	2.18	3.05	1.000	0	<.001	1.41	[-2.13, -0.70]

Furthermore, the inferential comparison of the evolution in *Writing* performance between the two groups revealed statistically significant differences (W[283] = 10.000, p = .025) in favour of the control group, albeit with a small effect size (d = 0.36, 95% CI = [-0.29, 1.00]). However, there were no statistically significant differences (p > .05) in terms of the "Language" descriptor between the evolution of both groups. It is worth noting that significant progress was identified in relation to the "Organisation" assessment descriptor, favouring the control group, as detailed in Table 57.

Table 57

Evolution of Writing Performance at B2 Level: A Comparison of Control and Experimental Groups

Writing performance	Variations (AI)		diff.	*W* score	*p* value	Cohen's *d*	95% CI
	Control	Experimental					
Overall	31.15	25.00	10.000	283	.025	0.36	[-0.29, 1.00]
Content	1.38	1.08	0.500	263	.077	0.50	[-0.15, 1.15]
Comm. achv.	1.25	0.98	0.500	269	.053	0.74	[0.08, 1.40]
Organisation	1.25	0.70	0.500	295	.008	0.96	[0.28, 1.64]
Language	1.05	0.88	0.000	233.50	.351	0.33	[-0.32, 0.97]

4.2.4. Concluding Remarks of RQ2

This section has examined the effect of metaphor-mediated instruction on the B2 achievement of L2 learners studying at the upper-intermediate level. The analysis has compared the English proficiency level of the control and experimental groups in the pre-testing and post-testing measures, as well as their evolution in terms of B2 overall, *Speaking* and *Writing* performance. The key findings of RQ2 are as follows:

1. In the **pre-test**:
    - No statistically significant differences were found between both groups in terms of their overall B2 achievement and *Speaking* and *Writing* performance. Although there were initial significant differences in specific areas but unrelated to vocabulary, these findings establish a strong baseline for comparing the English proficiency of the two groups, providing a reliable framework for further analysis.

2. In the **post-test**:
    - Although no statistically significant differences in overall B2 performance were observed between the control and experimental groups, it is worth noting that the control group exhibited a slightly higher performance.
    - The experimental group exhibited a statistically significantly higher global level of B2 proficiency in *Speaking* compared to the control group. Their performance in the "Grammar and Vocabulary" assessment descriptor was also statistically significantly better. These findings were further supported by large effect sizes.
    - The control group showed statistically significantly better mean results in *Writing* performance than the experimental group, accompanied by a medium effect size. However, when specifically examining the "Language" descriptor, no statistically significant differences were found between the control and experimental groups.
3. Comparison of control and experimental groups' **evolution**:
    - Both the control and experimental groups demonstrated a statistically significant increase in their overall B2 performance, with large and medium effect sizes, respectively. However, neither reached B2 level of English proficiency in their overall performance after the CL-oriented instructional intervention.
    - The experimental group exhibited statistically significant evolution in their *Speaking* proficiency, including statistically significant growth in the "Grammar and Vocabulary" assessment descriptor, supported by medium effect sizes in both cases. Findings highlight the impact of metaphor-mediated instruction on promoting *Speaking* skills, especially in relation to oral vocabulary production.
    - Both groups showed statistically significant evolution in *Writing* proficiency and the "Language" assessment descriptor, with large effect sizes observed in all cases. These findings emphasise the less pronounced impact of CL-style instruction on enhancing *Writing* skills, particularly regarding written vocabulary production.

- The comparison of the evolution of both groups concerning their overall and *Speaking* performance at B2 level did not show any statistically significant differences. However, the experimental group exhibited statistically significant progress in the *Speaking* "Grammar and Vocabulary" assessment descriptor compared to the control group, supported by a large effect size.
- The control group showed statistically significant progress in *Writing* performance compared to the experimental group, with a medium effect size. However, when it comes to written vocabulary production, no significant differences were observed in the evolution of both groups of participants.

4.3. RQ3: CORRELATION BETWEEN METAPHOR PERFORMANCE AND ASSESSMENT OF ENGLISH LANGUAGE PERFORMANCE AT B2 LEVEL

What relationship, if any, can be established between using metaphorical language in the L2 with English language proficiency at B2 level?

This section examines the potential relationship between the use of metaphorical language in the L2 and achievement at B2 level of participants from the control and experimental groups. The analysis primarily focuses on assessing the correlation between metaphor use (measured in metaphor density) and participants' proficiency in *Speaking* and *Writing*, as evaluated in the delayed post-test. While the study contributes to our understanding of the correlation between metaphor use and these productive areas, it does not encompass a comprehensive examination of the link between metaphorical language and overall B2 performance, considering global B2 level was not achieved by the participants as evaluated by the *B2 First for Schools Cambridge English* qualification at (see Section 4.2.2.1).

To address RQ3, this section is divided into two parts. The first subsection focuses on assessing the correlation between the density of metaphors used in the participants' post-test oral outputs and their proficiency level in *Speaking* at B2 level. Additionally, it examines the potential correlation between metaphor use and the assessment descriptor for vocabulary use. In the second subsection, the analysis shifts towards exploring the correlation between

the metaphor density measured in the post-test written essays and their performance in *Writing* at B2 level. This analysis also considers the assessment descriptor for vocabulary use.

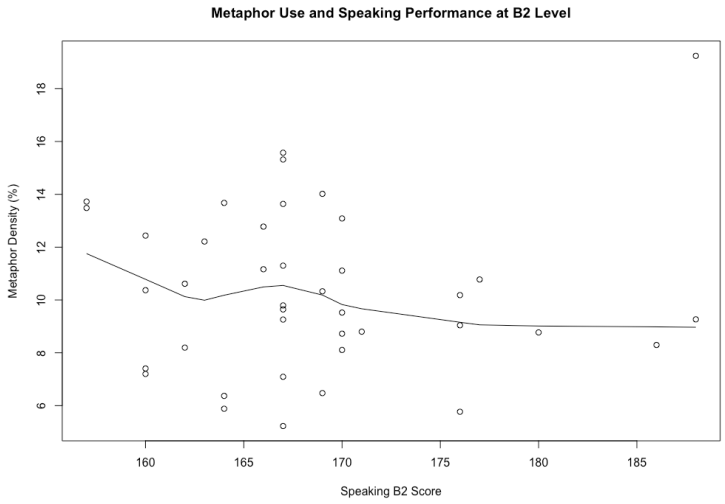
4.3.1. Metaphor Use and Assessment of *Speaking* Proficiency at B2 Level

This section examines the relationship between the density of metaphors used in the participants' post-test spoken productions and their *Speaking* proficiency at B2 level. The analysis aims to shed light on how metaphor use relates to *Speaking* proficiency, specifically in terms of vocabulary production in oral discourse.

The analysis revealed a non-linear, negative correlation between L2 learners' use of metaphors and their overall *Speaking* proficiency at the end of the study. Figure 1 visually represents this correlation, clearly illustrating that higher metaphor density was associated with lower *Speaking* proficiency scores.

Figure 1

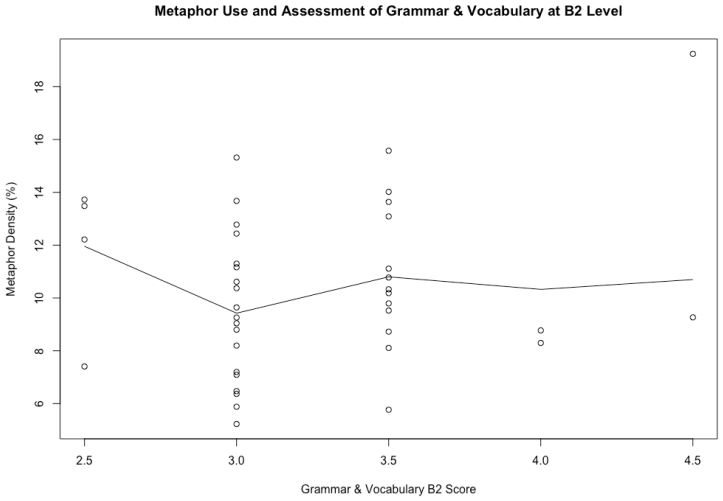
A Scatterplot with a Regression Line of the Correlation between L2 Learners' Metaphor Use and their Speaking Proficiency at B2 Level



Further analyses revealed a non-linear, positive correlation between the metaphors used by participants and their scores in the "Grammar and Vocabulary" assessment descriptor of the *Speaking* test, as shown in Figure 2. This relationship indicates that the more metaphors participants used, the better their performance was in terms of "Grammar and Vocabulary".

Figure 2

A Scatterplot with a Regression Line of the Correlation between L2 Learners' Metaphor Use and Assessment of "Grammar and Vocabulary" (Speaking)



However, it is important to note that the Spearman ρ test yielded weak and non-significant correlations between the density of metaphors in the participants' post-test spoken productions and their *Speaking* proficiency at B2 level (r = -.117, p > .05), as well as their "Grammar and Vocabulary" scores (r = .045, p > .05) at the end of the study.

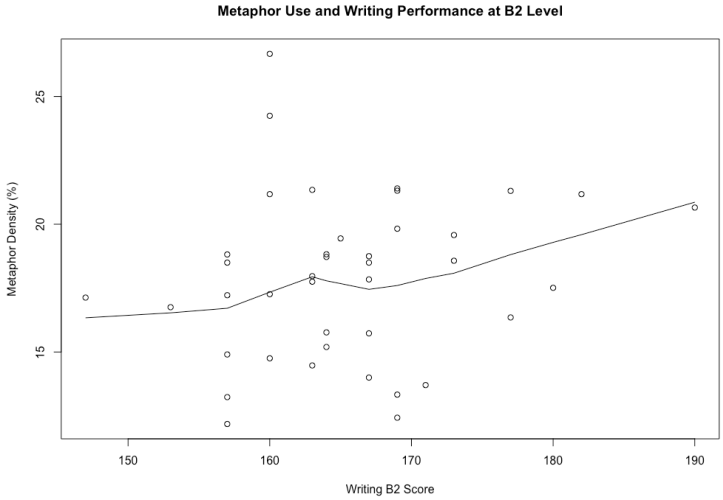
4.3.2. Metaphor Use and Assessment of *Writing* Proficiency at B2 Level

This section examines the relationship between the metaphor density in the participants' post-test written essays and their *Writing* proficiency at B2 level. The analysis aims to provide insights into how the use of metaphors relates to their proficiency in *Writing*, specifically in terms of vocabulary production in written discourse.

The analysis revealed a non-linear, positive correlation between L2 learners' metaphor use of metaphors and their post-test *Writing* performance, specifically Part 1 of the *Writing* test. Figure 3 illustrates this relationship, showing that as the metaphor density increased, so did the *Writing* scores.

Figure 3

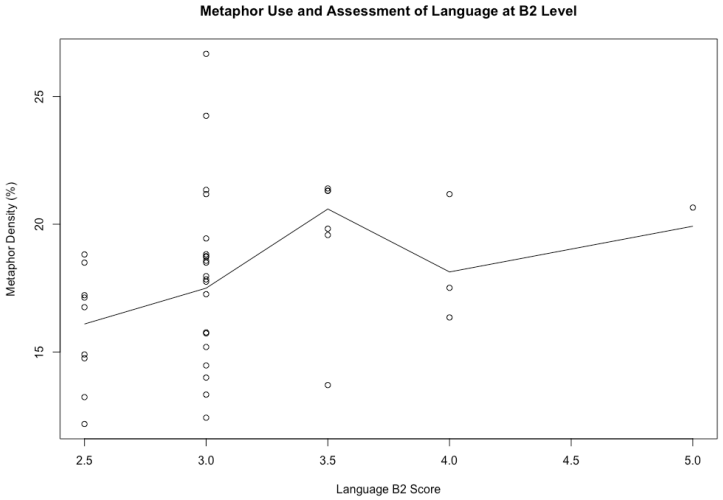
A Scatterplot with a Regression Line of the Correlation between L2 Learners' Metaphor Use and their Writing Proficiency at B2 Level



Furthermore, the analysis revealed a non-linear, positive correlation between participants' metaphor use in the post-test written essays and their scores in the "Language" assessment descriptor. As shown in Figure 4, this relationship indicates that the higher the metaphor density was, the higher the "Language" scores tended to be.

Figure 4

A Scatterplot with a Regression Line of the Correlation between L2 Learners' Metaphor Use and Assessment of "Language" (Writing)



The Spearman ρ test results showed a moderate and significant correlation (r = .414, p = .008) between the metaphor density observed in participants' written discourse and their "Language" scores at the end of the study. This indicates that metaphor use had a notable influence on participants' performance in terms of language usage in their written essays. However, the statistical test revealed a weak and non-significant correlation (r = .207, p > .05) between participants' metaphor use and their *Writing* scores obtained in the post-test essays.

4.3.3. Concluding Remarks of RQ3

This section has examined the relationship between metaphor use in the L2 and participants' CEFR B2 level of English proficiency in the post-test. The analysis has specifically focused on the correlation between metaphor density and the formal B2 assessment in English for both the control and experimental groups. The participants' *Speaking* and *Writing* proficiency at B2 level, as evaluated using the standard *B2 First for Schools* examination, were taken into consideration. The key findings can be summarised as follows:

1. L2 language **productive skills** (*Speaking* vs. *Writing*):
    - The analysis revealed weak, non-significant correlations between metaphor density and B2 proficiency scores. Despite the lack of statistical significance, the findings suggested a positive relationship between metaphor use and *Writing* proficiency, while a negative relationship was observed with *Speaking* proficiency.
2. **Vocabulary production** assessment descriptors (*Speaking* vs. *Writing*):
    - Positive correlations were observed between metaphor use and vocabulary production in both types of discourse at B2 level.
    - A significant, moderate correlation was found between the metaphors used in written discourse and the scores in the "Language" assessment descriptor. However, the correlation with the "Grammar and Vocabulary" assessment descriptor in oral discourse was weak and did not reach statistical significance.

CHAPTER 5

DISCUSSION

Chapter 5 discusses the research findings, aiming at responding to the three research questions posed in this study. It highlights some of the most important findings, specifically focusing on their implications for L2 metaphor production in oral and written discourse, rather than going into detail about form-function pairings.

The findings presented in this chapter can be considered meaningful across three major dimensions, each addressing a specific research question: L2 learners' metaphor performance (RQ1; see Section 5.1), L2 learners English language performance (RQ2; see Section 5.2), and the relationship between use of metaphor and the assessment of English language proficiency at B2 level (RQ3; see Section 5.3). Examples of MRWs used by L2 participants will be highlighted in bold throughout this chapter.

5.1. EXAMINING THE EFFECTS OF METAPHOR-MEDIATED INSTRUCTION ON METAPHOR PERFORMANCE AT B2 LEVEL (RQ1): KEY FINDINGS

This section analyses the results from examining the differences in metaphor use across oral and written discourse among the two groups of L2 learners who participated in the study: the experimental group (EXP) and the control group (CTRL).[41]

Results from the pre-test revealed no statistically significant differences in the use of organic metaphorical language (RQ1a) nor in the use of either open-class and closed-class metaphors (RQ1b) between the two groups of participants across oral and written discourse. This finding establishes a strong foundation for analysing and evaluating the effects of metaphor-mediated instruction on L2 learners' metaphor use and its potential implications for

[41] For easier understanding, participant references in Chapter 5 will be simplified. Each participant will be assigned a code using the following format: the abbreviation of "experimental" (EXP) or "control" (CTRL) and their original participant code.

learning outcomes at B2 level. Consequently, any differences observed in the post-test are likely attributable to the instructional treatment together with the learning progress taking place in the five months that went by, enabling a focused assessment of its potential benefits or limitations for learning gains at this level.

The forthcoming sections delve deeper into these findings. Section 5.1.1 explores the *amount* of metaphor use (RQ1a), measured in metaphor density, and examines the *type* of metaphorical language (RQ1b), distinguishing between open-class (i.e., nouns, verbs, adjectives, and adverbs) and closed-class metaphors (i.e., prepositions).[42] Section 5.1.2 then shifts the focus to assessing the long-term effectiveness of metaphor-mediated instruction on improving learning gains, particularly its impact on enhancing production of metaphorical language and deepening vocabulary knowledge.

5.1.1. Measuring Metaphor Use in Learner Discourse: Key Insights

L2 learners at the B2 level who received metaphor instruction exhibited higher metaphor usage in both oral and written discourse compared to those who did not. Specifically, L2 learners' metaphor usage rates were 11.07% (EXP) vs. 9.63% (CTRL) in oral discourse, and 19.31% (EXP) vs. 16.40% (CTRL) in written discourse. Similar to L1 English, L2 metaphor usage was more frequent in written than in oral discourse, supporting the idea that metaphors are more commonly used in written than in spoken language (Steen et al., 2010).

5.1.1.1. Metaphor Use in Learner Discourse: A Focus on Spoken Production

The study's findings highlight the beneficial effects of metaphor-mediated instruction in increasing L2 metaphor usage, especially within the Part 4 *Speaking* task of the *B2 First for Schools* examination. The results suggest that incorporating CL-inspired activities tailored to the EFL syllabus can enhance metaphor usage in scenarios that closely resemble oral real-life communication contexts, aligning to the CEFR descriptors.

[42] Examples of MRWs used by L2 participants will be highlighted in bold across across this chapter.

The current study provides compelling evidence of frequent L2 metaphor use in the oral discourse of Spanish learners of English at the B2 level. This finding, diverging from the traditional focus on written discourse, contributes to the existing body of research on learner English in spoken interactions within EFL contexts. It broadens our overall understanding of how metaphors are employed in real-world contexts by EFL learners (MacArthur & Alejo-González, 2024). The spoken production results of the experimental group in this study (11.07%) align closely with the findings from the EuroCoAT corpus, which reported a metaphor usage rate of 11.90% in office-hours consultations between L1 and L2 speakers (Alejo-González, 2022). These results are also comparable to the 14.22% metaphor usage rate found in the MetCLIL corpus, in another interactive oral genre involving L2 speakers in discussion seminars (Alejo-González, 2024). Nonetheless, it should be noted that variations in L2 spoken metaphor use may be influenced by specific communicative contexts (spoken productions in exam conditions vs. office-hour consultations and academic seminars), learning environments (EFL vs. ELF/EMI), topics, and learner populations (secondary school vs. university students). The similarity with the EuroCoAT findings is particularly relevant, as their study's face-to-face conversations provide an interactional context akin to the spoken discussions in the *Speaking* test of my study.

The control group's significant decrease in metaphor usage in their spoken productions, after not receiving metaphor-mediated instruction, suggests a potential negative impact from excluding this approach in L2 instruction. This indicates that the absence of explicit teaching of metaphor in TEFL might reduce metaphor usage, particularly open-class metaphors at the upper-intermediate level.

The choice of topics could partly explain the lack of significant differences in metaphor usage evolution between the two groups, despite the varying patterns. As detailed in Chapter 3, the study investigated the use of metaphorical language in L2 learners' speech through a variety of broad and relevant topics, such as *helping children with homework* (pre-test) and *spending time outdoors* (post-test), to ensure topic consistency. Nonetheless, this selection of topics may have negatively affected the occurrence and frequency of metaphors in L2

learners' spoken outputs. This observation aligns with existing research that suggests a topic-dependence in the use of metaphor in discourse (Deignan et al., 2013; Golden, 2012, 2021; Semino, 2008).

While employing different topics in each testing phase might introduce variability in metaphor density, the "ecological validity" of the findings is enhanced using authentic exam papers (Boers, 2011, p. 237). These not only mirror real-world language use in standard L2 assessment contexts but also accurately reflect actual pedagogical practices for both groups.

The focus on naturally produced learner discourse within testing environments ensured alignment with the EFL syllabus and CEFR descriptors, representing real-world language use. By prompting L2 learners to produce outputs spontaneously in a controlled setting, the study was able to analyse metaphor usage in response to specific tasks rather than general text production. This approach offers a more cohesive alternative to ad-hoc tests, which may deviate from standard teaching practices or unintentionally favor certain participant groups (Boers, 2013). In my view, this methodology might help bridge the gap between CL-oriented teaching approaches and the assessment techniques used to measure learning gains resulting from ACL research. However, it is crucial to acknowledge the limitations of data naturally produced in regular classroom activities, as it may not fully represent the breadth of L2 learners' metaphor production abilities.

The personal significance of topics may have also influenced the use of metaphors in L2 discourse leading to individual differences among participants in this study. Prior research suggests that discussing personally significant topics prompts an increased use of metaphors, especially in written texts (Golden, 2021; Littlemore et al., 2014). Expanding on this, the current study shows that this also influences L2 oral discourse, including topics that are not highly abstract and often leading to a potential decline in metaphor use.

As Boers (2004) points out, it is important to exercise caution when considering a group of L2 learners as a uniform entity. Enhancing metaphor awareness may not yield uniform effectiveness across all individuals, given the influence of diverse affective and cognitive factors. The learning experience itself, for instance, can have a crucial impact on metaphor

production. Motivated learners, for example, might be more inclined to engage with metaphors in their discourse. Additionally, aspects specific to the teaching intervention, such as patterns of class attendance, completion of assigned activities, or active cognitive engagement during classroom practice, could have influenced the development of metaphor use leading to individual differences.

For instance, a participant from the experimental group (EXP_ST0105) showed a keen interest in conversations about *spending time outdoors*, as indicated by a metaphor usage rate of 19.57%. The Shapiro-Wilk test confirmed the normal distribution of metaphor density in the post-test oral outputs for the experimental group ($p > .05$), suggesting that her high metaphor usage was a specific preference in response to a particular question in the discussion task. This particular interest is highlighted in bold in example (104), where she talks about the time she and her friends spend on outdoor activities, contextualising her answer within the COVID-19 pandemic lockdown.

(104) EXP_ST0105: <POSTB2.ST0105> […] just we **love**{*indirect-MRW*} to go to the cinema we watch lots of **things**{*indirect-MRW*} we **love**{*indirect-MRW*} to go to the park just to **relax**{*indirect-MRW*} and talk **about**{*indirect-MRW*} the week and all our **ups**{*indirect-MRW*} and **downs**{ *indirect-MRW*} during the week **with**{*indirect-MRW*} exams or different **situations** *indirect-MRW*} but yeah I think that it's very important to be outdoors with friends (.) yeah I now that we are like{*MFlag*} **isolated**{*direct-MRW*} I **realised**{*indirect-MRW*} that is very important to be with friends and outdoors now the **things**{*indirect-MRW*} that I **really**{*indirect-MRW*} **miss**{*indirect-MRW*} are my friends and I mean outdoors so that's it it's like{*MFlag*} the **key**{*direct-MRW*} for us when we are **in**{*indirect-MRW*} like{*MFlag*} **down**{*direct-MRW*} **ages**{*direct-MRW*} because are the very important **things**{*indirect-MRW*} for us and also being **out**{*indirect-MRW*} with friends usually **puts**{*indirect-MRW*} **up**{*indirect-MRW*} our **spirits**{*indirect-MRW*} because we sometimes **feel**{*indirect-MRW*} **low**{*indirect-MRW*} or **things**{*indirect-MRW*} like that and they always are **there**{*indirect-MRW*} to help you so (.) </POSTB2.ST0105>

Example (104) illustrates the varied responses of learners to the CL-oriented instructional intervention. Participant EXP_ST0105 demonstrated a significant increase in metaphor density. Her metaphor usage evolved by 114% in spoken discussions, rising from 8.98% in the pre-test to 19.24% in the post-test. This evolution was significantly more pronounced (p = .039) than the 6% increase observed in other participants of the same group. This suggests that EXP_ST0105 may have been particularly receptive to the instructional methods employed in the study, as evidenced by her significant increase in metaphor usage post-intervention.

Example (104) also illustrates how heightened metaphor awareness may facilitate conscious control over metaphorical expression, as seen with the use of "extended metaphor" stemming from their apparent awareness of the CM HAPPINESS IS UP, e.g., *ups and downs*, *feel low*, and so on (MacArthur, 2016a). Although not typical of the experimental group as a whole, this participant's use of metaphor confirms the impact of metaphor-mediated instruction on individual learners. What is clear is that for certain learners metaphor has proved an attractive and stimulating aspect of the learning experience. However, as Low (2017) cautions, "one cannot assume that the same person will create consistent, or at least convergent, metaphors for similar topics" (p. 253). Additional analysis is necessary to confirm these observations on facilitating topic development and generalise them, especially within the realm of L2 spoken production.

Factors related to B2 training and the input from the L2 classroom may partially account for metaphor usage among participants. The goal-oriented approach of B2 training, with textbooks primarily supporting examination preparation, might have played a role in the acquisition of new lexical items relevant to textbook topics. For instance, take the MRW verb *spend*[43] in (105) and (106).

(105) CTRL_ST0408: <*{Do you and your friends <u>spend</u> a lot of your time doing outdoor activities?}*> <POSTB2.ST0408> yeah we **spend** a lot of time we meet all the weeks

[43] Vocabulary term that participants who were preparing for B2 level in this study learnt in Unit 9 "Secrets of the mind" (Brook-Hart, 2014, pp. 96–105).

and we change the activities each week so we do different things (.) </POSTB2.ST0408>

(106) EXP_ST0102: <*{Do you think it's a good idea for families to go on holidays where they spend a lot of time outside, for example camping holidays?}*> <POSTB2.ST0102> if if this family is a family that **spend** a lot of time in outdoors I it's a good idea to to do camping or for example if a family only wants to go to to the beach on holidays it's a good idea to they go to the beach because they like it a lot and it's like if you if it's like a routine that you **spend** in another part of the world […] </POSTB2.ST0102>

The increased use of metaphors by L2 learners can also be explained by their use of highly frequent lexical items that are metaphorically used in those specific contexts. For example, consider the MRW verbs *have* in (107) and *take* in (108).

(107) CTRL_ST0307: <*{Is growing up in the countryside better for children than living in cities?}*> <POSTB2.ST0307> […] it's a good way to live in a countryside for example to **take** care of the environment and to **take** a rest what do you think? </POSTB2.ST0307>

(108) EXP_ST0103: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0103> friends are very important in our lives so now that we don't **have** them we miss them obviously […] </POSTB2.ST0103>

The positive progress of the experimental group in their metaphor usage can be attributed not only to the instructional approach but also to their incorporation of metaphorical expressions normally taught during the B2 training. Furthermore, it is important to bear in mind that many polysemous words are more frequently used in their figurative senses than in their literal ones (Hoffman, 1983).

This study demonstrates that the impact of specific task conventions on enhancing L2 metaphor discourse, previously observed in written production (e.g., Hoang & Boers, 2018; Littlemore et al., 2014; Nacey, 2013, 2022), extends to spoken production as well. The findings provide insights into L2 oral discourse, particularly in testing environments. In the *Speaking* test, the Part 4 task requires participants to engage deeply in topic-based discussions. This

format mirrors the Part 1 task of the *B2 First for Schools Writing* test, where L2 learners are asked to express and justify their opinions on given topics, often leading to increased use of metaphors in written outputs (Littlemore et al., 2014; Nacey, 2013, 2022).

The demand of specific vocabulary or fixed expressions containing MRWs, often required in the Part 4 *Speaking* task, may influence the metaphor usage of L2 learners. These MRWs are typically employed by learners to effectively complete the task, impacting their choice of language and metaphorical expressions. For instance, the opinion-giving phrase "from my point of view" in (109) and (110) featured the MRW nouns *point* and *view*.

(109) CTRL_ST0305: *<{Should students have to do outdoor sports lessons at school?}>* <POSTB2.ST0305> *from my* **point** *of* **view** I think that yes because outdoors you can do a lot of activities like running cycling that you can't do indoors what do you think?> </POSTB2.ST0305>

(110) EXP_ST0202: *<{Is growing up in the countryside better for children than living in cities?}>* <POSTB2.ST0202> […] *from my* **point** *of* **view** I think that it's nice to grow up in the countryside because you have a lot of things to do you have more space to do activities and what you want […] </POSTB2.ST0202>

These examples indicate that metaphor density in both groups' spoken outputs do not necessarily signify any awareness of metaphors. The use of pre-established language chunks, exemplified by phrase such as *from my point of view* in examples (109) and (110) might improve the speed and naturalness of speech delivery.

Spoken discussions in these settings often occur in pairs involving immediate interaction and collaboration. This format, influenced by the unique characteristics of face-to-face conversation, is likely to impact learners' oral performance and, consequently, their use of metaphors, as observed by MacArthur and Littlemore (2011). Speakers' spoken productions depend on their conversational partner, both in terms of time and of topic development. This makes spoken productions quite different from written essays. This pattern could partly explain the less frequent occurrence of metaphors in oral discourse compared to

written discourse observed in the current study, especially in standard L2 competence assessment contexts.

The use of metaphorical language in the oral productions of L2 learners in this study might have been influenced by factors such as the need for further practice and exposure to consolidate the metaphor usage in spoken productions. One key difference to consider is that written discourse is typically more planned and allows for reflection and editing, as seen in the teaching treatment. This treatment included CL-inspired activities aimed at enhancing both oral and written production skills, yet the execution of assessment tasks differed significantly. Written production tasks were assigned as follow-up homework, providing learners with ample time to plan and refine their language, a luxury not afforded in spoken discourse. In contrast, oral production tasks in the L2 classroom were completed without prior preparation, reflecting the spontaneous and less structured nature of oral communication. The need to think on-the-spot might constrain the use of more complex linguistic structures such as open-class metaphors, compared to written discourse where there is an opportunity to more careful crafting of sentences.

It should be also acknowledged, however, that in semi-structured yet spontaneous conversations such as those occurring in the *Speaking* task, collaborative efforts can lead to an increased use of metaphorical language. The face-to-face nature of these interactions often encourages the repetition or mirroring of MRWs, with participants echoing each other's language choices (MacArthur, 2016a; MacArthur & Littlemore, 2011). This pattern might be more pronounced in high-stakes ESOL examinations, where the pressure to excel can amplify these dynamics. An illustrative example of this phenomenon is observed in the discussion about whether a rural or urban upbringing is more beneficial for children. In example (111), the MRW noun *key* is strategically used to construct arguments, demonstrating the impact of dialogic interactions on the use of metaphors when one speaker repeats another's metaphorical use of *key*.

(111) EXP_ST0201: <POSTB2.ST0201> yes maybe in the city you have more abilities to to study than in the countryside but I think *as she said* the **key** is the family if you are with

your family I think you are going to have the same interest or or the things that you are going to think in your life (.) </POSTB2.ST0201>

The potential occurrence of repetitions was also attributed to participants often incorporating metaphorical language present in the speaking stimuli, explicitly prompted to guide their answers in the *Speaking* tests. Consider the MRW verbs *grow* and *live* in (112) and the MRW *verb* in (113).

(112) CTRL_ST0401: *<{Is growing up in the countryside better for children than living in cities?}>* <POSTB2.ST0401> I think that there are both advantages and disadvantages I think that if you **grow** up in the countryside you connects more with the nature whereas in the whereas if you are **growing** up in a **living** in the in the city you can make more friends or you can be you can have more relationship with the people you can be more social what about you? </POSTB2.ST0401>

(113) EXP_ST0101: *<{Do you think it's a good idea for families to go on holidays where they spend a lot of time outside, for example camping holidays?}>* <POSTB2.ST0101> […] I consider that it's a good idea now also for children that are actually currently they are always in front of a screen so it could be a good idea to disconnect for the technologies and **go** camping (.) </POSTB2.ST0101>

The demands of the spoken task and the versatility of prepositions may explain why MRW prepositions continued to be a dominant tool for expressing connections in oral discourse, despite a significant decrease in their use at the end of B2 training. Prepositions, given their high frequency in comparison with other closed-class items, might be more accesible for L2 learners, especially in spontaneous oral discourse. For instance, a single preposition such as *in* (defined by 26 entries in LM) can convey various metaphorical meanings, enhancing its utility in diverse contexts. To illustrate, consider ST0103 from the experimental group, as shown in (114).

(114) EXP_ST0103:*<{Do you and your friends spend a lot of your time doing outdoor activities?}>* <POSTB2.ST0103> […] friends are very important **in** *our lives* so now that

> we don't have them we miss them obviously and there are some things that for example if you are as you said you're not good you are sad so you have to talk with someone and that could be really important but it's better it's better to do it **in** *person* and outdoors instead of at home (.) </POSTB2.ST0103>

By grasping a single preposition such as *in* ("used with the name of a container, place, or area to say where someone or something is" [LM1]), L2 learners can convey a wide range of metaphorical meanings, such as experiential and temporal dimensions ("during a period of time" [LM5]) in "in our lives" or a manner of interaction or presence ("used to say how something is done or happens" [LM3]) in "in person". This breadth of applicability might explain the continued reliance on such prepositions, even as learners refine other linguistic aspects. This tendency highlights the fundamental role of MRW prepositions in oral communication throughout the development of L2 learners' proficiency.

Recent research by Alejo-González (2024) on the EuroCoAT and MetCLIL corpora demonstrates that in spoken academic contexts, MRW prepositions are the most frequently used lemmas among students, highlighting their crucial role in facilitating academic oral interactions. Alejo-González's (2024) analysis of metaphor density from in corpora reveals a high degree of overlap in the use of MRW prepositions such as *in*, *about*, *on*, *with*, *from*. In my study, the experimental group (*with*, *in*, *on*, *about*, and *from)* and the control group (*with*, *about*, *from*, *in*, *on)* predominantly used the same MRW prepositions, as detailed in Appendix P. This finding reinforces the idea of a foundational set of prepositions in oral communication.

The findings of this study suggest that excluding explicit metaphor instruction from the EFL syllabus could limit L2 learners' ability to incorporate metaphors in their speech. This is supported by the higher metaphor usage observed in participants from the experimental group, who engaged in spontaneous discourse, either unplanned or semi-planned, unlike their counterparts in the control group. These findings might prove useful in refining teaching methodologies, with the potential to enhance language proficiency and boost more expressive oral communication skills in L2 learners.

5.1.1.2. Metaphor Use in Learner Discourse: A Focus on Written Production

The study emphasises the relative effectiveness of metaphor-mediated instruction in helping secondary-school EFL learners use metaphors in their writing, particularly in the Part 1 *Writing* task of the *B2 First for Schools* examination.

The experimental group's significant evolution in using metaphors, compared to the control group, highlights the substantial impact of CL-inspired distributed learning of metaphor in enhancing metaphor usage in real-life written communication contexts at the B2 level. However, there is a lack of significant differences in the evolution of metaphor density between the control and experimental groups in their written essays. Despite not receiving metaphor-mediated instruction, the control group increase their use of metaphors in written discourse, although the difference between the groups was not statistically significant. This suggests that the development of the use of metaphor is relatively independent of any instructional method.

This challenges the perceived effectiveness of metaphor-mediated instruction, indicating that simply identifying metaphors in learner-produced texts and relying solely on metaphor density may not fully capture the extent of enhanced metaphor production. These observations underscore the necessity for a more comprehensive analysis of various factors that could influence L2 leaners' development of metaphor usage, particularly for those preparing for the B2 level without the benefit of metaphor-mediated instruction.

Although written discourse allows for reflection, it did not greatly favour open-class metaphors. Instead, L2 learners seemed to heavily lean on MRW prepositions in their writing, possibly due to their highly polysemous nature or language from the essay prompt itself to convey complex ideas (see Nacey, 2013). This reliance could stem from the inherent cognitive demands of writing, which challenges L2 learners to conceptualise, structure, and articulate concurrently while coping with exam pressure. Examples (115) and (116) demonstrate frequent usage of phrases such as "in the future", possibly drawn from the writing stimulus.

(115) CTRL_ST0302: Some might argue that schools are going to change **in** *the future*. This is because many things that there are now are in one form that **in** *the future* are going to be better because of all new facilities. (POSTB2.WOUTPUT.W1.ST0302)

(116) EXP_ST0205: The aim of this essay is to talk about how different schools will be **in** *the future*. It is a well-known fact that schools are going to be very different **in** *the future* because they will have more technology. (POSTB2.WOUTPUT.W1.ST0205)

The inherent nature of written discourse, typically devoid of immediate feedback, likely prompts L2 learners to rely more heavily on familiar linguistic patterns. This tendency might result in an amplified use of specific discourse elements. During their B2 training, all participants were exposed to intricate textual examples from the topic-based textbook that were rich in closed-class metaphors. In my study analysing the impact of input on metaphor use in learner discourse (Martín-Gilete, 2022b), I observed that these L2 learners were introduced to input texts with a notable higher metaphor density of closed-class metaphors (47.27%) as compared to open-class metaphors (22.71%). This is illustrated in the examples (117) and (118), which show how such input, encompassing both oral and written forms, might have influenced learners to replicate similar levels of metaphoricity in their writing.

(117) Psychology's quite a young science, which means psychologists haven't had time to measure how or whether our minds are evolving, or whether **in** *fact* we're getting cleverer. (TB9.OINPUT.L1)

(118) Well, now Lufthansa, the German airline, has made the calculation that on 80% of its flights, there is **in** *fact* a doctor amongst the passengers. (TB11.WINPUT.R3)

In the current study, this trend is particularly evident among L2 learners from the experimental group, who employ closed-class metaphors in their written essays more frequently after receiving metaphor-mediated instruction. This suggests a direct influence of the input material on their own metaphor usage, potentially deepening their understanding of preposition usage.

The observed use of metaphors among all participants might be also linked to the demands of the Part 1 *Writing* task in the *B2 First for Schools* examination. This task, akin to the Part 4 *Speaking* task, requires that L2 learners articulate their opinions and highlight their personal significance, which often leads to an increased use of metaphors (Littlemore et al.,

2014). This is particularly evident in argumentative essays, where abstract topics typically encourage more metaphorical language (see also Nacey, 2013, 2020, 2022).

These observations might explain the high use of open-class metaphor in this study, aligning with the increased usage observed by Littlemore et al. (2014) in B2 level *Writing* tasks. The experimental group, which received metaphor-mediated instruction, demonstrated a statistically significant higher usage of open-class metaphors than the control group. This was not only a positive enhancement but also indicated a balanced growth in their use of open-class metaphors. Thus, these learners appear to be better equipped for qualitative changes in their *Writing* skills at B2 level, likely due to the explicit focus on metaphors in their L2 instruction.

Despite these findings, it should be noted that the increased usage of metaphors might also be attributed to L2 learners' use of highly frequent lexical items. For example, among the top five MRW verbs use by the experimental group, we find *make* (see Appendix R). Example (119) illustrates how *make* was metaphorically used in this specific context ("used with some nouns to say that someone does something" [MM2]).

(119) EXP_ST0201: Therefore, from my point of view the key to change this are the books. We have studied in the same way all our life and we have to start to **make** changes. (POSTB2.WOUTPUT.W1.ST0201)

The MRW verb *have* is also a highly frequent lexical item in the writing of both groups of participants, as illustrated in (120) and (121). The verb *have* is defined by 24 sense entries in MM, ranging in meaning from "used for forming perfect tenses" (MM1) to "have sex with somebody" (MM24). Its broad range of metaphorical meanings, such as "used for saying what the qualities of someone's character are" (MM2b) or "to make something happen" (MM13), allows it to be applied in various contexts, making it a convenient choice for L2 learners when expressing ideas and concepts in English. The familiarity and versatility of the verb *have* likely contributed to its frequent incorporation in the participant's writing, enhancing the depth and nuance of their written expression.

(120) CTRL_ST0302: Sometimes the students **have** many homework to do and they do not **have** time because they **have** extra activities. (POSTB2.WOUTPUT.W1.ST0302)

(121) EXP_ST0201: At first sight, it's clearly seen that most of the thing in school will be different, for instance, the students will **have** other subjects. The subjects that they **have** now would be taught more practically. (POSTB2.WOUTPUT.W1.ST0201)

The conventions of B2 essay writing encourage the use of metaphors, as they require L2 learners to articulate their thoughts in sophisticated ways characteristic of written discourse. This sophistication often entails complex structures and vocabulary, including the use of specific phrases or fixed expressions that contain MRWs. For instance, opinion-expressing phrases containing MRWs are commonly employed by B2 level students to express their viewpoints in essay writing. Consider the phrases "share the view" in (122) and "in my view" in (123), which include the MRW noun *view*.

(122) CTRL_ST0406: Nowadays most people **share**{*MRW*} **the**{*non-MRW*} **view**{*MRW*} that schools are going to have an incredible evolution. (POSTB2.WOUTPUT.W1.ST0406)

(123) EXP_ST0209: All in all, *in my* **view**, schools will be very different in the future. (POSTB2.WOUTPUT.W1.ST0209)

As discussed earlier regarding spoken productions, another potential reason for the frequent use of metaphor could be learners' motivation to showcase their L2 proficiency using more sophisticated language. For instance, opting for phrases such as *share the view* in (109), as opposed to a simpler "think", demonstrates this tendency. This phenomenon aligns with findings from Nacey's (2010, 2013) and Littlemore et al.'s (2014) research, where learners were observed making concerted efforts to excel by employing more complex language.

Task familiarity is a key factor in language production, as highlighted by research such as that of Aas and Nacey (2019) and Nacey (2022). In this study, it seems likely that L2 learners' familiarity with the *Writing* task of the B2 exam, developed through intensive training, enhanced their ease and confidence in using pre-fabricated metaphorical expressions. These multiword units, explicitly taught and encouraged in the classroom for use in written, planned

discourse, likely became more intuitive for learners to use in their essays. Such training may have made these expressions readily accessible and integrated into their written responses.

The MRW verb *take* appears to be frequently used by both groups of participants due to its presence in specific or fixed expressions containing MRWs that are typically employed in essay writing at B2 level to highlight importance, as illustrated in (124) and (125). In these examples, the learners employed the phrases “take into account” and “taking all the above into consideration”, both of which contain the verb *take* used metaphorically.

(124) CTRL_ST0307: **Taking** *all the above into consideration*, I strongly believe that schools will be very different in the future because of the new knowledge and the big help of the Internet. (POSTB2.WOUTPUT.W1.ST0307)

(125) EXP_ST0208: Moreover, another point to **take** *into account* is the way in which new subjects will be introduced. (POSTB2.WOUTPUT.W1.ST0208)

These fixed expressions provide learners with ready-made language chunks that can be readily used in their writing. The familiarity and conventionality of such phrases might have encouraged learners to incorporate them into their essays, effectively integrating metaphorical language into their repertoire and allowing all the learners to express ideas more efficiently, regardless of any specific attention to metaphor.

Verbs such as *mention*, *share*, *consider*, *go*, and *say* were similarly commonly used by all participants due to their presence in fixed expressions frequently employed by L2 learners to strengthen arguments in the *Writing* task. Examples (125) through (129) exemplify these factors contributing to the participants’ MRW choices in their writing.

(126) CTRL_ST0305: Last but not least, tablets and computers should be **mentioned** because is a good way of teaching students instead of using books. (POSTB2.WOUTPUT.W1.ST0305)

(127) CTRL_ST0407: Nowadays, most people **share** *the view* that schools will have more resources in the future. (POSTB2.WOUTPUT.W1.ST0407)

(128) CTRL_ST0304: Secondly, subjects is also *an important point to* **consider** because high schools will add several subjects about mental illnesses and how can they take care of themselves. (POSTB2.WOUTPUT.W1.ST0304)

(129) EXP_ST0204: Last but not least, *it* **goes** *without* **saying** that classes have to be more practical because it is a better way to study something and easy to remember. (POSTB2.WOUTPUT.W1.ST0204)

The fact that both the control and experimental groups heavily relied on these specific MRW verbs does not necessarily indicate awareness of metaphors. Rather, it highlights the influence of set expressions and highly frequent lexical items in shaping the participants' MRW verb choices in their written discourse. These ready-made expressions and highly frequent verbs seem to provide learners with a convenient and familiar way to convey their ideas. Similarly, it also shows that they are familiar with the form and use of this conventional metaphorical expressions in English. These fixed expressions and versatile verbs allow them to adapt their language to different meaning and functions, leading to their frequent incorporation in written discourse.

A notable difference was observed between the control and experimental groups regarding participants' frequent use of MRW adjectives in their written discourse. This can be largely attributed to the use of specific vocabulary or fixed expressions containing MRWs, commonly employed in essay writing to introduce a key point or give emphasis. The control group primarily used the MRW *clear* in the phrase "it is clear" (see [130]), while the experimental group exhibited a more diverse usage by employing the MRW adjectives *clear* and *safe* in the phrases "it is clear" (see [131]) and "it is safe to say" (see [132]), respectively.

(130) CTRL_ST0407: To sum up, *it is* **clear** that teacher, subjects and construction are important facts to improve schools. (POSTB2.WOUTPUT.W1.ST0407)

(131) EXP_ST0108: Secondly, *it is* **clear** that students have now quite a lot of subjects to study. (POSTB2.WOUTPUT.W1.ST0108)

(132) EXP_ST0107: To begin with, *it is* **safe** *to say* that teachers will probably change the way they work. (POSTB2.WOUTPUT.W1.ST0107)

The employment of frequently used MRW adverbs was further shaped by participants' use of specific vocabulary or fixed expressions containing MRWs commonly used in B2 level written essays to introduce opinion. For instance, consider the MRW adverb *generally* in the phrases "generally speaking" and in "it is generally accepted", as depicted in examples (133) and (134), respectively.

(133) CTRL_ST0403: *It is* **generally** *accepted* that schools in the future will be very different because of the new technology and researchs done. (POSTB2.WOUTPUT.W1.ST0403)

(134) EXP_ST0108: **Generally** *speaking*, it is thought that in the near future the world will be very different, and schools are not an exception. (POSTB2.WOUTPUT.W1.ST0108)

A closer look at the results revealed a common tendency among both groups of participants to employ the MRW adverb *before*. This preference seems to be also influenced by its role within fixed expressions containing MRWs, which are frequently used in essays to refer to a previous statement or emphasise the relevance of a particular point, reinforcing its significance, as exemplified in (135) and (136), respectively.

(135) CTRL_ST0406: Regarding everything that *I have mentioned* **before**, I realise that all this could be happening in a few years. (POSTB2.WOUTPUT.W1.ST0406)

(136) EXP_ST0103: To conclude, it is clear that everything will change but, *as I said* **before**, we can not depend on technology. (POSTB2.WOUTPUT.W1.ST0103)

L2 learners also frequently employed phrasal verbs containing *up* in fixed expressions that helped them conclude their essays, as illustrated in (137) and (138).

(137) CTRL_ST0309: To **sum up**, it is clear that teachers, subjects and new technologies are the most important points for the schools change. (POSTB2.WOUTPUT.W1.ST0309)

(138) EXP_ST0102: To **sum up**, to my way of thinking, it's inevitable that the world will change. We, as humans, need to adapt to these new customs and take it easy. (POSTB2.WOUTPUT.W1.ST0102)

These observations provide some insights into the distinct aspects and challenges of metaphor usage among L2 learners, especially in written testing environments. In particular, the unique demands of the *Writing* task might have spurred a heightened use of specific vocabulary or fixed expressions rich in MRWs. The learners' frequent use of metaphorical language within specific or fixed expressions suggests that they often rely on pre-established language chunks to enhance their written expression.

While the frequent use of pre-fabricated metaphorical language may enhance learners' fluency and ease of expression, it may not necessarily indicate metaphor awareness or the ability to employ metaphors in context. These observations align with Low's (2017) question about "whether a metaphor was somehow pre-constructed or whether it was put together "on the fly", maybe as an artifact of the elicitation procedure" (p. 252). This raises questions about the nature and spontaneity of metaphor usage in L2 learning contexts. As such, it is crucial to look beyond the mere frequency of metaphor use and examine how L2 learners use them in various contexts and their ability to complete the exam tasks using metaphorical language.

5.1.2. Understanding Metaphor-Mediated Instruction: Benefits and Limitations

The study reveals a significantly higher occurrence of metaphor in written essays compared to the interactive spoken discussions of all participants studying at B2 level, both before and after the CL-oriented instructional intervention. This suggests that regardless of the teaching treatment focused on explicit metaphor instruction, L2 learners consistently used metaphors more frequently in their compositions than in their conversations. This observation is crucial in understanding how metaphor-mediated instruction affects metaphor usage across different communication modes (written vs. spoken), especially for L2 learners at the B2 level who all are engaged in a communicative approach.

The variations in metaphor production observed between oral and written discourse can likely be traced to the distinct characteristics and demands of spoken vs. written tasks. L2 learners may require more time and practice to incorporate metaphors into oral discourse, owing to its spontaneous and real-time nature, while written discourse provides a platform for

more deliberate and structured expression. This contrast in communication modes may have likely contributed to the differing levels of metaphor usage. Consequently, these differences might significantly impact metaphor use in *Speaking* and *Writing* tasks, particularly in high-stakes ESOL examinations where L2 proficiency is critically assessed under the unique pressures of real-time communication.

Despite the potential influence of external factors, findings suggest that metaphor-mediated instruction seems to have an important role in enhancing both spoken and written production. Learners who received explicit metaphor instruction and were aware of its role in language use showed increase usage of metaphors in their oral and written discourse, as evidenced in the delayed post-test. This metaphor-mediated instruction led to consistent improvement in the use of metaphors across both types of discourse, indicating a more evenly distributed use of metaphorical language in speech and writing.

L2 learners who did not receive CL-oriented distributed learning of metaphor showed inconsistent progress in using metaphors, particularly open-class words. While it seems that they can naturally enhance their metaphor usage in written discourse, which often involves more planning and pre-fabricated metaphorical expressions, they encounter significant challenges with metaphors in their spoken language. Oral discourse, often less planned or semi-planned and more spontaneous, presents notable obstacles for them without explicit metaphor instruction.

A five-month CL-oriented instructional period has been identified as sufficient for aiding L2 learners in increasing their metaphor usage, an important finding given the challenges of conducting long-term classroom-based research (Low, 2017). These findings underscore the value of sustained efforts in fostering metaphor awareness in L2 classroom practices, supporting the recommendations by Boers (2013) and MacArthur (2010). These findings also highlight the importance of contextualising different CL-oriented pedagogical techniques within the overall instructional framework, ensuring that they are integrated into regular activities aligning with the official EFL syllabus in real-life classroom environment.

The increased use of metaphors by the experimental group in this study, as compared to the control group, cannot be solely attributed to the metaphorical language taught during the CL-oriented instructional intervention. This difference may also be due to the experimental group's exposure to English metaphors in other contexts, such as in their secondary schools or through English discourse outside the English language school.

The six-week interval between the instructional intervention and the delayed post-test is key in highlighting the effectiveness of metaphor-mediated instruction, particularly in offering ongoing support in using metaphors at B2 level (Littlemore et al., 2014). The long-term improvement in the experimental group's use of metaphor, evident in both oral and written discourse, is an important outcome. The statistically significant changes and large effect sizes observed in the delayed-post-test underscore the practical relevance of this instructional approach in real-world L2 classrooms (Low, 2008). However, caution is warranted when considering a group of learners as a homogeneous population. The effectiveness of the technique of metaphor awareness may not be equally effective for all individuals due to diverse affective and cognitive factors (Boers, 2004; Littlemore, 2023).

Metaphor-mediated instruction appears to have increased L2 learners' confidence in using organic metaphors, leading to more frequent use in speech and writing and deepening their vocabulary knowledge (MacArthur, 2010). However, a five-month period may be insufficient for L2 learners to fully understand and consistently use metaphors according to conventional standards (Littlemore, 2023). The study's findings indicate that increased awareness of metaphors does not necessarily translate into better retention of form and/or meaning and effective usage. This suggests that CL-inspired instruction does not provide L2 learners with an infallible method for deciphering metaphorical meanings and applications autonomously (Boers, 2011).

The study findings highlight a contrast between CMs and MRWs. The learners might use MRWs such as *generally* in "generally speaking" as pre-fabricated expressions, meaning they use them as fixed phrases without fully grasping the nuanced meanings these words can convey. Findings also demonstrate the challenges of reconciling these uses of MRWs with

CL-inspired instruction. CL focuses on understanding language in the context of how the human mind processes and conceptualises it. The rigid or formulaic use of MRWs by L2 learners might not align well with the more fluid and context-dependent approach advocated by CL, indicating a gap in L2 teaching methodologies. These observations and their potential impact on L2 teaching strategies will be further examined in the subsequent discussion.

The metaphorical language used by the experimental group suggests a departure from surface-level understanding to a more profound grasp of the concepts. This is mirrored in L2 learners' discernment of when and how to employ these MRW prepositions, as demonstrated in example (139).

(139) EXP_ST0102: Last but not least, trips should be mentioned because are an important part of learning a subject. However, teachers can plan trips **at** *a knockdown price* where students can learn in a practical way. (POSTB2.WOUTPUT.W1.ST0102)

(140) EXP_ST0106: Last but not least, different materials is a crutial fact that determines the perfect classes. As long as you have better materials, you would expand your knowledge, so you will not feel **under** *the dumps*. (POSTB2.WOUTPUT.W1.ST0106)

However, this deeper understanding does not necessarily result in flawless usage. This is evident in example (140), where ST0106 unconventionally employed the idiom "feel down in the dumps", where "down" was replaced with "under". This variation indicates a lapse in the idiom's precise usage, yet it simultaneously reflects the learner's awareness of the metaphorical meanings of the preposition *under*. Drawing on its basic meaning ("below or at a lower level than something, or covered by something" [LM1]), the learner conveyed the idea of unhappiness or being affected by a negative disposition ("affected by a particular condition, influence, or situation" [LM4]). Such instances highlight that while metaphor-mediated instruction can foster awareness of metaphorical meanings, the precise application of these insights might still pose challenges for some learners (Boers et al., 2009).

Example (141) further illustrates the potential of metaphor-mediated instruction in enhancing awareness of varied metaphorical meanings during B2 training. The deeper comprehension of the preposition *under* ("below or at a lower level than something, or covered

by something" [LM1]) in example (141), conveying the notion of "affected by a particular condition, influence, or situation" LM4) is evident through the idiom "feel under the weather", which was employed in oral discourse.

(141) EXP_ST0106: <*{Should students have to do outdoor sports lessons at school?}*>
<POSTB2.ST0106> speaking personally it's a good thing because not always you have you have to hit the books because you can feel **under** the weather so while you are making doing sports you can keep fit while you are having a lot of fun close to the environment (.) </POSTB2.ST0106>

To gain deeper insights into the learning outcomes in each group, specific instances of how participants from the control and experimental groups used MRW nouns in their written essays are illustrated in examples (142) through (147).

(142) EXP_ST0103: To begin with, it goes without saying that teachers should learn how to deal with technology. Maybe, some people think they are over the **hill** to work with these techniques but, people should lift their spirits instead of judging them. (POSTB2.WOUTPUT.W1.ST0103)

(143) EXP_ST0109: Lastly, it goes without saying that students will have to face the **hurdle** of do not having electronical devices. Because in the future, all things will be digitalised, like homework or books. (POSTB2.WOUTPUT.W1.ST0109)

(144) EXP_ST0108: Nowadays, humanized robots are a work in **progress**, but it is highly likely that they will be the ones teaching students in the future. Thus, replacing actual human teachers. (POSTB2.WOUTPUT.W1.ST0108)

(145) EXP_ST0105: Secondly, it is safe to say that lots of subjects are underestimate. Normally, schools determine the subjects for their students, instead of letting them choose their **path**. (POSTB2.WOUTPUT.W1.ST0105)

(146) CTRL_ST0408: […] despite of all the **progress**, students will not learn more. (POSTB2.WOUTPUT.W1.ST0408)

(147) EXP_ST0108: All in all, I believe schools will, indeed, change in the future. But we have still got a long **way** to go until then. Easy does it! (POSTB2.WOUTPUT.W1.ST0108)

One intriguing finding was L2 learners in the experimental group's incorporation of vocabulary linked to the CM LIFE IS A JOURNEY in the written essays at the end of the study. This observation becomes evident in examples (142) through (144), where these L2 learners used metaphorical expressions such as *to be over the hill*, *hurdle*, and *progress*. It is important to note that these metaphors were introduced during Phase 1 of the CL-oriented instructional intervention, aimed at familiarising them with metaphors.

Particularly interesting is example (145), where a participant from the experimental group used the MRW noun *path*, which was not explicitly taught in the L2 classroom but underlies the CM LIFE IS A JOURNEY, showing enhanced depth of knowledge. This instance highlights the effectiveness of metaphor-mediated instruction in facilitating a deeper understanding and natural integration of metaphorical language in their written discourse. This participant using *path* demonstrates the success of metaphor-mediated instruction in helping learners develop a more profound and intuitive grasp of metaphorical language, even beyond the direct content of their lessons.

Participants from the control group also incorporated MRW nouns underlying the CM LIFE IS A JOURNEY. Upon examining the MRW noun *progress* in (146), it becomes apparent that there were no differences in meaning and usage when compared to how the experimental group used this MRW noun in (144). L2 learners in both groups used the same extended meaning ("the process of developing or improving" [MM1]) of *progress*, which refers to "forward movement" (MM2) in terms of its basic meaning.

Regarding the MRW noun *way* ("a road, path, direction, etc., that you take in order to get to a particular place" [LM3a]), the control group only made use of the extended meaning as "a method that you use to do or achieve something" (LM1), as illustrated in (147), or "the manner or style in which someone does something or in which something happens" (LM2). However, L2 learners in the experimental group exhibited vocabulary depth by also

incorporating the extended sense of “a distance or a length of time, especially a long one” (LM5), as illustrated in (147). These instances highlight the experimental group’s ability to apply a more varied range of metaphorical meanings.

Specific instances of MRW nouns used by the experimental group in oral discourse are illustrated in examples (148) through (150) to gain deeper understanding of their metaphor usage in spoken productions.

(148) EXP_ST0102: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0102> […] yes on weekends if we have time to to spend it together we spend together and for me doing activities outdoors is very important because if I am under the if I feel under the **weather** and my friends call me to to go outside and to spend great time with with with my friends (.) </POSTB2.ST0102>

(149) EXP_ST0105: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0105> […] we love to go to the park just to relax and talk about the week and all our **ups** and **downs** during the week with exams or different situations but yeah I think that it’s very important to be outdoors with friends (.) </POSTB2.ST0105>

(150) EXP_ST0205: <*{Should students have to do outdoor sports lessons at school?}*> <POSTB2.ST0205> I think it’s a good idea idea because doing sport in a in an indoor place it could be boring because is always the same place and you have you are like in four **walls** and you don’t have the the **liberty** that you have around in the in the outdoors (.) <POSTB2.ST0205>

Among the examples provided, one instance stands out in (148) where the idiom “to feel under the weather” is unconventionally used. This idiom is commonly used in English to describe physical illness or a general state of not feeling well. However, in this specific example, the unconventional use does not fit the standard meaning, as it does not refer to feeling unwell or sick. Instead, it seems that ST0102 intends to express being in a good mood, as shown in example (141). As MacArthur (2010) observes, enhanced metaphor awareness does not necessarily result in conventional ways of using metaphorical language, a point

illustrated by various examples throughout this chapter. Boers (2011) also highlights that exploiting linguistic motivation of language in the L2 classroom does not provide L2 learners with a "foolproof tool" to consistently decipher metaphorical meanings and usages.

As the example illustrates, the challenge for L2 learners in dealing with metaphorical language extends beyond linguistic metaphors to also include linguistic realisations of CMs, especially when these elements are included in their formal instruction. As Littlemore (2023) notes, "conceptual metaphors sometimes help us to understand linguistic metaphors, but they are not always a necessary prerequisite, nor a sufficient condition" (p. 131). This observation highlights the pedagogical disadvantages of drawing explicit attention to the conceptual metaphorical underpinnings of L2 vocabulary (cf. O'Reilly & Marsden, 2023). For L2 teaching purposes, it seems important to distinguish between the conceptual and the linguistic levels.

This finding, which was also observed through participant observation in classroom-based research, raises intriguing questions about the learner's understanding of metaphorical language, especially concerning the CMs GOOD IS UP and BAD IS DOWN. In Unit 9 "Secrets of the mind", L2 learners were initially exposed to these higher-level CMs to explain the link between vocabulary and the underlying CM theme of *happiness and the mind*. Specifically, L2 learners were explicitly made aware of the CM relevant to the textbook topic, HAPPINESS IS VERTICALITY, to enhance their comprehension of metaphors in the oral and written input texts.

It is possible that the CM HAPPINESS IS VERTICALITY might have influenced the later learners' comprehension of the idiom "to feel under the weather" in Unit 11 "Medical Matters" when discussing *health and fitness*. Despite being explicitly made aware of the specific CM HEALTH IS VERTICALITY, explained through different CL-style techniques and provided with pictorials elucidating the meaning of the idiom, it appears that the L2 learners tended to associate the higher-level CMs GOOD IS UP and BAD IS DOWN more directly with feelings, as illustrated in (140), possibly eclipsing its narrower health-related meaning. This confusion could have led to ST0102 misconstruing the preposition "under" as indicating being "down" or "low" emotionally, thereby confusing being sad or unhappy with being unwell or sick.

Example (150) illustrates the use of a direct metaphor. In this case, the lexical marker *like* indicates a direct comparison between *walls* (["an upright flat structure made of stone or brick, that divides one area from another or surrounds an area" LM1]) and the feeling of being restricted and lacking *liberty*. This direct metaphor usage highlights how the participant draws on the concept of *walls* to convey the sense of confinement and limited freedom.

It is worth noting that direct metaphors were not frequently used in the participants' oral discourse at the end of the study. Specifically, the experimental group employed 13 instances (2.59%), with 6 of them (1.20%) being MRW nouns. In contrast, the control group only used one direct metaphor (0.32%), as illustrated in (111), which was also a MRW noun. These findings in L2 discourse align with previous research that shows an underuse of direct metaphors (0.04%) in L1 conversation (see Kaal, 2012, pp. 117–178), EFL conversations (MacArthur, 2016a), and in CLIL textbooks (Alejo-González & García-Bermejo, 2020). This underuse could be attributed to a less elaborate and arguably less deliberate metaphor use (cf. Dorst, 2011; Krennmayr, 2011).

To gain a more comprehensive understanding of the learning outcomes in each group, specific instances of how the control and experimental groups used MRW verbs in their written essays are shown in examples (151) through (154). These examples highlight the differences in the participants' usage of MRW verbs and provide insights into their development of open-class metaphors over the course of the study. By examining these examples, we can better understand how the CL-oriented instructional intervention influenced their language use and metaphorical expressions in their writing, shedding light on the effectiveness of metaphor-mediated instruction in enhancing their vocabulary and potentially written discourse.

The examples provided indicate that the experimental group successfully learnt the vocabulary linked to the CM LIFE IS A JOURNEY, and the effort to raise metaphor awareness proved fruitful. This is evident in their skillful incorporation of metaphorical expressions in the form of MRW verbs, as discussed with MRW nouns earlier. However, it is essential to highlight that the control group also employed metaphors related to the CM LIFE IS A JOURNEY, similar to the experimental group. For instance, participants in both groups used the verb *progress*

("to move forward slowly" [MM2]) to convey the idea of making advancements, particularly in the context of technologies ("to continue to develop or move forward" [MM1]), as shown in (151) and (152).

(151) CTRL_ST0302: […] also because the technology is **progressing** so the schools that use it a lot too. (POSTB2.WOUTPUT.W1.ST0302)

(152) EXP_ST0102: Nowadays with the new technologies, it is easier to **progress** rather than the past. (POSTB2.WOUTPUT.W1.ST0102)

When the control group expressed their views of schools in the future and the possibility of important changes, no additional MRW verbs underlying the CM LIFE IS A JOURNEY were found. For instance, example (151) illustrates the control group's usage of MRW verbs related to travel, but without directly invoking the specific CM. In this instance, ST0407 used the verb *go* ("to travel or move to a place that is away from where you are or where you live" [LM1a) metaphorically to convey the notion of regularly attending school ("to regularly attend school, a church etc" [LM3b]).

(153) CTRL_ST0407: Last but not least, the construction of schools should be mentioned because when you **go** to a educational centre you look if is ugly, old, etc, and you are not going to be happy there. (POSTB2.WOUTPUT.W1.ST0407)

(154) EXP_ST0101: As a matter of fact, our whole life **go** through a process of change just like our future schools. That is it, it is future. (POSTB2.WOUTPUT.W1.ST0101)

While the control group demonstrated some level of metaphoricity in their writing to discuss schooling and changes, they did not explicitly draw upon the CM LIFE IS A JOURNEY in their MRW verb choices. Their usage of MRW verbs appeared to be less directly aligned with the specific CM linking "life" and "future schools", as seen in the experimental group's written essays. For instance, in example (154), ST0101 used the MRW verb *go* to liken the passage of life to a process involving various changes and transformations, much like the future schools ("to experience a difficult or unpleasant situation, feeling etc" [LM1]).

When comparing the usage of MRW verbs in relation to "life" and "future schools", the experimental group exhibited a higher frequency of MRW verbs that emphasised the idea that

both life experiences and educational institutions undergo continuous transformation and evolution, shaping individuals and institutions alike over time. This trend is evident in examples (155) through (159), where the experimental group employed a more diverse range of verbs to establish metaphorical connections between the two contexts. These findings suggest a heightened metaphorical awareness and potential improved written production quality in the experimental group, likely influenced by the metaphor-mediated instruction they received.

(155) EXP_ST0108: The time has yet to **come**, but it is most likely that education will change drastically, as follows. (POSTB2.WOUTPUT.W1.ST0108)

(156) EXP_ST0107: In general, I believe we should **follow** the lead of other countries, like Finland. (POSTB2.WOUTPUT.W1.ST0107)

(157) EXP_ST0106: In conclusion, schools will be very different in the future. As long as we **advance**, schools are changing. (POSTB2.WOUTPUT.W1.ST0106)

(158) EXP_ST0102: In general, if the world keeps on **moving**, there will be a revolution in education. (POSTB2.WOUTPUT.W1.ST0102)

(159) EXP_ST0110: To sum up, it is clear that each era has like its customs and as time **pass**, things are getting more different. (POSTB2.WOUTPUT.W1.ST0110)

Taking an alternative perspective, the experimental group's successful learning of the vocabulary linked to the CM LIFE IS A JOURNEY cannot be solely attributed to mere instruction. Participant observation provided valuable insights into how the experimental group used the metaphors they learnt during Phase 1 of the teaching treatment across various topics from the mainstream textbook. One notable example of this was the recycling of the CM LIFE IS A JOURNEY in Unit 8 "Dream of the stars", which focused on the topic of *career and aspirations*. This observation indicates that the experimental group not only acquired the metaphorical vocabulary but also effectively transferred their classroom practice to new input and integrated it into their oral discourse, adapting it to new contexts such as discussions related to careers.

As shown in (155) through (159), the findings suggest the presence of recycling in the use of CM LIFE IS A JOURNEY when writing their essays about possible important changes in future schools. This indicates that the experimental group also integrated the metaphors into

their writing, adapting it to new contexts such as discussions about future schools. This phenomenon underscores the benefits of enhancing metaphor awareness among L2 learners, enabling them to use a single CM to elucidate multiple abstract concepts. As Littlemore (2023) highlights:

> Conceptual metaphors are though to exist for every abstract concept that we have, although there is no one-to-one mapping; a single abstract concept can be understood through several conceptual metaphors, and a single conceptual metaphor can be used to explain several abstract concepts. (p. 129)

Examples (160) through (172) provide evidence that the experimental group went beyond the incorporation of MRW verbs underlying the CM LIFE IS A JOURNEY when discussing future schools. They enriched their writing by using other CMs, such as CAREER IS A BUILDING, which was presented in Unit 8 “Dream of the stars”. The recycling of metaphors across different topics and themes suggests that the experimental group did not merely learn and apply the CMs in isolation but integrated them into their broader language repertoire, extending their usage to diverse contexts.

(160) EXP_ST0202: Also, they could add other subjects, for example, about how to **get** a job in the future. (POSTB2.WOUTPUT.W1.ST0202)

(161) EXP_ST0104: Finally, another point to take into consideration is that classes could be made at home because the tecnological advantages will **make** it possible. (POSTB2.WOUTPUT.W1.ST0104)

(162) EXP_ST0206: Another idea is that, in the future there will be new jobs that will encourage teenagers to **hit** *the big time*. What is more, students will be able to learn how to discover things, and these will help them to become a star. (POSTB2.WOUTPUT.W1.ST0206)

It must be acknowledged that participants from control group used the verb *reach* (“to touch something by stretching out your arm” [LM5b]) metaphorically to convey the idea of *success*, specifically that technology will enable or facilitate the accomplishment of the needed transformation or improvement in the field of education, as shown in (163).

(163) CTRL_ST0401: Moreover, I strongly believe that education is in need of a change and technology will help us to **reach** it. (POSTB2.WOUTPUT.W1.ST0401)

However, it should be noted that ST0401 from the control group used this MRW unconventionally to suggest that technology can act as a means or tool to bring about the desired changes in the educational system. Instead of using more conventional verbs such as "bring about" or "achieve", ST0101 opted for *reach* to imply that technology has the potential to bridge the gap between the current state of education and the desired future state of education. This alternative use of the MRW verb *reach* may be influenced by the L1, as the Spanish translation of *reach* (*alcanzar*) is often used metaphorically in contexts related to *careers and aspirations*. This finding suggests a potential transfer of metaphoric language from the L1 to the L2, reflecting a less heightened metaphor awareness for the control group, or it could be metaphor awareness of L1 metaphors.

The success of the experimental group in learning the CM LIFE IS A JOURNEY vocabulary extends beyond direct instruction. The findings draw on the potential overlapping of CMs (Lakoff & Johnson, 1980) and provide insights into why the vocabulary related to one topic, such as *life*, can be learned and used more frequently than others. This increased frequency of usage is likely due to encountering the CM LIFE IS A JOURNEY in different contexts, such as when discussing *career and aspirations* or *changes*. This versatility in metaphorical expression likely contributed to their enhanced written production quality. The experimental group's successful incorporation of these MRW verbs suggests that metaphor-mediated instruction can play a crucial role in bridging the gap between learners' language proficiency and the expectations of native-like expression.

In example (162), participant ST0206 provides a remarkable example of recycling of metaphors. Not only does this student used the phrase "hit the big time" to address the theme of *careers and aspirations* — covered in Unit 8 "Dream of the stars" from the mainstream textbook — but also, she skillfully incorporated the MRW noun *star*. This noun was explicitly introduced through CL-inspired instruction as part of their B2 training in Unit 8 to offer support in fully grasping the semantic potential of newly acquired words from their classroom learning.

Having received metaphor-mediated instruction, the experimental group used some of the metaphorical expressions encountered throughout Unit 9 “Secrets of the mind”, showing the application of their learning. As detailed in Section 3.2 of Chapter 3, the CM HAPPINESS IS VERTICALITY was used to support L2 learners in dealing with the metaphorical uses of the language, addressing the topic of *happiness and the mind*. The experimental group included the phrasal verb *cheer up* (see [164]), verbal phrases such as *come to terms with* (see [165]) and *put up your spirits* (see [166]), and the idiom *walk on air* (see [167]) in their oral outputs.

(164) EXP_ST0206: <*{Should students have to do outdoor sports lessons at school?}*> <POSTB2.ST0206> […] I believe that this is an an important point because teenagers spend a lot of time for example in front of books or a computer and I think that schools should like **cheer** *up* them to practise sport outdoors (.) </POSTB2.ST0206>

(165) EXP_ST0203: <*{Do you think it’s true that if people spent more time outside, they’d care more about protecting the environment?}*> <POSTB2.ST0203> yes I agree with him because sometimes the people are don’t try to **come** *to terms* with the situation so maybe they think okay it’s a problem but it doesn’t matter so but maybe the people who live in the nature out sometimes think okay I have to face the problem […] </POSTB2.ST0203>

(166) EXP_ST0105: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0105> […] because are the very important things for us and also being out with friends usually **puts** *up our spirits* because we sometimes **feel** *low* or things like that and they always are there to help you so (.) </POSTB2.ST0105>

(167) EXP_ST0203: <*{Should students have to do outdoor sports lessons at school?}*> <POSTB2.ST0203> […] I think it’s a different way to learn and also have fun and I think for example in my case when we have physical education and we go out to do sports I’m like **walking** *on air* [...] </POSTB2.ST0203>

These findings show that raising metaphor awareness is not sufficient to ensure the effective use of metaphors, as illustrated for example in (164), considering this study primarily

focused on enhancing metaphor awareness among L2 learners, rather than specifically investigating vocabulary acquisition and retention of metaphors as pre-taught vocabulary.

As Littlemore and Low (2006) point out, “the usefulness of querying routines may be more limited when it comes to [metaphorical] language production […], you need to remember the form of the expression, check that its meaning is appropriate to the topic, the audience and the type of discourse and check that the grammar and collocation are appropriate” (p. 25). In example (164), the phrasal verb *cheer* up was used instead of the verb *encourage* to refer to “to provide conditions that help something to happen” (MM2). A possible explanation could be that the Spanish verb *animar* is used to denote both English terms. More detailed examination is needed to discern whether these unconventional uses arise from rote learning or possible mistranslation of meaning in the context of metaphor-mediated instruction.

The enhanced depth of vocabulary knowledge of the experimental group is evident in their usage of the verb *feel* (see [166]), which was accompanied by the MRW adjective *low* when referring to “sad”. In contrast, the control group did not display a MRW usage when referring to feelings (see [168]).

(168) CTRL_ST0405: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0405> yes always we can go outside we try to do it because we **feel** *better* we have some kind of more freedom outside at the street as we are at home and we can go everywhere and no more people I think it’s the best way to spend the time with the friends (.) </POSTB2.ST0405>

These qualitative findings provide compelling evidence that the experimental group’s spoken production quality was notably improved, indicating they may be better supported in completing the *Speaking* test at B2 level.

Another interesting observation was made regarding the usage of *meet out* in (169) by ST0106 from the experimental group. In this specific context, the particle *out* was not correct as the phrasal verb *meet up* is typically used to denote “to come together with someone, either unexpectedly or as planned” (MM1). However, this participant used *out* instead of “up” as an alternative to referring to meeting in a more informal and friendly context.

(169) EXP_ST0106: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0106> yes on Fridays we **meet out** and go to running or for a walk and also going for a tour on bicycle and but other we do indoors like doing projects at my or their or my friends' houses […] </POSTB2.ST0106>

This unconventional use highlights how L2 learners frequently process phrasal verbs as full decomposable expressions or novel compounds, instead of fixed chunks of language (MacArthur & Littlemore, 2011). For example, they may add non-standard particles or adverbs (see e.g., Alejo-González, 2010). It is possible that this L2 learner inserted the MRW adverb *out* to emphasise the sense of *meet* ("to come together in order to talk to someone who you have arranged to see" [MM1]) by adding one of the extended meanings of *out* defined as "to or in a place that is not your home, in order to enjoy yourself" (LM3a). However, this observation suggests that possibly this learner was not aware of the extended meanings of "up". Consequently, L2 learners might need more support, particularly in fostering awareness of the underlying motivation of particles in phrasal verbs.

The use the MRW adjective *big* by the participants showed variations when discussing the important changes expected in future schools, as shown in examples (170) and (171). In this context, the participants from the control group used *big* ("large in size" [MM1]) metaphorically to convey the idea that the role of teachers and the method of instruction are likely to undergo a substantial transformation, highlighting the magnitude of the shift in educational practices envisioned in the future ("large in degree, or having a strong effect" [MM2]), as illustrated in example (170).

(170) CTRL_ST0402: In addition, I strongly believe that the **biggest** change will be the teacher because are used to treat with our teachers face to face not to a robot online. (POSTB2.WOUTPUT.W1.ST0402)

(171) EXP_ST0106: First of all, people say that robots are going to be the teachers so we can learn without hitting the books. Basically, this way would help people become a **star** and hit the **big** time. (POSTB2.WOUTPUT.W1.ST0106)

Conversely, the experimental group leaned on the use of the MRW adjective *big* to convey the idea of achieving significant success or prominence ("important or major" [MM3]). For instance, in example (171), *big* is used in the phrase "hit the big time", suggesting that schools need to change their teaching methods into a more modern approach to achieve significant success or make a substantial impact.

These distinct usages highlight that both groups of participants used the MRW *big* for evaluative purposes, each with different emphases. Of particular interest is the usage of *big* by the experimental group, particularly in the context of *success*. This specific use resonates with the idea of recycling metaphors L2 learners are familiar with, thereby potentially enhancing their ability to convey their ideas in the *Writing* task.

Another intriguing example is (172), where participant ST0103 from the experimental group incorporated the MRW adjective *rushed* into her written essay. This adjective is a term frequently employed in the idiomatic expression "rushed off one's feet" to convey the idea of being constantly on the move and having very little time to rest or relax.

(172) EXP_ST0103: On the top of that, being **rushed** *of the feet* will not be a problem because of writing on a computer instead of by hand. (POSTB2.WOUTPUT.W1.ST0103)

Importantly, it should be noted that this idiom was intended to be learnt as part of Unit 11 "Medical Matters", focusing on *health and fitness*. L2 learners received support through CL-oriented methods, including pictorial elucidation, to understand the meaning of various idioms, such as "be rushed off your feet" or "be thrown in at the deep end".

It is worth noting that, in this particular instance, ST0103 mistakenly used this idiom as "rushed of (written with a single "f") the (using the wrong determiner) feet" intending to convey the idea of extremely busy or overwhelmed with a lot of work or tasks. This observation highlights that raising metaphor awareness alone might not guarantee flawless recall of the exact wording of metaphorical language that L2 learners have been explicitly introduced to, as Low (2008) emphasises. However, it should be stressed that despite this minor spelling oversight and the error with the determiner, the student successfully deployed the correct

usage in context within her writing. This finding strengthens the observation that CM, while it can be exploited for the teaching of metaphor, may not be useful in helping learners acquire the phraseological patterns that accompany the metaphors and that play an important role in determining their meaning (Philip, 2011). Further attention should be paid to phraseological patterns when developing CL-inspired activities, emphasising the idea that CM needs to be supplemented by other approaches that help learners use metaphorical language in "phraseologically appropriate ways", as Littlemore (2023, p. 136) points out. In this regard, Philip's (2011) "discovery learning" approach to encourage learners compare collocational patterns of words in the L2 with those in their L1 might be useful for two aims: material design and for learner awareness of metaphor.

Consistent with the observations concerning MRW verbs in oral discourse, where unconventional uses could stem from rote learning or potential mistranslation, example (173) presents a scenario where raising metaphor awareness did not effectively facilitate the L2 learner in using the idiom "be thrown in at the deep end" correctly. This inaccuracy arose not only from the substitution of the article "a" for "the", but also from the misalignment of the idiom with the context.

(173) EXP_ST0101: Last but not least, it is interesting to think about books. Nowadays, computers and technological instruments are used instead of them. It results hard and misunderstanding. For this reason, people will be thrown in at a **deep** end. (POSTB2.WOUTPUT.W1.ST0101)

Typically, this idiom describes a situation where someone is put into a challenging or difficult situation without adequate preparation or support. ST010's attempt to incorporate the idiom "be thrown in at the deep end" did not yield the intended outcome, as the phrase was inaccurately applied to convey the notion that people facing challenges and difficulties during the transition from using books to computers and technological instruments. The idiom seems somewhat out of place in this instance, as it is primarily linked to scenarios involving unfamiliar tasks or responsibilities, and not inherently tied to the context of technology usage. A more conventional expression might be "thrown into unfamiliar territory" or "thrown into the

unknown", which would better convey the idea of facing difficulties when adapting to new technology or tools. This example aligns with Littlemore's (2023) observation that while linguistic motivation can assist L2 learners in understanding language input, it does not necessarily ensure their ability to produce appropriate language forms.

Among these examples, all instances followed conventional uses, except for the experimental group's usage of *thrown in* in the idiomatic expression "be thrown in at the deep end", which has been earlier discussed regarding MRW adjectives in written discourse.

(174) EXP_ST0101: Last but not least, it is interesting to think about books. Nowadays, computers and technological instruments are used instead of them. It results hard and misunderstanding. For this reason, people will be **thrown in** *at a deep end*.

The experimental group's enhanced depth of vocabulary knowledge is also evident in their unconventional uses of MRW adjectives in oral discourse. For instance, in the case of ST0105, she used the phrase *down ages* instead of "down times" (see [175]), demonstrating awareness of the extended meaning of *down* as referring to sadness, low spirits, or challenges (MM15). ST0105 opted for a metaphorical use of *ages* defined as "a period in history" (MM4), intending to covey "times", defined as "a period in history" (MM2a) or "a period in someone's life" (MM2b), which referred to the challenging period of the COVID-19 pandemic.

(175) EXP_ST0105: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0105> yeah I now that we are like isolated I realised that is very important to be with friends and outdoors now the things that I really miss are my friends and I mean outdoors so that's it it's like the key for us when we are in like **down ages** because are the very important things for us and also being out with friends usually puts up our spirits because we sometimes feel low […] </POSTB2.ST0105>

This specific instance illustrates how the experimental group's heightened metaphor awareness allowed them to incorporate alternative expressions, even if some collocational nuances were not entirely adhered to. The experimental group's ability to use metaphorical language reflects their deeper understanding and mastery of English vocabulary, contributing to potentially enhanced spoken production.

The experimental group exhibited a more enhanced command of phrasal verbs, as illustrated in (176) and (177). For instance, ST0105's use of the phrasal verb *come out* in (176) reflects understanding of the extended meaning of the verb *come* ("to move towards you or arrive at the place where you are" [LM1]) paired with the particle *out* ("from inside an object, container, building, or plane" [LM1]), particularly in the context of something becoming publicly known or revealed. Notably, the L2 learner combines the phrasal verb *come out* with the idiom "open a can of worms". This suggests that the learner not only grasps the individual nuances of each expression but is also able to coherently combine them in a meaningful manner to discuss a controversial topic she views as contentious. However, it is important to note that the more conventional phrasal verb in such contexts is typically "come up". Unlike *come out*, which implies revelation or disclosure, "come up" often refers to issues that arise or are brought to attention in discussions.

(176) EXP_ST0105: It is a well-known fact that schools later this years have become a controversial subject. Every time it **comes out** is *like* **opening a can of worms**. (POSTB2.WOUTPUT.W1.ST0105)

(177) EXP_ST0203: Furthermore, another reason to consider is the fact that the school can **wear** student **out**. (POSTB2.WOUTPUT.W1.ST0203)

Additionally, example (176) shows the use of a direct metaphor. Here, the lexical marker *like* signals a direct comparison between the act of "coming out" and the notion of probing a particular issue. This direct metaphor underscores how the participant draws on the idea of "becoming publicly known or revealed" to convey the idea of exposing a sensitive or controversial topic. It should be highlighted that direct metaphors were infrequently used in the participants' written discourse, mirroring the patterns observed in the oral discourse findings. Specifically, the experimental group employed six instances (0.80%), while the control group only recorded three direct metaphors (0.48%).

Furthermore, ST0203's usage of *wear out* in (177) illustrates an enhanced understanding of this phrasal verb. In the given context, it means to make someone feel

extremely tired. By applying this phrasal verb, the L2 learner effectively conveys the idea that school can be mentally and physically taxing, leading students to feel exhausted.

Overall, the qualitative findings highlight the experimental group's positive progress in incorporating MRW adverbs, especially when used as particles in phrasal verbs. Yet, mirroring insights from oral discourse, these L2 learners could further benefit from targeted assistance to deepen their understanding of the metaphorical motivation of particles in English.

5.2. ANALYSING THE IMPACT OF METAPHOR AWARENESS IN THE EFL SYLLABUS ON B2 PERFORMANCE (RQ2): KEY FINDINGS

No statistically significant differences were observed between the two groups of participants in terms of their overall B2 achievement and performance in *Speaking* and *Writing* in the pre-testing phase. Both groups demonstrated a global attainment of B2 level in oral skills, while their overall performance and written skills were below this level prior to implementing the CL-oriented instructional intervention. However, some initial statistically significant differences were identified in specific areas, such as the use of organisational patterns in both oral and written discourse. Importantly, these differences were unrelated to vocabulary production, which was the primary focus of study in this investigation.

5.2.1. Achievement of L2 Learners at B2 Level: *Overall* Performance

The findings of the study reveal that while raising metaphor awareness among L2 learners can be impactful, it does not necessarily guarantee improved overall performance in the *B2 First for School* examination. Despite the experimental group showing significant improvement, they failed to meet the minimum score threshold of 160 (UCLES, 2015). In contrast, the control group, which did not receive explicit metaphor instruction, unexpectedly attained a higher score (EXP = 154.20 vs. CTRL = 156.65). This finding raises questions about the direct impact of metaphor awareness on exam success.

Caution is warranted when interpreting the lack of significant differences in B2 level proficiency and progress between the control and experimental groups. It is essential to recall

that the CL-oriented instructional intervention focused on raising metaphor awareness and promoting metaphor usage in L2 learners' oral and written discourse, aligning with established instructional practices, following MacArthur (2010) and Boers (2013). The non-significant difference in effectiveness between both teaching approaches for achieving overall B2 level proficiency suggests that other factors beyond the CL-oriented instructional intervention could have contributed to these results. For example, individual differences, external influences, or the incomplete coverage of the EFL syllabus due to the COVID-19 pandemic.

The findings also indicate that the CL-oriented instructional intervention had a positive impact on the experimental group's productive skills, especially in *Speaking*, as detailed in Section 5.2.2 of this chapter. However, this intervention alone was insufficient to elevate their overall performance to the desired B2 level, as assessed by the examination used in this study. This highlights the need to incorporate additional B2 training within metaphor-mediated instruction to assist L2 learners in achieving their overall performance targets at the B2 level.

5.2.2. Achievement of L2 Learners at B2 Level: *Speaking* Performance

5.2.2.1. Key Findings from Speaking Performance

It is important to remember that both participant groups had already achieved a global B2 level of English proficiency in *Speaking* (Grade C) before the teaching intervention began. Both groups demonstrated positive progress in their *Speaking* performance by the end of the study. This was despite a noticeable decline in the use of metaphors in the spoken productions of the control group.

The CL-oriented instructional intervention seems to have been more effective in aiding the experimental group to achieve a significantly higher global B2 proficiency in *Speaking* (EXP = 171.85 vs. CTRL = 165.30). This group showed significant growth, bringing them closer to a Grade B (173–179) level. In comparison, the control group attained a lower Grade C (160–172) at the end of the study, as per UCLES (2015) standards.

It is important to note that, despite the lack of statistically significant differences in progress between the groups, the control group also showed improvement in their *Speaking* skills. This suggests that factors beyond the CL-inspired instructional intervention may have influenced these results (see further discussion in Section 5.3 of this chapter).

The CL-oriented instructional intervention significantly enhanced various aspects of L2 learners' *Speaking* performance, both in terms of level and progress, with a special focus on enriching vocabulary production. In contrast, the control group, which did not receive CL-oriented pedagogical training, showed a notable decline in this specific area. This seems to have led to marked differences between the two groups in terms of vocabulary usage in oral discourse. These findings emphasise the efficacy of metaphor-mediated instruction in improving L2 learners' *Speaking* skills, particularly in vocabulary enrichment.

When interpreting these results, considering the "Grammar and Vocabulary" assessment descriptor becomes essential, as it evaluates both grammatical and lexical knowledge. Thus, the findings indicate advancements in both grammar and vocabulary skills among the experimental group in their speech.

5.2.2.2. Topic-Related Uses of Metaphors: Some Insights from Spoken Production

The significance of topic similarity in successfully demonstrating L2 competence in high-stakes ESOL exams should be noted (Council of Europe, 2020). According to the *Speaking* assessment criteria for *B2 First for Schools* examination, the ability to use appropriate vocabulary and effectively express thoughts on a wide range of familiar topics is essential for achieving success in the L2 assessment (UCLES, 2019a, p. 82). Hence, incorporating vocabulary related to the topic discussed in the task can provide support for L2 learners to successfully complete the task in the *Speaking* test, thereby enhancing their English language performance at B2 level.

It is important to acknowledge that this part of the study will be of a more exploratory nature than the preceding analyses reported in Chapter 4. By focusing on the occurrence and effectiveness of topic-based metaphors, we can shed more light on how the CL-oriented

instructional intervention influenced learners' use of metaphorical language, ultimately preparing them for real-life communication. By extension, the examination of topic-related uses of metaphor can provide further insights into the potential impact on their *Speaking* performance. An enhanced ability to employ metaphors in a manner relevant to different topics can indicate a higher level of language proficiency, as it reflects the participant's adaptability and command of metaphorical expressions in various contexts.

At the conclusion of the study, distinct patterns emerged in the MRW-type rates of both the control and experimental groups within their spoken productions (see Section 4.1.2.2 of Chapter 4). Upon conducting a closer examination of MRW nouns, the most frequently used open-class metaphor in the spoken productions of both groups, it becomes evident that participants occasionally employed topic-based metaphorical language when discussing the topic of *spending time outdoors* in the post-test.

Table 58 presents examples 178–186, showing the 14 out of 40 (35%) topic-based MRW nouns[44] (RF = 17.69%) that participants from the experimental group used in their post-test oral discourse. For further details, see Table 116 in Appendix Q.

Table 58

Topic-Related Metaphor Uses in the Oral Discourse of Experimental Group Participants: Post-Test

Spoken prompt	Examples	
Do you and your friends spend a lot of your time doing outdoor activities?	(178)	<POSTB2.ST0105> […] now that we are like isolated I realised that is very important to be with friends and outdoors now the things that I really miss are my friends and I mean outdoors so that's it it's *like* the **key** [B2 level] for us when we are in like down ages because are the very important things for us and also being out with friends usually puts up our **spirits** [B2 level] because we sometimes feel low or things like that and they always are there to help you so (.) </POSTB2.ST0105>
Should students have to do outdoor sports lessons at school?	(179)	<POSTB2.ST0205> […] doing sport in a in an indoor place it could be boring because is always the same place and you have you are *like* in four **walls** [A1 level] and you don't have the the **liberty** [B2 level] that you have around in the in the outdoors (.) </POSTB2.ST0205>
	(180)	<POSTB2.ST0208> […] I think that is a be- a better option to do it outdoors with the natural **world** [B1 level] around you and when you can breathe very fresh air and be more comfortable than inside (.) </POSTB2.ST0208>

[44] See Section 3.3.3 of Chapter 3 for explanation.

Do you think it's a good idea for families to go on holidays where they spend a lot of time outside, for example camping holidays? Is growing up in the countryside better for children than living in cities?	(181)	<POSTB2.ST0102> […] it's a good idea to they go to the beach because they like it a lot and it's like if you if it's *like* a **routine** [B1 level] that you spend in another part of the world knowing new things […] </POSTB2.ST0102>
	(182)	<POSTB2.ST0101> […] I consider that it's a good idea now also for children that are actually currently they are always in front of a **screen** [A2 level] so it could be a good idea to disconnect for the technologies and go camping (.) </POSTB2.ST0101>
	(183)	<POSTB2.ST0210> […] in the city it's worse because here is more difficult to practise all the activities we want and if we want to do something we have to go to the city **centre** [A2 level] and is it would be difficult what do you think? </POSTB2.ST0210>
	(184)	<POSTB2.ST0104> […] it's more safe to be in the **country** [A2 level] that in the city but is also real that you have more **facilities** [B1 level] in the city because like more people live there and also you have more shops. </POSTB2.ST0104>
	(185)	<POSTB2.ST0205> I totally agree with you because the villages or in the countryside you have more **independence** [B1 level] because there is less **traffic** [A2 level] and also everybody knows everyone […] </POSTB2.ST0205>
	(186)	<POSTB2.ST0105> […] it's like a safer place and there are like different things that makes **life** [B1 level] in countryside safer but I don't know being in a city also makes you like being very independent and to know different **problems** [A1 level] or things that can happen to you […] </POSTB2.ST0105>

To assess the proficiency levels of these nouns, the *Cambridge English Dictionary* online (CAMD) was employed, as it provides information about the assigned CEFR levels to words. According to the CAMD, three topic-based MRW nouns (*key* [178], *spirit* [178], and *liberty* [179]) were used at a B2 level, constituting 21.42% of the total topic-related usage. The above examples clearly show that the participants in the experimental group did not use any of the topic-based MRW nouns that were explicitly prompted in the *Speaking* post-test.

Examples 187–192 in Table 59 shed light on which topic-based MRW nouns participants from the control group used in the *Speaking* post-test. These participants employed 7 out of 19 (36.84%) topic-based MRW nouns (RF = 13.09%) in their post-test spoken productions. For further details, see Table 116 in Appendix Q.

Table 59

Topic-Related Metaphor Uses in the Oral Discourse of Control Group Participants: Post-Test

Spoken prompt		Examples
Do you and your friends spend a lot of your time doing outdoor activities?	(187)	<POSTB2.ST0304> yes we always go to the to the sport **centre** [A2 level] and we practise football or bask- or basketball at the weekend (.) </POSTB2.ST0304>

Should students have to do outdoor sports lessons at school?	(188)	<POSTB2.ST0306> […] if we if others don't say to us that we have to do exercise we don't going we aren't going to do it so I think it's really important to have **lessons** [A1 level] outdoors and practise some different sports. </POSTB2.ST0306>
Is growing up in the countryside better for children than living in cities?	(189)	<POSTB2.ST0301> […] I think it's there are this ter- are different ways to grow up no one is better than other because maybe in the countryside you have more **freedom** [B2 level] but there are less people than in the cities […] </POSTB2.ST0301>
	(190)	<POSTB2.ST0307> […] it's a good way to live in a countryside for example to take care of the environment and to take a **rest** [A2 level] what do you think? </POSTB2.ST0307>
	(191)	<POSTB2.ST0405> […] I think we we learn more things if we grow in a in a countryside because as our yes we discover new **world** [B1 level] instead of doing the same thing as the people we have surround </POSTB2.ST0405>
Do you think it's true that if people spent more time outside, they'd care more about protecting the environment?	(192)	<POSTB2.ST0405> […] if we go outside we see how is the world being damaged and we know the pollu- what the what **damage** [B1 level] are doing the pollution to to the world we are going to see all the **rise** [B2 level] in seas and those types of things (.) </POSTB2.ST0405>

Similar to the experimental group, participants in the control group had limited usage of topic-based MRW nouns targeted at B2 level. According to the CAMD, two MRW nouns (*freedom* [189] and *rise* [192]) were used at this level, constituting 28.57% of their overall topic-related usage. Additionally, it is worth noting that the two specific MRW nouns (*centre* and *world*) used by both groups in their spoken productions after the teaching treatment did not meet the B2 level of proficiency. This is not unexpected, as inclusion of lexis from lower proficiency levels is common. Furthermore, it is noteworthy that the control group's usage of the topic-based MRW noun *lesson* was explicitly prompted by the *Speaking* test materials, which was not the case for the experimental group.

In the spoken productions of both groups, a notable example was the limited range and frequency of topic-based MRW nouns at the B2 level, as defined by the CAMD. This limitation may be attributed to the intrinsic characteristics of oral discourse, which tends to favor more spontaneous and less complex language structures. Despite the challenges in using more sophisticated B2-level MRW nouns, the experimental group distinguished itself by demonstrating a higher frequency of such usage, even without explicit prompts for MRW nouns. This contrast with the control group reinforces the positive impact of the CL-oriented

instructional intervention on the experimental group's metaphor production, vocabulary improvement, and their overall *Speaking* performance.

The findings of the study provide insights into the significant support that metaphors offer for the development of L2 learners' *Speaking* proficiency at the B2 level, especially in terms of vocabulary usage. Yet, the study also indicates that incorporating topic-based MRWs into L2 learners' spoken discourse may not be essential for achieving success in *Speaking*. Instead, the broader positive impact of the CL-oriented intervention on metaphor usage and oral skills appears to be more influential, leading to enhancement in participants' *Speaking* abilities without an excessive focus on specific topic-based MRWs. Future research should delve deeper into investigating the specific benefits of metaphor usage in enhancing oral skills, particularly in relation to the CEFR can-do statements at B2 level.

It is essential to acknowledge that the exploration of topic-related uses of metaphors was limited to nouns, while other open-class metaphors produced in oral discourse (i.e., verbs, adjectives, or adverbs) were not examined. Considering that different word classes can have varying impacts on L2 learners' metaphor production (see Section 5.1.2 of this chapter), further examination is required to delve deeper into the role of topic-related uses of various word classes in metaphorical language usage among L2 learners.

5.2.3. Achievement of L2 Learners at B2 Level: *Writing* Performance

5.2.3.1. Key Findings from Writing Performance

The study's findings suggest that L2 learners can be effectively supported in developing their *Writing* skills without necessarily relying on CL-inspired instruction. Initially, neither the control nor the experimental group attained the B2 level in the pre-test *Writing* assessment, highlighting areas for improvement. Nonetheless, both groups demonstrated notable advancement by the end of the study (EXP = 163.40 vs. CTRL = 167.50), each achieving a global B2 level of English proficiency in *Writing*, as evidenced by their Grade C (160–172) achievement (UCLES, 2015).

The control group, which did not receive metaphor-mediated instruction, showed both significant progress and a higher proficiency level in *Writing* as assessed by the examination used in this study. When comparing the evolution of both groups' *Writing* proficiency at the end of the study, the findings reveal that metaphor-mediated instruction had a more limited impact, with significant differences in favour of the control group.

The control group exhibited superior performance across all four *Writing* assessment descriptors, underscoring that L2 learners can effectively enhance their written skills without receiving metaphor awareness for B2 training. This superiority was particularly noticeable in areas such as "Communicative Achievement" and "Organisation". While these aspects were not the primary focus of the study, acknowledging these significant differences is crucial as they offer insights into the control group's strengths in specific *Writing* proficiency areas.

These findings suggest that the control group might have been more adept at using the conventions of the communicative task to engage the target reader effectively and express their ideas clearly, as demonstrated by their use of specific vocabulary and key structures. Furthermore, their ability to develop well-organised and coherent texts, employing various cohesive devices and organisational patterns, further indicates their proficiency in written communication. This may indicate the disadvantage for L2 learners of the time devoted to explicit teaching of metaphor as opposed to time devoted to exam training as regards the creation of cohesive texts.

The proficiency of the control group in these assessment descriptors likely played a key role in the statistically significant difference observed between the two groups regarding the development of their *Writing* skills. These findings prompt further questions about the role of pre-fabricated language, both metaphorical and non-metaphorical, in contributing to their higher scores across various areas of the *Writing* test.

It is crucial to consider various domains, including vocabulary production, to achieve a comprehensive and balanced performance in *Writing*. At the end of the study, significant progress was observed in both groups concerning the "Language" assessment descriptor. Similar to the "Vocabulary and Grammar" criteria in the *Speaking* assessment, this descriptor

evaluates both grammatical and lexical knowledge in *Writing* (UCLES, 2019a, p. 34). The findings indicate notable improvements in both grammar and vocabulary in the participants' writing. The control group showed better performance in this aspect, although there were no significant differences in the proficiency level and evolution between the two groups.

The findings suggest that the CL-oriented instructional intervention had a more limited impact on supporting written skills compared to oral skills, as discussed in Section 5.2.1 of this chapter. Despite the experimental group's significant improvement in both modalities, the influence of metaphor-mediated instruction on *Writing* skills appears somewhat less pronounced than on *Speaking* skills. This disparity underlines the necessity of further research to better understand how different instructional approaches affect various language skills.

5.2.3.2. Topic-Related Uses of Metaphors: Some Insights from Written Production

Effectively addressing the set topic is crucial for demonstrating L2 competence in high-stakes ESOL exams, as outlined by the Council of Europe (2020). Enhanced vocabulary usage, particularly when it aligns with the *Writing* task's topic, plays a significant role in achieving this. By using topic-relevant vocabulary, L2 learners are better supported in skillfully crafting their essays, which contributes to improved *Writing* skills at the B2 level.

Consistent with oral discourse, the post-test written productions of both the control and experimental groups exhibited distinct patterns concerning MRW types, as detailed in Section 4.1.2.2 of Chapter 4. A notable finding emerged from an analysis of MRW verbs, identified as one of the open-class metaphors showing substantial differences between the groups. Specifically, the experimental group used topic-based metaphors more frequently than the control group in their essays. This was particularly evident when expressing views on the future of schools and the possibility of significant changes, as per the instructions for the post-test written task.

Examples 193–213 in Table 60 illustrate the topic-based metaphors employed by the experimental group in their post-test essays. These participants used 25 out of 66 (37.87%) topic-based MRW verbs (RF = 35.16%). For further details see Table 122 in Appendix R.

Table 60

Topic-Related Metaphor Uses in the Written Discourse of Experimental Group Participants: Post-Test

Written prompt	Examples
Some people say that schools will be very different in the future. What do you think?	(193) In England, there are subjects very useful which focus on students future and dreams. So, hopefully these will be **accepted** [B2 level] in schools all around the world. (POSTB2.WOUTPUT.W1.ST0105) (194) In terms of learning, most goverments will need to **adapt** [B2 level] to a new form of teaching such as online. (POSTB2.WOUTPUT.W1.ST0102) (195) In conclusion, schools will be very different in the future. As long as we **advance** [B2 level], schools are changing. (POSTB2.WOUTPUT.W1.ST0106) (196) The time has yet to **come** [B2 level], but it is most likely that education will change drastically, as follows. (POSTB2.WOUTPUT.W1.ST0108) (197) All in all, in my view, schools will be very different in the future. Because, as I see it, everybody will have to know how to **deal** [B1 level] with technologies. (POSTB2.WOUTPUT.W1.ST0209) (198) It is often thought that schools have improved a lot so it is not necessary to **do** [A2 level] any change. In contrast, it is important to **make** [A2 level] changes that will improve the education. (POSTB2.WOUTPUT.W1.ST0203) (199) It is a well-known fact that humans are going to **evolve** [C1 level]. Nowadays with the new technologies, it is easier to **progress** [B2 level] rather than the past. (POSTB2.WOUTPUT.W1.ST0102) (200) Last but not least, different materials is a crutial fact that determines the perfect classes. As long as you have better materials, you would **expand** [B2 level] your knowledge, so you will not **feel** [A1 level] under the dumps. (POSTB2.WOUTPUT.W1.ST0106) (201) It is a well-known fact that all the people **expect** [B2 level] some changes in their schools in the future. (POSTB2.WOUTPUT.W1.ST0202) (202) Lastly, it goes without saying that students will have to **face** [B2 level] the hurdle of do not having electronical devices. (POSTB2.WOUTPUT.W1.ST0109) (203) In general, I believe we should **follow** [C2 level] the lead of other countries, like Finland. (POSTB2.WOUTPUT.W1.ST0107) (204) To sum up, it is clear that each era has like its customs and as time **pass** [B1 level], things are **getting** [B1 level] more different. (POSTB2.WOUTPUT.W1.ST0110) (205) As a matter of fact, our whole life **go** [B2 level] through a process of change just like our future schools. (POSTB2.WOUTPUT.W1.ST0101) (206) Therefore, people who play an important role can change them and **insert** [C1 level] subjects related to people life. (POSTB2.WOUTPUT.W1.ST0210) (207) By large and far, it is interesting to imagine how the school of the future will **look** [A2 level] like. (POSTB2.WOUTPUT.W1.ST0101) (208) In general, if the world keeps on **moving** [B1 level], there will be a revolution in education. (POSTB2.WOUTPUT.W1.ST0102) (209) Maybe the different people who nowadays play a role in school will be **replaced** [B1 level] for machines for saving time due to time is money. (POSTB2.WOUTPUT.W1.ST0208) (210) However, the actual subjects are going to **stay** [B1 level] but updated with some new information. (POSTB2.WOUTPUT.W1.ST0205) (211) As far as I know, I get the impression that teachers can become **substituted** [B2 level] by machines due to that actually, the

		machines are doing the work that was done by hand years ago. (POSTB2.WOUTPUT.W1.ST0201)
	(212)	It is often thought that schools will always be something traditional and will not **suffer** [B2 level] huge changes. (POSTB2.WOUTPUT.W1.ST0209)
	(213)	Nowadays, computers and technological instruments are used instead of them. It results hard and misunderstanding. For this reason, people will be **thrown** [B2 level] in at a deep end. (POSTB2.WOUTPUT.W1.ST0101)

According to the CAMD, a total of 12 topic-based MRW verbs (*accept* [193], *adapt* [194], *advance* [195], *come* [196], *progress* [199], *expand* [200], *expect* [201], *face* [202], *go* [205], *substitute* [211], *suffer* [212], and *throw* [213]) were used at the B2 level, constituting 48% of the overall topic-related usage. Additionally, it is worth noting that these participants also used two topic-based MRW verbs targeted at C1 level (*evolve* [199] and *insert* [206]) and one at C2 level (*follow* [203]).

Taken together, the experimental group's frequent use of topic-based MRW verbs at B2 level and above accounted for 60% of the total topic-related usage. The examples above clearly show that none of the topic-based MRW verbs used by the participants from the experimental group were explicitly prompted in the essay question.

Table 61 contains examples 214–221, showing the topic-based metaphors the control group employed in the written post-test. These L2 learners used 8 out of 45 (17.78%) topic-based MRW verbs (RF = 12.93%). For further details, see Table 122 in Appendix R.

Table 61

Topic-Related Metaphor Uses in the Written Discourse of Experimental Group Participants: Post-Test

Written prompt		Examples
Some people say that schools will be very different in the future. What do you think?	(214)	Taking all the above into consideration, I strongly believe that kids and teenagers will **experiment** a huge change in schools. (POSTB2.WOUTPUT.W1.ST0304)
	(215)	This is because of how new technologies are being adapted and how the use of paper and pen is **getting** old. (POSTB2.WOUTPUT.W1.ST0404)
	(216)	Taking all the above into consideration, I strongly feel that in the future schools will change and **give** more people the change to have a variety of subjects and skills useful for they future job and life. (POSTB2.WOUTPUT.W1.ST0306)
	(217)	Last but not least, new technologies should be mentioned because most people say that is the most important point for the change of

schools and the different changes that schools will **make**. (POSTB2.WOUTPUT.W1.ST0409)

(218) All things considered, the schools are going to change. I am sure because of the three points I already mention before and also because the technology is **progressing** so the schools that use it a lot too. (POSTB2.WOUTPUT.W1.ST0302)

(219) Moreover, I strongly believe that education is in need of a change and technology will help us to **reach** it. (POSTB2.WOUTPUT.W1.ST0401)

(220) First and foremost, it cannot be denied that teachers are an essential part of schools and maybe in the future they will be **replaced** by robots. (POSTB2.WOUTPUT.W1.ST0402)

(221) First and foremost, it cannot be denied the fact that teachers will be **suffering** these changes. (POSTB2.WOUTPUT.W1.ST0401)

The above examples indicate that the control group demonstrated limited use of topic-based MRW verbs at B2 level, with only four of them used at this level (*experiment* [214], *progress* [218], *reach* [219] and *suffer* [221]), accounting for 50% of their more limited overall topic-related usage, according to the CAMD. Further analysis revealed that among the five specific MRW verbs (*get*, *make*, *progress*, *replace*, and *suffer*) used by both groups in their post-test written essays, only the MRW verbs *progress* and *suffer* were classified as B2-level items. Yet, similar to the experimental group, none of the topic-based MRW verbs used by the participants from the control group were explicitly included in the essay question.

An important observation from the study is that the experimental group exhibited a greater range and more frequent use of topic-based MRW verbs, particularly at B2 level, in their written essays compared to the control group. While neither group relied on MRW verbs that were explicitly prompted in the essay instructions, what sets the experimental group apart is their ability to incorporate more sophisticated MRW verbs targeted at C1 and C2 levels into their written essays. This not only indicates a higher proficiency in vocabulary usage but also suggests a deeper understanding and application of language nuances.

This finding aligns with Semino's (2008) observation that "[m]etaphors can make topic clearer, more accessible, and easier to imagine and remember" (p. 148), thereby enhancing our understanding of metaphors as topic facilitators in L2 written productions. Metaphor-mediated instruction not only assists L2 learners in contextual metaphor usage but also

supports them in addressing the set topics in written tasks, a crucial skill for demonstrating L2 competence in high-stakes ESOL examinations.

The study's findings highlight the substantial benefits of the CL-oriented instructional intervention in enhancing the experimental group's use of metaphors in their written discourse, particularly concerning the set topic. The qualitative analysis revealed the experimental group's nuanced understanding and sophisticated use of metaphors, contributing to the richness of their written expression. However, it is important to note that these qualitative findings do not fully align with the quantitative results, as the enhanced metaphor usage did not translate into a significantly greater improvement in *Writing* proficiency at the B2 level, as assessed by the *B2 First for Schools* qualification.

The examples from the experimental group, particularly noteworthy, were not free from inaccuracies. They demonstrated unconventional uses of metaphor, as illustrated by example (198). As previously explained, the phrase "thrown in at the deep end" is commonly used to describe being put into a new and challenging situation without adequate preparation. However, in the context of discussing the use of computers and technological instruments, this metaphor does not align well and implies a sudden, challenging shift to technology. This implication is not clearly articulated in the surrounding text, and the more conventional expression is "thrown in the deep end", making the inclusion of *at* grammatically unusual and contributing to its unconventional usage.

This situation raises questions about the double-edged nature of fostering metaphor awareness. While it can deepen vocabulary knowledge, it may also lead learners to use unconventional expressions that examiners might find inappropriate or incorrect. At the B2 level, there is an expectation for learners to use a range of vocabulary accurately and contextually. Although encouraging metaphorical language use can be beneficial, it must be appropriate and grammatically correct. The unconventional and slightly incorrect use of the phrase in this case fails to meet these standards.

The findings suggest that unconventional metaphor usage, particularly when enhanced by CL-inspired approaches, can result in non-standard language production. This

raises concerns about its impact on achieving L2 proficiency, especially in standard L2 competence exams. It highlights the importance of balancing unconventional expressions with mastery of standard language forms in L2 learning.

These findings underscore the critical need to balance unconventional metaphor usage with the mastery of standard language forms in L2 learning. This balance is especially pivotal when aiming for proficiency that is both linguistically fluent and contextually appropriate. The examples from the experimental group, while showing metaphor awareness, often veered into unconventional metaphor usage, questioning their appropriateness in high-stakes ESOL exams. In contrast, the control group, which did not receive metaphor-mediated instruction, exhibited significantly higher proficiency levels and progress in their *Writing* performance, including vocabulary use, compared to the experimental group. This unexpected outcome suggests that other factors may have influenced the overall *Writing* performance assessment of both groups, warranting further investigation.

5.3. EVALUATING METAPHOR AT B2 LEVEL: RELATIONSHIP BETWEEN METAPHOR USE AND ASSESSMENT OF ENGLISH LANGUAGE PROFICIENCY (RQ3): KEY FINDINGS

5.3.1. Linking Metaphor Use to *Speaking* Proficiency Assessment at B2 Level

The current research findings shed some light on the relationship between metaphor usage and B2 level *Speaking* proficiency, particularly focusing on the Part 4 *Speaking* task of the *B2 First for Schools* examination. The results indicate that participants who incorporated more metaphors into their spoken responses tended to achieve higher vocabulary scores. This positive correlation suggests that metaphor use can play a key role in enhancing oral lexical knowledge in high-stakes language assessments.

While the findings indicating a link between metaphor usage and vocabulary scores in B2 level *Speaking* proficiency are encouraging, they must be interpreted with caution. The

identified correlation was statistically weak and non-significant, suggesting a potential association but not definitively confirmed by the data. It is crucial to recognise that correlation does not imply causation. This relationship between metaphor usage and vocabulary scores might be influenced more by specific task demands that prompt metaphor usage, rather than a natural integration of metaphors in speech, as discussed in Section 5.1 of this chapter.

The study's findings offer insights into the complex relationship between metaphor usage and overall *Speaking* proficiency. Contrary to expectations, increased use of metaphorical language does not necessarily correlate with improved quality in spoken production. This finding is particularly intriguing given the pervasive role of metaphor in the language system (Lakoff & Johnson, 1980; Steen et al., 2010) and its relevance to all aspects of communicative competence (Littlemore et al., 2006a), including its known contributions to vocabulary building and enhancing cultural awareness (MacArthur, 2010, 2016c). Indeed, the MC construct, as proposed by Littlemore and Low (2006a), has been empirically tested with theory-driven, valid, and reliable instrumentation by O'Reilly and Marsden (2021, 2023).

However, the results suggests that in some instances, overall *Speaking* proficiency might diminish with more frequent use of metaphors. This could be attributed to various factors, such as the overuse or misapplication of metaphors, which may complicate the clarity of speech rather than enhance it.

It is crucial to remember that correlation does not establish causation. The observed negative correlation does not necessarily imply that higher metaphor usage directly diminishes *Speaking* proficiency, as there could be other external factors influencing this relationship, especially given the lack of statistical significance in the findings.

Although metaphorical language use can enrich the depth of vocabulary knowledge, as indicated by this study (see Section 5.1.2), the non-standard use of metaphors might overshadow other critical aspects of *Speaking* proficiency, such as clarity, coherence, and the appropriate use of vocabulary and structures.

In this study, the experimental group used unconventional metaphors less frequently than the control group, with rates of 3.94% (*SD* = 4.37) and 5.98% (*SD* = 10.60), respectively.

However, this difference was not statistically significant (W[187.50]) = 0.000, p = .731). It should be noted that this study did not delve deeply into the nuances of unconventional metaphor usage. Instead, it followed MacArthur's (2019) approach, classifying metaphors as either conventional or unconventional, as detailed in Section 3.3 of Chapter 3.

To better understand whether the advantages of metaphor usage align with the *Speaking* task demands in the *B2 First for Schools* examination, it is essential to revisit the findings discussed earlier in this study. The experimental group, after receiving metaphor-mediated instruction during their B2 training, showed enhanced metaphor performance in oral discourse. This enhancement was evident not only in their increased usage of metaphors (RQ1a) but also in the depth of their vocabulary knowledge (RQ1b). Consequently, their improved *Speaking* performance (RQ2), particularly marked by enhanced vocabulary production and infrequent use of B2 level topic-based metaphors, seems to stem from their heightened metaphor awareness.

A closer examination of the assessment criteria, as outlined in Table 74 in Appendix I, reveals that factors beyond mere vocabulary are crucial in evaluating oral skills. These criteria emphasise the importance of using varied and relevant vocabulary to express and exchange ideas smoothly on familiar topics, focusing more on sustained fluency and the effective exchange of ideas rather than pinpointing errors. It appears that assessments prioritise consistent speech flow, allowing for some leniency towards occasional vocabulary errors, including unconventional uses. This suggests that deviating from standard vocabulary might not significantly hinder L2 learners' oral skills. Such an approach, emphasising fluency and task achievement over strict adherence to standard vocabulary usage, might contribute to the observed negative correlation between metaphor usage and *Speaking* proficiency.

The higher grades achieved by the experimental group in oral skills can be largely attributed to their improved quality of spoken production, particularly notable in their enhanced fluency and performance in aspects beyond vocabulary. While the use of metaphorical language likely played a role in their overall *Speaking* performance, it was not the only factor contributing to their success. This belief is supported by the fact that the experimental group

outperformed the control group across all five *Speaking* assessment descriptors by the end of the study, showing a comprehensive improvement in various facets of oral skills.

The assessment criteria, which allow for flexibility in spoken language use, present a challenge to the observed positive correlation between metaphor usage and B2 vocabulary production scores. Depth of vocabulary is not explicitly highlighted in the *Speaking* assessment descriptors (see Table 76 in Appendix I). This suggests that improvements in spoken quality might be more attributable to L2 learners' adaptive vocabulary use in response to task demands, rather than an increased use of metaphorical language. It is conceivable that the use of high frequent lexical items, which may also serve metaphorical functions in specific contexts, could have enhanced the learners' depth of vocabulary knowledge, thereby contributing to improved oral fluency.

In a similar vein, as discussed throughout this chapter, the use of specific vocabulary or fixed expressions that include MRWs may have facilitated fluency and ease of oral expression, thereby aiding L2 learners in more effectively expressing their thoughts. However, while reliance on prefabricated metaphorical expressions featuring MRWs can offer advantages for task completion, such usage does not inherently demonstrate a deep understanding of metaphors or adeptness in their contextual application.

The frequent use of topic-based metaphors may not significantly contribute to enhanced *Speaking* proficiency. For example, the use of the MRW noun *interest* in example (222) demonstrates that extended word meanings are often not necessarily sophisticated. Instead, they tend to involve more highly frequent lexical items, which may align more closely with lower CEFR levels. This suggests that reliance on such metaphors might not elevate the level of proficiency as much as anticipated.

(222) EXP_ST0109: <*{Do you and your friends spend a lot of your time doing outdoor activities?}*> <POSTB2.ST0109> when I am with my friends we usually go to the park and we eat some candies that we have bought in shops previously and we talk with

each other about our things our **interests** [B1 level] and we usually go to the cinema too for watching films […] </POSTB2.ST0109>

While the CAMD categorises the extended meaning of *interest* ("an activity that you enjoy doing when you are not working" [MM3]) at the B1 level, its basic sense ("a connection with something that influences your attitude or behaviour because you can gain an advantage from it" [MM6]) is classified at the C1 level. Incorporating *interest* with its extended meaning in *Speaking* tasks might enhance performance by adequately addressing the provided topic and contributing to oral fluency. However, such usage does not necessarily reflect L2 learners' mastery of advanced vocabulary, given its lower CEFR categorisation.

The findings from this study suggest that achieving overall proficiency in *Speaking* is not directly correlated with learners' fluency in metaphor usage. While MC and general language proficiency have been linked, as demonstrated in the research by O'Reilly and Marsden (2023), the strength of this relationship can vary based on several factors. These include the specific MC constructs and measures targeted (this study focused solely on production), the individual proficiency levels of the learners (with not all students reaching the B2 level), and the confounding effects of vocabulary knowledge, particularly given the emphasis on vocabulary in assessment criteria.

In their research, O'Reilly and Marsden (2023) explored the relationships between specific MC construct measures and high-stakes proficiency measures, focusing on the OOPT and the IELTS tests. They discovered a clear association between vocabulary depth and MC, although the precise role of MC within higher-stakes testing environments remained somewhat ambiguous. Despite differences in research focus, instruments, and participants, my analysis of the *B2 First for Schools* examination has revealed similar insights. Specifically, it highlighted the complexities of overall *Speaking* proficiency and demonstrated a more pronounced relationship with vocabulary knowledge.

5.3.2. Linking Metaphor Use to *Writing* Proficiency Assessment at B2 Level

The results of this study shed some light on the relationship between metaphor usage and *Writing* proficiency, specifically within the context of the Part 1 task of the *B2 First for Schools* examination. A notable finding is that L2 learners who more frequently incorporated metaphorical language into their written essays tended to achieve higher scores. This trend was observed not only in overall *Writing* performance but was particularly pronounced in the area of vocabulary production. The non-linear nature of these positive correlations suggests that metaphor usage varied among participants, indicating different levels of metaphor awareness within the cohorts.

These findings highlight the value of metaphors in enhancing the perceived quality of written essays, with a notable impact on lexical content. However, the contribution of metaphors to overall *Writing* performance is more complex, as evidenced by a weak, non-significant correlation. A potential concern is the over-reliance on metaphors, particularly through the frequent use of specific vocabulary or fixed expressions containing MRWs, as detailed in Section 5.1.2 of this chapter. Such reliance on pre-fabricated metaphorical expressions might detract from the naturalness of learners' expressions, potentially leading to penalties in formal assessment.

The study found a moderate, yet statistically significant, correlation between the use of metaphors and vocabulary, suggesting that metaphors could play an integral role in enriching lexical knowledge. However, it is vital to emphasise that correlation does not imply causation. External factors, such as the individual writing styles of learners and their familiarity with specific task demands, could also significantly influence these observed relationships.

To fully understand the alignment of learning gains from metaphor use with the *Writing* task demands of the *B2 First for Schools* examination, it is crucial to revisit the outcomes of the CL-oriented instructional intervention. The experimental group, trained with CL-inspired pedagogical practices, exhibited heightened metaphor awareness in their written essays. This was evidenced by their increased metaphor usage (RQ1a) and enhanced depth of vocabulary knowledge (RQ1b). However, despite their consistent use of topic-based metaphors at B2

level and above, this advantage did not translate into a significant improvement in overall *Writing* performance. In contrast, the control group, despite demonstrating weaker metaphor performance, showed greater proficiency in *Writing*. Notably, they achieved excellence in vocabulary use, even without the frequent use of B2 level topic-based metaphors.

These findings underscore the potential influence of specific evaluation benchmarks in the *B2 First for Schools* examination, where there is a notable emphasis on vocabulary use in assessing written skills (see Table 75 in Appendix I). L2 learners are evaluated based on their ability to employ a wide range of vocabulary, particularly less frequent words and phrases that are well-suited to the given topic. While the assessment criteria allow for occasional errors, they emphasise that these should not impede understanding, seemingly favoring conventional language use.

This implies that deviating from standard language patterns, especially in the use of metaphorical expressions, could be disadvantageous for L2 learners in their *Writing* assessments. Supporting this notion, Littlemore et al. (2014) noted that at the B2 level, "learners are more likely to make more errors when using metaphor than when using other types of language" (p. 139), highlighting the risk associated with metaphor usage in high-stakes language evaluations.

The experimental group used unconventional metaphors in their writing somewhat more frequently than the control group, with rates of 7.04% (*SD* = 6.64) as opposed to 4.12% (*SD* = 2.61). However, this difference was not statistically significant (W[162] = 1.015, p = .308). As a result, the experimental group, despite taking more linguistic risks by using unconventional metaphors, might not receive any reward for this approach, and could be potentially unrewarded or even face penalties in their evaluations.

The unexpectedly superior performance of the control group in *Writing* raises questions about potential inconsistencies in the evaluation process. The influence of subjective assessments, especially by L1 speaker examiners, needs careful consideration. This subjectivity may lead to biases, causing assessors to view unconventional metaphors

from L2 learners as errors rather than valid linguistic alternatives. Yet, as discussed earlier, it should also be acknowledged that errors were still identified in idiom usage, for example.

Gibbs (2016) points out that mixed metaphors, often seen as indicative of poor writing or speaking, do not necessarily reflect cognitive errors nor impede understanding (p. ix). This could partly explain why the experimental group, despite displaying richer vocabulary depth and lesser reliance on pre-fabricated language, did not outperform the control group in *Writing*.

This ambiguity extends to L2 teaching practices. Instructors lack clear guidelines on how to support learners in using both conventional metaphors and more creative or "hybrid metaphors" that mix linguistic and conceptual systems (MacArthur, 2016c). Therefore, L2 instructors are left to navigate whether to prioritise communicative effectiveness (MacArthur, 2016c) or conventional linguistic form (Philip, 2010) in their assessment of metaphorical language, a decision that remains challenging in the absence of clear guidance.

Consider example (223), where the experimental group's student ST2010 used the verb *insert* metaphorically in an unconventional manner to convey the act of introducing something new into a specific context. More commonly, verbs such as *introduce*, *include*, *add*, or *incorporate* are used for this purpose. However, *insert*, defined as "to put something into something else, or into a hole or space" (MM1), also embodies the idea of physically placing one thing within another, similar to the previously mentioned verbs. The deliberate choice of *insert* in this context suggests that the learner is not merely replicating familiar expressions. Instead, it indicates an active engagement with the language, as the learner draws from a broader understanding of English to creatively construct sentences.

(223) EXP_ST0210: Secondly, subjects are also an important point to consider because some of them can open a can of worms, because of their topic. Therefore, people who play an important role can change them and **insert** subjects related to people life. (POST.WOUTPUT.ST0210)

The choice to use *insert* in an unconventional metaphorical context reveals heightened metaphor awareness, yet it also underscores the complex balance between such awareness

and adherence to standard language use. This deviation sheds light on the challenges that L2 learners face in understanding and employing established expressions.

These insights align with the arguments of Littlemore et al. (2014), who advocate for increased leniency from L2 assessors towards deviations from conventional L1 language patterns. This is particularly important considering the developmental phase of L2 learners, especially at the B2 level.

Piquer-Píriz (2021) further emphasises this point, stating that "L2 learners need to be aware of the normative conventions that regulate figurative uses in the target language, but, at the same time, […] it would also seem worth exploring how [they] make use of their rich linguistic repertoires through their figurative ability to communicate effectively" (p. 193). This highlights the critical need to encourage L2 learners to experiment with language, recognising its importance not just in their learning and communication process, but also in assessment contexts. However, the findings of the present study indicate that this experimentation might not be sufficiently recognised or valued in the way L2 proficiency is currently evaluated aligning with standard descriptors outlined in the CEFR, specifically at B2 level. This discrepancy suggests a potential gap between the pedagogical approaches in L2 learning and the criteria applied in standard L2 proficiency assessments.

It is important to note that the reasons behind the use of these unconventional metaphors remain unclear. They could stem from various sources, such as L1 influence or transfer/calque, creative language use, rote learning, potential mistranslations, or other factors. Further research is crucial to delve into the motivations for these unconventional metaphor choices. Specifically, it would be valuable to investigate whether they are influenced by CL-oriented instructional interventions and to evaluate their impact on L2 learner assessments conducted by L1 speakers.

Boers (2004) explored how L1 English speakers react to unconventional metaphorical expressions derived from established CMs in writing, comparing this with their responses to non-standard idiomatic expressions. His pilot study showed that L1 speakers are more tolerant of unconventional metaphorical expressions than of deviant idioms. In contrast, the current

study indicates a possible trend of reduced tolerance among L1 examiners towards unconventional metaphors, especially when assessing written organic metaphorical uses by L2 learners in testing environments.

The findings underscore the potential importance of conventional language chunks in essay writing. The control group's superior performance could be attributed to their familiarity with their use of specific vocabulary or fixed expressions containing MRWs. Employing these pre-fabricated metaphorical expressions might have improved the quality of their written work, allowing them to more effectively meet task conventions and articulate their ideas with greater clarity. However, this enhanced written quality might stem more from the use of these language chunks, which happen to contain MRWs, rather than from a profound understanding of the metaphors themselves. Given that these conventional language chunks may not fully represent the learners' authentic linguistic capabilities, this raises intriguing questions about the true relationship between metaphor usage and *Writing* proficiency.

The control group's higher scores in *Writing* proficiency can be largely attributed to the enhanced quality of their written production, particularly evident in their cohesive writing flow and in areas beyond mere vocabulary. Although metaphorical language might have positively contributed to their *Writing* performance, it was not the sole key to success. While two L2 learners might employ metaphors to a similar extent, the way these incorporate the metaphors into their essays can vary, reflecting their individual writing styles. This observation is bolstered by the control group's superior performance across all four *Writing* assessment descriptors.

The findings underscore that the quality of written outputs is largely influenced by the strategic use of vocabulary, tailored to the specific demands of the task. This highlights the importance of task-appropriateness, as outlined in the assessment criteria, which places value on the use of varied and less frequent, context-specific vocabulary.

Despite the experimental group's proficiency in employing topic-based metaphors, particularly at the B2 level and above, and their demonstration of a wide range of metaphorical expressions, these sophisticated elements of their writing seemed to go unnoticed in the

evaluation process. The focus in assessment appears to be more on the breadth of lexical knowledge, with less emphasis on depth, such as the nuanced use of metaphors.

This creates a notable disparity: the experimental group's advanced metaphorical proficiency did not result in higher scores for writing quality. This suggests that the assessment criteria might either overlook the complexities of metaphor usage or consider such nuances as non-essential to the overall perceived quality of writing. Additionally, it raises the possibility that the integration of sophisticated metaphor usage with grammatical errors may have negatively impacted their scores. This potential interplay between high-level metaphorical language and grammatical accuracy warrants further exploration to understand its influence on *Writing* assessments.

The findings suggest that employing topic-based metaphors does not automatically lead to improved *Writing* skills. While these metaphors can enhance written performance by closely addressing the set topic and aiding in task completion, their association with lower CEFR levels may not accurately reflect an L2 learner's mastery of advanced vocabulary, as detailed in Section 5.3.1 of this chapter. Furthermore, the use of topic-based metaphors at the B2 level or higher may not receive adequate recognition in assessments, given their subtle incorporation within the established criteria.

A central issue in this context is the significance placed on metaphor usage in L2 competence assessments by major examination boards. This raises critical questions about the validity of such assessments in high-stakes exams, particularly concerning whether they accurately measure essential components of L2 competence (Littlemore & Low, 2006a, 2006b). Future research should further explore whether these assessments truly reflect meaningful L2 proficiency or overlook these key aspects. It is essential to ascertain if this investment in instructional time to metaphor teaching aligns with the actual learning outcomes and contributes significantly to the overall development of L2 proficiency.

CHAPTER 6

CONCLUSIONS

In this doctoral thesis, I have explored the long-term effects of integrating CL-oriented distributed learning of metaphor into an EFL syllabus, specifically aligned with the CEFR B2 level descriptors. Focusing on L2 learners of English receiving private tuition at a language school to achieve B2 level, this quasi-experimental study aimed to investigate how raising metaphor awareness in topic-based teaching through regular instruction by a CL-trained EFL teacher can assist their use of metaphor in speech and writing as their language skills develop.

The research tracked the evolution of organic metaphor usage, in terms of frequency and variety, among Spanish secondary school students preparing for the CEFR B2 level. By exploiting the semantic potential of topic-based vocabulary from the goal-oriented textbook used in class as the primary source of input, the CL-oriented instructional intervention sought to enhance L2 learners' conventional use of metaphors. This approach was integral to a broader effort to foster vocabulary growth and improve their English language proficiency.

Additionally, considering the importance of topic similarity in preparing for high-stakes ESOL examinations, the study aimed to determine whether learning gains from metaphor-mediated instruction could lead to enhanced production quality in various real-life oral and written communication contexts, as assessed by the *B2 First for Schools* qualification.

6.1. SUMMARY OF KEY FINDINGS AND CONTRIBUTIONS OF THE STUDY

This study has attempted to broaden our understanding of L2 metaphor production, providing some insights for EFL instructional and formal assessment contexts at the B2 level. It identifies distinct patterns of metaphor use, both in spoken and written forms, emerging from metaphor-mediated instruction during B2 training, and elucidates how learning outcomes from these CL-inspired pedagogical techniques are assessed in high-stakes ESOL examinations.

The current research has shown that metaphor-mediated instruction can be beneficial yet also very challenging. Employing a longitudinal approach to raise metaphor awareness using CL-oriented methods alongside the development of language skills has successfully increased L2 learners' confidence in using metaphorical language organically, extending beyond idioms and phrasal verbs. This reduced hesitance seems to significantly enhance the frequency and diversity of metaphor usage, marked by a higher density and broader range of metaphorical expressions in both open-class and closed-class lexical items. These findings highlight that CL-oriented teaching methods are effective not just in teaching but also in actively encouraging the use of metaphors among L2 learners for expressive purposes in real-life communication contexts, particularly supporting topic discussion.

This study expands on the existing understanding of the pervasive role of metaphors in English learner discourse, delving into their usage among Spanish secondary-school students. While the findings confirm that metaphors are a frequent linguistic feature in L2 speech, they are even more dominant in L2 writing. The research's ecological validity is enhanced by analysing naturally produced learner discourse in controlled environments that replicate standard testing conditions, aligning with the EFL syllabus and the CEFR assessment descriptors. By also ensuring consistency with pedagogical practices, this approach mirrors real-world language use in specific tasks and helps to bridge the gap between CL-oriented approaches and assessment methods for measuring learning gains. Furthermore, the slight adaption of MIPVU (L2-MIPVU), ensuring replicability, acknowledges the heterogeneity in L2 language use when analysing learner discourse across various communicative contexts.

This study provides some insights into how heightened awareness of metaphors can lead to their more frequent and consistent use in speech and writing. Yet, it also shows that metaphor usage in L2 discourse may not be always a consequence of the teaching-learning process as it could be also quite independent of it. The effectiveness of metaphor-mediated instruction in L2 learning appears to be influenced by several factors. These include the specific nature of discourse, task requirements, and input texts from the mainstream textbook

used in class for B2 training, without necessarily reflecting a deeper understanding of metaphors and the ability to employ them contextually.

The study's findings reveal that L2 learners naturally improve their metaphor usage, especially in written discourse, which benefits from more planning and the use of pre-fabricated metaphorical expressions. In contrast, the less planned and spontaneous nature of oral discourse may require support from explicit metaphor teaching. Metaphor-mediated instruction has been particularly effective in enhancing spoken metaphor production, helping L2 learners cope with the communicative challenges of oral discourse in testing environments.

The delayed post-test demonstrates the sustained effectiveness of distributed teaching of metaphor, especially when integrated with a diverse range of CL-oriented methods in the EFL syllabus for regular L2 classroom practices. This holistic approach to metaphor instruction enabled learners to apply metaphorical language in different contexts, thereby enhancing their learning experience beyond the scope of vocabulary presented in the mainstream materials. However, the varying degrees of benefit observed among learners from these CL-oriented approaches highlight individual differences in response to explicit metaphor instruction, as has been shown in the case of participant ST0105 from the experimental group.

My dual role as an EFL instructor and a researcher served as a two-way bridge, merging key findings from CL with the realities of the L2 classroom and providing insights into learner performance from the actual practitioner's view. The experimentation at B2 level noted by Littlemore et al. (2014), and confirmed in the present study, appears resistant to the support offered by metaphor-mediated instruction. Although they tend to use metaphors in a judicious way extending their usage across diverse contexts when these are encountered in different topics and underling overlapping CMs, additional corrective feedback and direct teaching of metaphors are essential to help L2 learners finetune their oral and written productions.

While a five-month CL-oriented instructional period may be sufficient to yield long-term gains in vocabulary growth across various word classes, this study reveals that it does not necessarily guarantee effective metaphor usage in terms of form and meanings. CL-inspired classroom activities enhance metaphor awareness and learner autonomy to explore

alternatives to pre-fabricated metaphorical expressions. However, its effectiveness in fostering metaphor usage according to conventional standards appears limited. The findings suggest that metaphor-mediated instruction using CL-inspired approaches may not effectively address unconventional metaphor uses, despite the gain in depth of vocabulary knowledge.

The study sheds light on the inherent challenges of CL-oriented methods, particularly their focus on the conceptual dimension of metaphors, which fosters deeper cognitive engagement. It underscores the need to develop more effective pedagogical strategies for enhancing accuracy in using metaphor. Towards this end, explicit metaphor instruction should adopt a dynamic and holistic approach. This involves integrating linguistic elements, such as lexico-grammatical patterns, with pragmatic aspects such as contextual usage, in methods designed to raise metaphor awareness.

In my view, a key consideration is to balance unconventional language use and mastery of standard language forms, especially when aiming for fluency and contextual appropriateness in the L2. Adherence to conventional language norms is often crucial for mastery, particularly in formal training that targets native-like linguistic accuracy and testing environments, where deviations from these norms might be deemed inappropriate.

The current study demonstrates that metaphor-mediated instruction can significantly enhance L2 learners' chances of success in high-stakes English ESOL examinations at the B2 level. This is particularly evident in *Speaking* proficiency, with a notable improvement in vocabulary knowledge. However, this study also reveals that alternative methods can particularly support L2 learners in improving their *Writing* proficiency, without necessarily relying on CL-inspired instruction.

From a qualitative perspective, this research highlights the benefits of CL-oriented distributed learning of metaphor in equipping L2 learners with the tools to integrate topic-based metaphors, targeted at B2 level and beyond, with a particular emphasis on enhancing their written outputs. This approach provides some insights into metaphors as topic facilitators, aiding learners in effectively addressing the set topics in exam tasks, a crucial skill for demonstrating proficiency in high-stakes ESOL examinations.

However, the findings suggest that merely employing topic-based metaphors does not automatically lead to improved oral and written skills. Instead, the use of prefabricated metaphorical expressions, tailored to the task requirements, and unconventional uses of metaphor might play a crucial role in L2 learners' proficiency. In writing, precision in form is prioritised, contrasting with the emphasis on fluency in speech. Hence, this variation in impact across different language skills may be more attributed to the specific assessment criteria than to the differential effect of the teaching approach.

This study sheds some light on the relationship between metaphor usage and the perceived quality of learner production in standard L2 competence exams, specifically at the B2 level. It demonstrates that while metaphor-mediated instruction enhances metaphor usage in both *Speaking* and *Writing*, such learning gains do not necessarily translate to higher scores. These findings suggest that previous research has perhaps been overly optimistic in promoting the benefits of raising metaphor awareness in relation to overall communicative competence.

Metaphors, although enriching the vocabulary repertoire, are not the sole predictor of success in high-stakes ESOL examinations. In the *B2 First for Schools* examination, other vital linguistic skills take precedence. These include sustained fluency, an individual writing style, and effective task completion, all aligned with the can-do statements at the B2 level of the CEFR descriptors. This research further suggests that a strategic use of task-specific vocabulary is crucial for broadening vocabulary scope. Conversely, employing unconventional metaphors may hinder clarity and coherence in learner discourse, potentially impacting the appropriate use of vocabulary.

An important insight from this research is the potential misalignment between the integral role of metaphor in L2 learning and its assessment in high-stakes ESOL examinations. While metaphors display both vocabulary depth and cognitive engagement, they might be undervalued or overlooked in standard testing environments. The current assessment criteria, while rewarding lexical breadth, especially less frequent and relevant words, do not explicitly acknowledge vocabulary depth in the descriptors. Indeed, this study shows that metaphors do

not necessarily reflect L2 learners' mastery of advanced vocabulary. This could imply that the varied usage of topic-based metaphors often goes unnoticed, even when directly addressing the set topic. These oversights may inadvertently diminish the centrality of metaphor in learner discourse, undervaluing its significance in assessing the quality of spoken and written outputs.

This study offers some insights into the types of metaphors that are either rewarded or simply tolerated in real-world testing environments. It reveals that deviations from standard language patterns may not be advantageous. L2 learners' linguistic risks might not only go unnoticed but could also be penalised in assessments, often being interpreted as errors rather than valid linguistic alternatives. This suggests a potentially limited degree of tolerance by L1 speakers towards unconventional metaphorical expressions. Therefore, experimentation with language, which could demonstrate advanced proficiency, might not be sufficiently recognised or valued in current assessments of L2 proficiency.

A key issue is whether major examination boards consider metaphor usage a vital aspect of assessing L2 competence. This brings into question the legitimacy of evaluating metaphor use in high-stakes ESOL exams and whether it effectively reflects a significant and required element of L2 proficiency. Additionally, it requires a reflection on the value of devoting instructional time to metaphor in EFL contexts, to determine whether the efforts invested in this area of language instruction are warranted by the resulting learning outcomes.

6.2. IMPLICATIONS AND APPLICATIONS

This PhD dissertation provides some insights for both L2 instruction and assessment in relation to metaphor usage at B2 level. The following subsections suggest some pedagogical applications and assessment considerations essential for EFL instructors, L2 learners, ESOL examiners and curriculum designers.

6.2.1. Pedagogical Applications for the EFL Classroom

Metaphor-mediated instruction provides a valuable framework for fostering metaphor usage in real-world contexts. This study shows that bringing CL to the real L2 classroom to

enhance metaphor production among L2 learners can be viewed by EFL instructors as doable, despite practical constraints such as limited time in exam training.

A key insight from this research is the value of recycling metaphors that L2 learners encounter across various themes in topic-based instruction, which facilitates their application in novel contexts. Vocabulary linked to overlapping CMs is mastered more effectively in terms of form, meaning, and usage in both spoken and written productions.

Lexical organisation along CMs could be advantageous to develop CL-inspired activities that target overlapping CMs relevant to exam topics, especially abstract ones, given the role of topic similarity in demonstrating L2 competence in higher-stakes ESOL examinations. Boers (2004) suggests that CMs can provide a framework for the integration of knowledge in L2 learning. This means that topic-based expressions originating from a common source domain can be grouped together, considering that source domains are broad providing a rich source of mappings (Littlemore & Low, 2006a). Presenting these to L2 learners as belonging to the same category might help provide structure and organisation in their approach to vocabulary learning.

For instance, as shown in this study, L2 learners can be introduced to the CM TIME IS MONEY to equip them with language that may inform the metaphors learners are likely to use when talking or writing about “shopping” (e.g., *shop for the best deal*) and “leisure” (e.g., *spend time with friends*). Furthermore, this can also inform discussions about “education and work” in terms of productivity (e.g., *budget your time*), opportunity costs (e.g., *time cost*), prioritisation (e.g., *best return on your time*), or time management (e.g., *time blocks*).

This study has shown that presenting L2 learners with CMs has great appeal, not only with those who are adult and/or are at advanced levels (Low, 2008) but also with secondary school students. Yet, the research also spotlights potential hurdles in enhancing metaphor awareness through the CMA. A possible alternative might be a shift to a bottom-up approach in lieu of the top-down approach used in this study. This method would focus on the narrower theme highlighted by the lower-level mappings instead of the broader themes highlighted by the higher-level mapping. To illustrate, presenting the lower-level CMs HAPPINESS IS

VERTICALITY and HEALTH IS VERTICALITY concurrently with the higher-level CMs GOOD IS UP and BAD IS DOWN could enhance learners' understanding of metaphors.

Although metaphor-mediated instruction enriches the lexical range of L2 learners, they might greatly benefit from extended practice, particularly in spoken production. Implementing CL-inspired activities that offer guided practice and corrective feedback on metaphorical language use in oral discourse could be beneficial. A potential instructional approach might involve distinguishing between teacher-driven and learner-driven CL-inspired activities. The former encompasses direct instructional support, while the latter champions learner autonomy under the instructor guidance. Such a balance could foster a more supportive environment, gradually steering learners towards independent mastery in metaphor usage.

6.2.2. Implications for Formal L2 Assessment and Syllabus Design

This study presents empirical evidence underscoring the lack of detailed description of metaphor-related language skills within the CEFR descriptors (Council of Europe, 2020). Understanding how metaphorical language is formally assessed in standard L2 competence exams is crucial. Knowing which metaphors satisfy examiner expectations will better inform EFL instructors to enhance metaphor production among L2 learners, thus optimising training for high-stakes ESOL examinations. Without clear rubrics for evaluating metaphors in spoken and written outputs, raising metaphor awareness in instructed L2 settings may risk becoming "a hit or miss affair", as pointed out by MacArthur (2021, p. 356). Notably, this study highlights that a lack of pedagogical direction can lead to gaps in metaphor-related learning outcomes.

For metaphors to be effectively evaluated in standard L2 competence exams, its foundational concept should be clearly defined within the CEFR descriptors. As MacArthur (2021) notes, "[e]ducational processes are indissolubly linked to educational products" (p. 357). In this view, the metaphor-related descriptors for the CEFR levels proposed by Littlemore et al. (2014) are challenging to incorporate in training materials, primarily because they are not part of the marking criteria for high-stakes ESOL examinations.

Indeed, integrating these descriptors into the established CEFR standards for each level is problematic for assessment, as many ESOL examiners might not have the expertise in metaphor studies to fully understand them. Therefore, if metaphors are overlooked in standard L2 assessment and are not included in the curriculum, EFL instructors may hesitate to teach them. This raises questions about the value of fostering teacher awareness of the CL view of metaphor and investing classroom time in pedagogical practices that enhance metaphor usage among L2 learners. These concerns become more accentuated given the substantial investment of time and cognitive effort CL-inspired instruction requires, potentially daunting for regular EFL instructors.

Considering that training materials align with the syllabus design and exam materials, implementing specific but user-friendly guidelines on metaphor assessment in the CEFR descriptors would help bridge the research-practice-assessment gap. Such an inclusion would not only enrich mainstream materials but also enhance EFL teachers' understanding and appreciation of the utility and feasibility of CL-inspire instruction (see Llopis-García et al., 2022). While not all practitioners might be trained in CL, textbooks can provide them with effective strategies and lesson plans to teach metaphors, thereby boosting teachers' confidence in implementing CL-inspired activities into the L2 classroom.

Textbooks should incorporate opportunities for metaphorical thinking throughout their units, using metalanguage that is user-friendly for both instructors and learners. Emphasis should be placed on the most common words or phrases used in discourse, but also addressing both clarity and teachability of their metaphorical senses, as well as the practicality of CL-oriented teaching methods (MacArthur & Littlemore, 2008).

Focusing on specific vocabulary sets as learning targets (e.g., Brezina & Gablasova, 2015) can make the challenging task of learning metaphorical language more manageable for teachers and learners, a crucial consideration in time-limited language programmes. This approach could lead to a more standardised teaching method, reducing cognitive load and ensuring consistent engagement with useful, topic-based metaphors at their CEFR levels.

The textbook should serve as a vital ally in the teaching process, enhancing both the effectiveness of CL-inspired instruction and learners' proficiency in using metaphor. Raising metaphor awareness, especially regarding the conceptual domain, into classroom dynamics can enrich L2 instruction, as highlighted by this study. Teachers, however, need specifically designed activities to improve students' phraseology and pragmatic metaphor use. These activities should align with curricular objectives and criteria for formal evaluations, such as official exams, and be integral to the main teaching materials. This includes coverage of both conceptual and linguistic/pragmatic domains of metaphor. Facilitating this integration requires incorporating the concept of "metaphor" into major language descriptors like the CEFR, which would support the inclusion of CL-inspired instruction in syllabi and, therefore, in textbooks.

6.3. LIMITATIONS AND FUTURE DIRECTIONS

Several limitations to this study need to be acknowledged. First, the research primarily focused on metaphor production outcomes arising from a small-scale investigation. Therefore, the participant pool may not fully represent the broader L2 population, thus potentially limiting the applicability of the findings to L2 learners of varied ages or from different L1 backgrounds. However, employing the *First for Schools Cambridge English* qualification as a measure for oral and written production underscores the relevance and applicability of the findings to the specified cohort of B2 level learners. To achieve a more holistic understanding of the effects of metaphor-mediated instruction on L2 metaphor production, future research should broaden its scope and sample diversity. While an extended delayed post-test might have added depth to these findings, the challenges brought about by the COVID-19 pandemic made this difficult.

Given that the sample was drawn from small EFL classes within private tuition settings, caution is warranted when generalising these findings. The CL-inspired activities implemented in this research were tailored for a specific EFL context. Future studies would benefit from applying this methodology in a variety of instructed L2 settings, such as state schools. In such environments, the dynamics between EFL instructors and L2 learners might differ, especially

in larger classrooms. Yet, securing a participant sample as uniform as the one in this study might pose challenges in these alternative instructed L2 settings.

The experimental group's increased use of metaphors compared to the control group cannot solely be attributed to the metaphorical language introduced during the CL-oriented instructional intervention. It is possible that the experimental group may have learned additional metaphors in various contexts or from exposure to English discourse outside the English language school, such as in their secondary schools.

While this study's evaluation of B2 level in the *B2 First for Schools* examination aligns with the CEFR descriptors, it represents just one method to gauge L2 learner performance. Broader research across various CEFR levels and different standardised testing environments is crucial to gain comprehensive insights. Furthermore, future studies analysing metaphor usage in L2 discourse should strike a careful balance between topic variation and ecological validity in their measurement choices. Such methodological refinement could provide valuable insights into how topic selection influences metaphor production. Further research is needed to investigate the impact of varying task demands on L2 learners' metaphorical language use.

This investigation provides a limited perspective on MC. Although it delved deeply into metaphor production by examining the metaphorically used words within learner discourse, it primarily focused on metaphor density as an indicator of lexical *growth*. It did not evaluate the development of metaphor-related language skills as *outgrowth* of CL-inspired instruction, which might result in enhanced MC. Future studies should incorporate additional aspects such as creativity, appropriateness, or effective employment to refine the teaching approach, aiming at improving L2 learners' overall English language proficiency. Additionally, future research should explore the degree of tolerance by ESOL examiners (often L1 speakers) towards unconventional language in testing environments.

The similarities and discrepancies between the findings of this study and those by Littlemore et al. (2014) as well as Nacey (2020, 2022) suggest the need for further exploration of open-class vs. closed-class metaphors in learner discourse beyond just the CEFR B2 level. Future research should delve deeper into the role of input in L2 speech and writing. Such

exploration could be beneficial for developing CL-inspired activities tailored to the EFL syllabus, especially in providing support in the use of topic-based metaphors across various word classes. Aligning assessment with training input, ensuring task similarity, will offer additional insights into how topic similarity can enhance L2 metaphor production.

To gain a deeper understanding of the impact of metaphor-mediated instruction on L2 learners, further investigation into individual performance differences is necessary alongside group-level differences. Future studies should provide a detailed description of the teaching-learning experience, shedding light on the advantages and challenges of incorporating CL-oriented methods into the EFL syllabus. Feedback from both L2 learners and CL-trained EFL teachers will offer invaluable perspectives on understanding metaphor-mediated instruction.

6.4. CONCLUDING THOUGHTS

This thesis, while raising numerous questions, provides a foundation to view L2 metaphor production not just as an isolated skill but as a concept that benefits from long-term explicit instruction. Although L2 learners' use of metaphorical language is likely to develop independently as part of a growing lexicon, it might not always be meaningfully employed.

The observed improvements in metaphor production among L2 learners who did not receive metaphor-mediated instruction suggest that solely identifying metaphors in learner-produced text and relying strictly on quantitative results might not capture the full scope of improved production quality. Procedures such as MIPVU (Steen et al., 2010) provide insights into metaphor analysis in linguistic data. However, focusing on metaphor density as a primary indicator of growing metaphor usage provides just a glimpse of the broader metaphor production landscape. Indeed, metaphors identified by such procedures might not always be viewed or processed as MRWs by L2 learners (see Steen, 2023). This research indicates that factors such as task parameters and the adaptable nature of highly frequent lexical items largely influence L2 learners' increased metaphor usage. Rote learning of language patterns containing MRWs, while beneficial for task completion, does not necessarily reflect a deeper understanding or meaningful use of metaphorical language.

This study emphasises that the benefits of explicit metaphor instruction in L2 learners' language development extend beyond the scope of statistical measures. By raising metaphor awareness, L2 learners not only see a significant increase in their metaphor use — evidenced by density rates — but they also experience an expansion of their lexical diversity (e.g., deeper vocabulary), and thematic integration into both spoken and written production. This piece of research has shown that raising metaphor awareness prevents fossilisation of vocabulary development, particularly at B2 level.

As Boers (2022) argues, it is essential to weigh both statistical and pedagogical significance in research results. While statistical significance offers insights into overall trends and patterns, they may overlook the subtle intricacies of L2 learners' language skill development and their sophisticated deployment of metaphors. Recognising the divide between quantitative outcomes, which may sometimes be disappointing, and more encouraging qualitative insights can pave the way for a richer understanding of classroom-based research implications. Therefore, the evaluation of metaphorical language use should be approached holistically, considering more than just density rates, and always within its broader context.

Finally, the study prompts questions about the role of metaphor in the CEFR descriptors and its significance to formal L2 proficiency assessment. If high-stakes ESOL examinations truly aim to reflect real-world language use, then the integration of metaphor-related language skills is essential, particularly considering its fundamental role in overall communicative competence (Littlemore & Low, 2006a).

Without clear assessment criteria, the evaluation of metaphorical language could become overly dependent on examiner judgment, which might be different depending on whether the interlocutor is a native speaker or share the same L1 with the L2 learners. MacArthur (2016c) argues that using native speaker norms as a benchmark for metaphor production might be an unattainable objective for learners: "hybrid metaphors are an almost inevitable outcome of language contact, and emerge in the speech and writing of even highlight proficient users of English as an L2" (p. 133). However, this type of mixing metaphors

might be seen as an infelicitous metaphor use by L2 learners of English when judged against L1 speakers (Philip, 2010).

This study highlights that metaphor may be seen as a double-edge sword for L2 learners aiming to achieve proficiency, particularly regarding standard L2 assessment. The unconventional uses arising from enhanced metaphor awareness and deeper vocabulary knowledge, may add complexity to the teaching-learning process, learning outcomes, as well as the assessment practices. This study sets the stage for future investigations in this field and possible pedagogical applications of metaphor in the EFL classroom practice. This encompasses both instructional methods and teacher feedback, which can enhance learner performance in formal L2 assessment.

REFERENCES

Aas, H. L., & Nacey, S. (2019). Methodological concerns for investigating pause behaviour in spoken corpora. In L. Degand, G. Gilquin, L. Meurant & A. C. Simon (Eds.), *Fluency and disfluency across languages and language varieties* (pp. 41–64). Presses Universitaires de Louvain.

Achard, M. (2018). Teaching usage and concepts: Toward a cognitive pedagogical grammar. In A. Tyler, L. Huang & H. Jan (Eds.), *What is applied cognitive linguistics? Answers from current SLA research* (pp. 37–62). De Gruyter Mouton. https://doi.org/10.1515/9783110572186-002

Achard, M., & Niemeier, S. (Eds.). (2004). *Cognitive linguistics, second language acquisition, and foreign language teaching*. De Gruyter Mouton. https://doi.org/10.1515/9783110199857

Alejo-González, R. (2010). Making sense of phrasal verbs: A cognitive linguistics account. *AILA Review*, *23*(1), 50–71. https://doi.org/10.1075/aila.23.04ale

Alejo-González, R. (2022). Metaphor in the academic mentoring of international undergraduate students: The Erasmus experience. *Metaphor and Symbol*, *37*(1), 1–20. https://doi.org/10.1080/10926488.2021.1941969

Alejo-González, R. (2024). *Metaphor and corpus linguistics: Building and investigating an English as a medium of instruction corpus*. Routledge. https://doi.org/10.4324/9781003400905

Alejo-González, R., & García-Bermejo, V. (2020). The manage of two kingdoms must: An analysis of metaphor in two CLIL textbooks. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 241–262). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-012

Alejo-González, R., Piquer-Píriz, A. M., & Reveriego-Sierra, G. (2010). Phrasal verbs in EFL course books. In S. D. Knop, F. Boers & A. D. Rycker (Eds.), *Fostering language*

teaching efficiency through cognitive linguistics (pp. 59–78). De Gruyter Mouton. https://doi.org/10.1515/9783110245837.59

Alejo-González, R., Piquer-Píriz, A. M., Castellano-Risco, I., Martín-Gilete, M., Fielden-Burns, L., MacArthur, F., Nacey, S., Philips, G., Krennmayr, T., Coelho, M., Littlemore, J., & Ädel, A. (2021). The METCLIL corpus. v1. https://www.sketchengine.eu/metclil-corpus-of-metaphor-in-academic-talk/

Aleshtar, M. H., & Dowlatabadi, H. (2014). Metaphoric competence and language proficiency in the same boat. *Procedia – Social and Behavioral Sciences*, *98*, 1895–1904. https://doi.org/10.1016/j.sbspro.2014.03.620

Alonso-Aparicio, I., & Llopis-García, R. (2019). La didáctica de la oposición imperfecto/perfecto simple desde una perspectiva cognitiva. In I. Ibarretxe-Antuñano, T. Cadierno & A. Castañeda Castro (Eds.), *Lingüística cognitiva y español LE/L2* (pp. 274–299). Routledge.

Alsadi, H. Z. (2016). Metaphors production and comprehension by Qatari EFL learners: A cognitive approach. *Arab World English Journal (AWEJ)*, *7*(1), 408–421. https://dx.doi.org/10.24093/awej/vol7no1.24

Amaya-Chávez, E. (2010). The gaps to be filled: The (mis)treatment of the polysemous sense of hand, cool and run in EFL text books. In G. Low, Z. Tood, A. Deignan & L. Cameron (Eds.), *Researching and applying metaphor in the real world* (pp. 81–104). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.26.06ama

Asher, J. J. (1981). The total physical response: Theory and practice. *Annals of the New York Academy of Sciences*, *379*, 324–331. https://doi.org/10.1111/j.1749-6632.1981.tb42019.x

Azuma, M. (2005). *Metaphorical competence in an EFL context*. Toshindo.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and Developing useful language tests*. Oxford University Press.

Baese-Berk, M. M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *138*(3), 223–228. https://doi.org/10.1121/1.4929622

Beréndi, M., Csábi, S., & Kövecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 65–100). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.2.65

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, D., & Gray, B. (2016*). Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511920776

Bielak, J. (2011). Cognitive linguistics and foreign language pedagogy: An overview of recent trends and developments. In M. Pawlak (Ed.), *Extending the boundaries of research on second language learning and teaching* (pp. 241–262). Springer.

Boers, F. (2000a). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes*, *19*(2), 137–147. https://doi-org.ezproxy.unex.es/10.1016/S0889-4906(98)00017-9

Boers, F. (2000b). Metaphor awareness and vocabulary retention. *Applied Linguistics*, *21*(4), 553–571. https://doi.org/10.1093/applin/21.4.553

Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect*, *16*(3), 35–43.

Boers, F. (2003). Applied linguistics perspectives on cross-cultural variation in conceptual metaphor. *Metaphor and Symbol*, *18*(4), 231–238. https://doi.org/10.1207/S15327868MS1804_1

Boers, F. (2004). Expanding learners' vocabulary through metaphor awareness: What expansion, what learners, what vocabulary? In M. Achard & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 211–232). De Gruyter Mouton. https://doi.org/10.1515/9783110199857.211

Boers, F. (2011). Cognitive semantic ways of teaching figurative phrases: An assessment. *Review of Cognitive Linguistics*, *9*(1), 227–261. https://www.jbe-platform.com/content/journals/10.1075/rcl.9.1.11boe

Boers, F. (2013). Cognitive linguistic approaches to teaching vocabulary: Assessment and integration. *Language Teaching*, *46*(2), 208–224. https://doi.org/10.1017/S0261444811000450

Boers, F. (2022, October 27–28). *Instructional approaches to multiword expressions in a second language: A critical review*. [Plenary session]. VI Congreso Internacional de Lingüística y Literatura, Santander, Cantabria, Spain. https://congresolinguisticayliteratura.unican.es/

Boers, F. (2023). Vocabulary as a pedagogical target. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–6). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405198431.wbeal20454

Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *English Language Teaching Journal*, *55*(3), 255–262. https://doi.org/10.1093/elt/55.3.255

Boers, F., & Lindstromberg, S. (2006). Cognitive linguistic applications in second or foreign language instruction: Rationale, proposals, and evaluation. In G. Kristiansen, M. Achard, R. Dirven & F. J. Ruiz de Mendoza Ibáñez (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 305–358). De Gruyter Mouton. https://doi.org/10.1515/9783110197761.4.305

Boers, F., & Lindstromberg, S. (2008a). *Cognitive linguistic approaches to teaching vocabulary and phraseology*. De Gruyter Mouton. https://doi.org/10.1515/9783110199161

Boers, F., & Lindstromberg, S. (2008b). How cognitive linguistics can foster effective vocabulary teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 1–64). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.1.1

Boers, F., & Lindstromberg, S. (2009). *Optimising a lexical approach to instructed second language acquisition*. Palgrave Macmillan.

Boers, F., Bui, T., Deconinck, J., Stengers, H., & Coxhead, A. (2023). Helping learners develop autonomy in acquiring multiword expressions. *The Modern Language Journal*, *107*(1), 222–241. https://doi.org/10.1111/modl.12829

Boers, F., De Rycker, A., & De Knop, S. (2010). Fostering language teaching efficiency through cognitive linguistics: Introduction. In S. De Knop, F. Boers & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 1–26). De Gruyter Mouton. https://doi.org/10.1515/9783110245837.1

Boers, F., Demecheleer, M., & Eyckmans, J. (2004a). Cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies*, *8*(3), 375–388. https://doi.org/10.1080/1382557042000277449

Boers, F., Demecheleer, M., & Eyckmans, J. (2004b). Etymological elaboration as a strategy for learning figurative idioms. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 53–78). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.10.07boe

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 189–116). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.2.189

Boers, F., Piquer-Píriz, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research*, *13*(4), 367–382. https://doi.org/10.1177/1362168809341505

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, *36*(1), 1–22. https://doi.org/10.1093/applin/amt018

Brook-Hart, G. (2014). *Complete First for Spanish speakers* (2nd ed.). Cambridge University Press.

Cambridge dictionary online. (2022). Cambridge University Press. Available at https://dictionary.cambridge.org/

Cambridge idioms dictionary. (2006). Cambridge University Press.

Cameron, D., & Deignan, A. (2006). *The emergence of metaphor in discourse*. *Applied Linguistics*, *27*(4), 671–690. https://doi.org/10.1093/applin/aml032

Cameron, L. (2003). *Metaphor in educational discourse*. Continuum.

Cameron, L. (2008). Metaphor and talk. In R. W. Gibbs (Ed.), *Cambridge handbook of metaphor and thought* (pp. 197–211). Cambridge University Press. https://doi.org/10.1017/CBO9780511816802.013

Cameron, L., & Maslen, R. (2010). Identifying metaphors in discourse data. In L. Cameron & R. Maslen (Eds.), *Metaphor analysis: Research practice in applied linguistics, social sciences and the humanities* (pp. 97–115). Equinox.

Cameron, L., & Stelma, J. (2004). Metaphor clusters in discourse. *Journal of Applied Linguistics*, *1*(2), 107–136. https://doi.org/10.1558/japl.2004.1.2.107

Casasanto, D., & Gijssels, T. (2015). What makes a metaphor an embodied metaphor? *Linguistics Vanguard*, *1*(1), 327–337. https://doi.org/10.1515/lingvan-2014-1015

Castellano-Risco, I., & Piquer-Píriz, A. M. (2020). Measuring secondary-school L2 learners vocabulary knowledge: Metaphorical competence as part of general lexical competence. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 199–218). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-010

Castellano-Risco, I., Martín-Gilete, M., Hijazo-Gascón, A., & Ibarretxe-Antuñano, I. (2023). Metaphors set in motion in the context of L2 academic spoken discourse. *VIAL*, *20*, 77–105. https://doi.org/10.35869/vial.v0i20.4355

Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics*, *23*(1), 104–133. https://doi.org/10.1093/applin/23.1.104

Charteris-Black, J. (2004). *Corpus approaches to critical metaphor analysis*. Palgrave Macmillan. https://doi.org/10.1057/9780230000612

Charteris-Black, J., & Musolff, A. (2003). 'Battered hero' or 'innocent victim'? A comparative study of metaphors for euro trading in British and German financial reporting. *English for Specific Purposes*, *22*(2), 153–176. https://doi.org/10.1016/S0889-4906(02)00012-1

Chen, Y., & Lai, H. (2011). The effects of EFL learners' awareness and retention in learning metaphoric and metonymic expressions. In E. Hong & M. Dong (Eds.), *Proceedings of the 25th Pacific Asia conference on language, information and computation* (pp. 541–548).

Clandfield, L. (2003). Vocabulary: Metaphor. Available at https://www.onestopenglish.com/

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Condon, N. (2008). How cognitive linguistics motivations influence the learning of phrasal verbs. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 133–158). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.2.133

Condon, N., & Kelly, P. (2002). Does Cognitive Linguistics have anything to offer English language learners in their efforts to master phrasal verbs? *I.T.L. – International Journal of Applied Linguistics*, *137*(1), 205–231. https://doi.org/10.1075/itl.137-138.03con

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Council of Europe Publishing. Available at www.coe.int/lang-cefr

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. Available at www.coe.int/lang-cefr

Cuberos, R., Rosado, E., & Perera, J. (2019). Using deliberate metaphor in discourse: Native vs. non-native text production. In I. Navarro i Ferrando (Ed.), *Current approaches to*

metaphor analysis in discourse (pp. 235–256). De Gruyter Mouton. https://doi.org/10.1515/9783110629460-011

Danesi, M. (1992). Metaphorical competence in second language acquisition and second language teaching: The neglected dimension. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. 489–500). Georgetown UP.

Danesi, M. (1995). Learning and teaching languages: The role of "conceptual fluency". *International Journal of Applied Linguistics*, *5*(1), 3–20. https://doi.org/10.1111/j.1473-4192.1995.tb00069.x

Danesi, M. (2008). Conceptual errors in second-language learning. In S. de Knop and T. de Rycker (Eds.), *Cognitive approaches to pedagogical grammar* (pp. 231–256). De Gruyter Mouton. https://doi.org/10.1515/9783110205381.2.231

De Knop, S., & Dirven, R. (2008). Motion and location events in German, French and English: A typological, contrastive and pedagogical approach. In S. de Knop & T. de Rycker (Eds.), *Cognitive approaches to pedagogical grammar* (pp. 295–324). De Gruyter Mouton. https://doi.org/10.1515/9783110205381.3.295

De Knop, S., Boers, F., & De Rycker, A. (Eds.). (2010). *Fostering language teaching efficiency through cognitive linguistics*. De Gruyter Mouton. https://doi.org/10.1515/9783110245837

Deignan, A. (2005). *Metaphors and corpus linguistics*. John Benjamins Publishing Company. https://doi.org/10.1075/celcr.6

Deignan, A. (2011). Deliberateness is not unique to metaphor: A response to Gibbs. *Metaphor in the Social World*, *1*(1), 57–60. https://doi.org/10.1075/msw.1.1.05dei

Deignan, A., Gabryś, D., & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal*, *51*(4), 352–360. https://doi.org/10.1093/elt/51.4.352

Deignan, A., Littlemore, J., & Semino, E. (2013). *Figurative language, genre and register*. Cambridge University Press.

Deignan, A., Semino, E., & Paul, S.-A. (2019). Metaphors of climate science in three genres: Research articles, educational texts, and secondary school student talk. *Applied Linguistics*, *40*(2), 379–403. https://doi.org/10.1093/applin/amx035

Dirdal, H., Hasund, I. K., Drange, E.-M., Vold, E., & Berg E.-M. (2022). Design and construction of the Tracking Written Learner Language (TRAWL) corpus: A longitudinal and multilingual young learner corpus. *Nordic Journal of Language Teaching and Learning*, *10*(2), 115–135. https://doi.org/10.46364/njltl.v10i2.1005

Dirven, R. (2001). English phrasal verbs: Theory and didactic application. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied cognitive linguistics II: Language pedagogy* (pp. 3–28). De Gruyter Mouton. https://doi.org/10.1515/9783110866254.3

Doiz, A., & Elizari, C. (2013). Metaphoric competence and the acquisition of figurative vocabulary in foreign language learning. *Estudios de Lingüística Inglesa Aplicada*, *13*, 47–82. http://doi.org/10.12795/elia.2013.i13.02

Dorst, A. G. (2011). *Metaphor in fiction: Language, thought and communication* [Doctoral dissertation, Vrije Universiteit Amsterdam]. Uitgeverij BoxPress. https://research.vu.nl/en/publications/metaphor-in-fiction-language-thought-and-communication

Dunn, G. (1989). *Design and analysis of reliability studies: The statistical evaluation of measurement errors*. Oxford University Press.

Ellis, N. C., & Robinson, P. (2008). An introduction to cognitive linguistics, second language acquisition, and language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 3–24). Routledge.

Ervas, E., Gola, E., & Rossi, M. G. (Eds). (2019). *Metaphor in communication, science and education*. De Gruyter Mouton. https://doi.org/10.1515/9783110549928

Eyckmans, J., & Lindstromberg, S. (2016). The power of sound in L2 idiom learning. *Language Teaching Research SAGE Journals*, *21*(3), 341–361. https://doi.org/10.1177/1362168816655831

Fielden-Burns, L. V., & Piquer-Píriz, A. M. (2022). Personification and relationships in English as a medium of instruction business discourse: Crossing paths in metaphorical constructions. *Journal of Pragmatics*, *202*, 145–159. https://doi.org/10.1016/j.pragma.2022.11.002

Gao, L., & Meng, G. (2010). A study of the effect of metaphor awareness raising on Chinese EFL learners' vocabulary acquisition and retention. *Canadian Social Science*, *6*(2), 110–124. http://dx.doi.org/10.3968/j.css.1923669720100602.012

Gibbs, R. W. (1994). *The poetics of mind*. Cambridge University Press.

Gibbs, R. W. (2011). Are "deliberate" metaphors really deliberate? A question of human consciousness and action. *Metaphor and the Social World*, *1*(1), 26–52. https://doi.org/10.1075/msw.1.1.03gib

Gibbs, R. W. (2012). Metaphor, snowflakes, and termite nests: How nature creates such beautiful things. In F. MacArthur, J. L. Oncins-Martínez, M. Sánchez-García & A. M. Piquer-Píriz (Eds.), *Metaphor in use: Context, culture, and communication* (pp. 347–372). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.38.25gib

Gibbs, R. W. (2014). Why do some people dislike conceptual metaphor theory? *Cognitive Semiotics*, *5*(1–2) 14–36. https://doi.org/10.1515/cogsem.2013.5.12.14

Gibbs, R. W. (2016). Introduction. In R. W. Gibbs (Ed.), *Mixing metaphor* (pp. vii–xiv). John Benjamins Publishing Company. https://doi.org/10.1075/milcc.6.001int

Gibbs, R. W. (2021). Metaphorical embodiment. In M. D. Robinson & L. E. Thomas (Eds.), *Handbook of embodied psychology*. Springer. https://doi.org/10.1007/978-3-030-78471-3_5

Gibbs, R. W. (Ed.). (2008). *The Cambridge handbook of metaphor and thought*. Cambridge University Press. https://doi.org/10.1017/CBO9780511816802

Giessler, R. (2012). Teacher language awareness and cognitive linguistics (CL): Building a CL-inspired perspective on teaching lexis in EFL student teachers. *Language Awareness*, *21*(1–2), 113–135. https://doi.org/10.1080/09658416.2011.639891

Goatly, A. (1997). *The language of metaphors*. Routledge.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago University Press.

Golden, A. (2010). Grasping the point: A study of 15-year-old students' comprehension of metaphorical expressions in schoolbooks. In G. Low, Z. Todd, A. Deignan & L. Cameron (Eds.), *Researching and applying metaphor in the real world* (pp. 35–62). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.26.04gol

Golden, A. (2012). Metaphorical expressions in L2 production: The importance of the text topic in corpus research. In F. MacArthur, J. L. Oncins-Martínez, M. Sánchez-García & A. M. Piquer-Píriz (Eds.), *Metaphor in use: Context, culture, and communication* (pp. 135–148). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.38.11gol

Golden, A. (2021). The relationship between topic and metaphor in second-language learners' essays. *Metaphor and the Social World*, *11*(2), 261–278. https://doi.org/10.1075/msw.00018.gol

Guo, S. (2007). Is idiom comprehension influenced by metaphor awareness of the learners? A case study of Chinese EFL learners. *Linguistics Journal*, *2*(3), 148–166.

Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar*. Routledge.

Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Pearson Education Ltd.

Hashemian, M. (2013). Effects of focus-on-form(s) instruction on Iranian intermediate L2 learners' metaphorical competence development. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *32*(1), 1–17. https://doi.org/10.22099/jtls.2013.1491

Herrera, H., & White, M. (2000). Cognitive linguistics and the language learning process: A case from Economics. *Estudios Ingleses de la Universidad Complutense de Madrid*, *8*, 55–78.

Herrmann, J. B. (2013). *Metaphor in academic discourse: Linguistic forms, conceptual structures, communicative functions and cognitive representations* [Doctoral

dissertation, Vrije Universiteit Amsterdam]. LOT. https://research.vu.nl/en/publications/metaphor-in-academic-discourse-linguistic-forms-conceptual-struct-2

Hijazo-Gascón, A., & Llopis-García, R. (2019). Applied cognitive linguistics and foreign language learning. Introduction to the special issue. *International Review of Applied Linguistics in Language Teaching*, *57*(1), 1–20. https://doi.org/10.1515/iral-2018-2004

Hoang, H. (2014). Metaphor and second language learning: The state of the field. *TESL-EJ*, *18*(2). Retrieved from https://tesl-ej.org/wordpress/issues/volume18/ej70/ej70a5/

Hoang, H., & Boers, F. (2018). Gauging the association of EFL learners' writing proficiency and their use of metaphorical language. *System*, *74*, 1–8. https://doi.org/10.1016/j.system.2018.02.004

Hoffman, R. R. (1983). Recent research on metaphor. *Semiotic Inquiry*, *3*, 35–62.

Holme, R. (2004). *Mind, metaphor and language teaching*. Palgrave Macmillan. https://doi.org/10.1057/9780230503007

Holme, R. (2009). *Cognitive linguistics and language teaching*. Palgrave Macmillan. https://doi.org/10.1057/9780230233676

Howell, D. C. (2010). *Statistical methods for psychology*. Wadsworth, Cengage Learning.

Hughes, B., & Heasley, B. (2018). *Linguistics: The Cambridge survey*. Cambridge University Press.

Ibarretxe-Antuñano, I., Cadierno, T., & Castañeda Castro, A (Eds.). (2019). *Lingüística cognitiva y español LE/L2*. Routledge.

Kaal, A. A. (2012). *Metaphor in conversation* [Doctoral dissertation, Vrije Universiteit Amsterdam]. BOXPress. https://research.vu.nl/en/publications/metaphor-in-conversation

Kamberi, L. (2014). Using metaphors in language teaching and learning. *European Journal of Research in Education,* *2*(2), 92–97. https://api.semanticscholar.org/CorpusID:26526579

Kathpalia, S. S., & Carmel, H. K. H. (2011). Metaphorical competence in ESL student writing. *RELC Journal*, *42*(3), 273–290. https://doi.org/10.1177/0033688211419379

Kecskes, I. (2006). On my mind: Thoughts about salience, context and figurative language from a second language perspective. *Second Language Research*, *22*(2), 219–237. https://doi.org/10.1191/0267658306sr266ra

Kövecses, Z. (2002). *Metaphor: A practical introduction*. Oxford University Press.

Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511614408

Kövecses, Z. (2009). Metaphor, culture and discourse: The pressure of coherence. In A. Musolff & J. Zinken (Eds.), *Metaphor and discourse* (pp. 11–24). Palgrave Macmillan. https://doi.org/10.1057/9780230594647_2

Kövecses, Z., & Szabó, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics*, *17*(3), 326–355. https://doi.org/10.1093/applin/17.3.326

Krennmayr, T. (2011). *Metaphor in newspapers* [Doctoral dissertation, Vrije Universiteit Amsterdam]. LOT. https://research.vu.nl/en/publications/metaphor-in-newspapers

Krennmayr, T. (2017). Metaphor and parts-of-speech. In E. Semino & Z. Demjén (Eds.), *The Routledge handbook of metaphor and language* (pp. 165–177). Routledge.

Krennmayr, T., MacArthur, F., & Nacey, S. (2022). *UNDERSTANDING IS SEEING in business seminars* [Paper presentation]. 15th Researching and Applying Metaphor Conference. Białystok, Poland. https://raam15.uwb.edu.pl/

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 202–251). Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.013

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge as western thought*. Basic Books.

Langacker, R. W. (1987*). Foundations of cognitive grammar. Volumen 1: Theoretical prerequisites*. Standford University Press.

Langacker, R. W. (1990). *Concept, image, and symbol: The cognitive basis of grammar*. De Gruyter Mouton. https://doi.org/10.1515/9783110857733

Lazar, G. (2003). *Meaning and metaphors: Activities to practise figurative language*. Cambridge University Press.

Li, F. T. (2002). *The acquisition of metaphorical expressions, idioms and proverbs by Chinese learners of English: A conceptual metaphor and image schema-based approach* [Doctoral dissertation]. Chinese University of Hong Kong.

Lindstromberg, S. (1996). Prepositions: Meaning and method. *ELT Journal*, *50*(3), 225–236. https://doi.org/10.1093/elt/50.3.225

Lindstromberg, S. (2010). *English prepositions explained* (2nd ed.). John Benjamins Publishing Company.

Lindstromberg, S., & Boers, F. (2005). From movement to metaphor with manner-of-movement verbs. *Applied Linguistics*, *26*(2), 241–261. https://doi.org/10.1093/applin/ami002

Lindstromberg, S., & Boers, F. (2008). *Teaching chunks of language: From noticing to remembering*. Helbling Languages.

Littlemore, J. (2001a). Metaphoric competence: A language learning strength of students with a holistic cognitive style? *TESOL Quarterly*, *35*(3), 459–491. https://doi.org/10.2307/3588031

Littlemore, J. (2001b). The use of metaphor in university lectures and the problems that it causes for overseas students. *Teaching in Higher Education*, *6*(3), 333–349. https://doi.org/10.1080/13562510120061205

Littlemore, J. (2003). The effect of cultural background on metaphor interpretation. *Metaphor and Symbol*, *18*(4), 273–288. https://doi.org/10.1207/S15327868MS1804_4

Littlemore, J. (2004). Interpreting metaphors in the EFL classroom. *Cahiers de l'Apliut*, *23*(2), 57–70. https://doi.org/10.4000/apliut.3339

Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching* (1st ed.). Palgrave Macmillan. https://doi.org/10.1057/9780230245259

Littlemore, J. (2010). Metaphoric competence in the first and second language: Similarities and differences. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind* (pp. 293–315). John Benjamins Publishing Company. https://doi.org/10.1075/celcr.13.20lit

Littlemore, J. (2023). *Applying cognitive linguistics to second language learning and teaching* (2nd ed.). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-39796-7

Littlemore, J., & Fielden-Burns, L. V. (2023). On the fringes of metaphor: Using ambiguously figurative vague language to pragmatically negotiate sensitive topics in the English as a medium of instruction classroom. *Journal of Pragmatics*, *209*, 1–14. https://doi.org/10.1016/j.pragma.2023.02.016

Littlemore, J., & Juchem-Grundmann, C. (Eds.). (2010). *Applied cognitive linguistics in second language learning and teaching. AILA Review, 23*. John Benjamins Publishing Company. https://doi.org/10.1075/aila.23

Littlemore, J., & Low, G. (2006a). *Figurative thinking and foreign language learning*. Palgrave Macmillan.

Littlemore, J., & Low, G. (2006b). Metaphoric competence, second language learning, communicative language ability. *Applied Linguistics*, *27*(2), 268–294. https://doi.org/10.1093/applin/aml004

Littlemore, J., Chen, P. R., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics*, *32*(4), 408–429. https://doi.org/10.1093/applin/amr009

Littlemore, J., Krennmayr, T., Turner, J., & Turner, S. (2014). An investigation into metaphor use at different levels of second language writing. *Applied Linguistics*, *35*(2), 117–144. https://doi.org/10.1093/applin/amt004

Littlemore, J., MacArthur, F., Rubienska, A., & Turner, J. (2013). Communicating academic content to international students: Interplay and variations in the use of verbal and

gestural metaphor. *Revista Española de Lingüística Aplicada*, *1*, 23–50. https://acortar.link/Xxks1g

Littlemore, J., Turner, S., & Tuck, P. (2023). *Creative metaphor, emotion and evaluation in conversations about work.* Routledge.

Llopis-García, R., Martín-Gascón, B., & Alonso-Aparicio, I. (2022, June 27–29). *Challenges of empirical testing in L2 and new pathways for better results* [Paper presentation]. 12th International Conference of the Spanish Cognitive Linguistics Association. Logroño, La Rioja, Spain. https://aelco2020.unirioja.es/

Longman dictionary of contemporary English online (2022). Pearson Education. Available at https://www.ldoceonline.com/

Low, G. (1988). On teaching metaphor. *Applied Linguistics*, *9*(2), 125–147. https://doi.org/10.1093/applin/9.2.125

Low, G. (2008). Metaphor and education. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 212–231). Cambridge University Press. https://doi.org/10.1017/CBO9780511816802.014

Low, G. (2010). Wot no similes? The curious absence of simile in university lectures. In G. Low, Z. Todd, A. Deignan & L. Cameron (Eds.), *Researching and Applying Metaphor in the Real World* (pp. 291–308). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.26.17low

Low, G. (2017). Eliciting metaphor in education research: Is it really worth the effort? In F. Ervas, E. Gola & M. G. Rossi (Eds.), *Metaphor in communication, science and education* (pp. 249–266). De Gruyter Mouton. https://doi.org/10.1515/9783110549928-014

Low, G. (2020). Taking stock after three decades: “On teaching metaphor” revisited. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 37–56). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-003

Low, G., Littlemore, J., & Koester, A. (2008). Metaphor use in three UK university lectures. *Applied Linguistics*, *29*(3), 428–455. https://doi.org/10.1093/applin/amn008

Lowery, D. (2013). Helping metaphors take root in the EFL classroom. *English Teaching Forum*, *51*(1), 12–17.

MacArthur, F. (2005). The competent horseman in a horseless world: Observations on a conventional metaphor in Spanish and English. *Metaphor and Symbol*, *20*(1), 71–94. https://doi.org/10.1207/s15327868ms2001_3

MacArthur, F. (2010). Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review*, *23*(1), 155–173. https://doi.org/10.1075/aila.23.09mac

MacArthur, F. (2015). On using a dictionary to identify the basic senses of words. *Metaphor and the Social World*, *5*(1), 137–144. https://doi.org/10.1075/msw.5.1.07mac

MacArthur, F. (2016a). Overt and covert uses of metaphor in the academic mentoring in English of Spanish undergraduate students at five European universities. *Review of Cognitive Linguistics*, *14*(1), 23–50. https://doi.org/10.1075/rcl.14.1.02mac

MacArthur, F. (2016b). Beyond engaged listenership: Assessing Spanish undergraduates' active participation in academic mentoring sessions in English as academic lingua franca. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2016: Global Implications for Culture and Society in the Networked Age* (pp. 153–178). Springer. https://doi.org/10.1007/978-3-319-41733-2_8

MacArthur, F. (2016c). When languages and cultures meet: Mixed metaphors in the discourse of Spanish speakers of English. In R. W. Gibbs (Ed.), *Mixing metaphor* (pp. 133–154). John Benjamins Publishing Company. https://doi.org/10.1075/milcc.6.07mac

MacArthur, F. (2017). Using metaphor in the teaching of second/foreign languages. In E. Semino & Z. Demjén (Eds.), *The Routledge handbook of metaphor and language* (pp. 413–425). Routledge.

MacArthur, F. (2019). Linguistic metaphor identification in English as a lingua franca. In S. Nacey, A. G. Dorst, T. Krennmayr & W. G. Reijnierse (Eds.), *Metaphor identification in multiple languages: MIPVU around the world* (pp. 289–312). John Benjamins Publishing Company. https://doi.org/10.1075/celcr.22.14mac

MacArthur, F. (2021). Afterword. *Metaphor and the Social World*, *11*(2), 352–359. https://doi.org/10.1075/msw.00022.mac

MacArthur, F., & Alejo-González, R. (2024). Beyond idioms, the use of metaphor in ELF academic settings: A comprehensive review. *Journal of Pragmatics*, *219*, 48–57. https://doi.org/10.1016/j.pragma.2023.11.002

MacArthur, F., & Boers, F. (forthcoming). Using visuals to illustrate the source domains of idioms: Can they help learners appreciate usage restrictions too? In C. Jurchen & S. Nieimier (Eds.), *KNOWING IS SEEING: Metaphor and language pedagogy*. De Gruyter Mouton.

MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 159–188). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.2.159

MacArthur, F., & Littlemore, J. (2011). On the repetition of words with the potential for metaphoric extension in conversations between native and non-native speakers of English. *Metaphor and the Social World*, *1*(2), 201–239. https://doi.org/10.1075/msw.1.2.05mac

MacArthur, F., & Piquer-Píriz, A. M. (2007). Staging the introduction of figurative extensions of familiar vocabulary items in EFL: Some preliminary considerations. *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies*, *53*, 123–134.

MacArthur, F., Alejo-González, R., Piquer-Píriz, A. M., Amador-Moreno, C., Littlemore, J., Ädel, A., Krennmayr, T., & Vaughn, E. (2014). EuroCoAT. The European corpus of academic talk. http://www.eurocoat.es

MacArthur, F., Krennmayr, T., & Littlemore, J. (2015). How basic is "UNDERSTANDING IS SEEING" when reasoning about Knowledge? Asymmetric uses of sight metaphors in office hours consultations in English as Academic Lingua Franca. *Metaphor and Symbol*, *30*(3), 184–217. https://doi.org/10.1080/10926488.2015.1049507

Macmillan English dictionary online. (2023). Macmillan Education. Available at https://www.macmillandictionary.com/

Martín-Gascón, B. (2023). Developing L2 learners' metaphoric competence: A case study of figurative motion constructions. *International Review of Applied Linguistics in Language Teaching*, *61*(1), 79–109. https://doi.org/10.1515/iral-2022-0043

Martín-Gascón, B., Llopis-García, R., & Alonso-Aparicio, I. (2023). Does L2 assessment make a difference? Testing the empirical validity of applied cognitive linguistics in the acquisition of the Spanish/L2 psych-verb construction. *Language Teaching Research*. https://doi.org/10.1177/13621688231190981

Martín-Gilete, M. (2022a). CL-oriented approaches to teaching phrasal verbs: A report on EFL classroom-based research. *Human Review. International Humanities Review*, *11*, 1–11. https://doi.org/10.37467/revhuman.v11.3911

Martín-Gilete, M. (2022b). The role of input in the use of metaphor in L2 writing. *ES Review: Spanish Journal of English Studies*, *43*, 207–241. https://doi.org/10.24197/ersjes.43.2022.207-241

Masuda, K. (2023). Learning Japanese interactional particles through a usage-based and concept-based language instruction. *East Asian Pragmatics*, *8*(1), 27–56. https://doi.org/10.1558/eap.22036

McCarthy, M., & O'Dell, F. (2017a). *English idioms in use. Intermediate* (2nd ed.). Cambridge University Press.

McCarthy, M., & O'Dell, F. (2017b). *English phrasal verbs in use. Intermediate* (2nd ed.). Cambridge University Press.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge University Press.

Meissner, V. H. (2010). *Metaphorical comprehension in grade 8 English language learners from one town in Norway* [Unpublished master's thesis]. Hogskolen I Hedmark.

Microsoft Corporation. (2021). *Microsoft Excel* [Computer software]. Available at https://office.microsoft.com/excel

Microsoft Corporation. (2021). *Microsoft PowerPoint* [Computer software]. Available at https://office.microsoft.com/powerpoint

Millar, E. (2023). *Exploring linguistic research influence in mainstream English language teaching: The case of multi-word verbs*. [Unpublished doctoral dissertation]. University of Cantabria. https://catalogo.unican.es/cgi-bin/abnetopac?TITN=446587

Morimoto, S., & Loewen, S. (2007). A comparison of the effects of image-schema-based instruction and translation-based instruction on the acquisition of L2 polysemous words. *Language Teaching Research*, *11*(3), 347–372. https://doi.org/10.1177/1362168807081

Nacey, S. (2010). *Comparing linguistic metaphors in L1 and L2 English*. [Unpublished doctoral dissertation]. University of Oslo. Retrieved from https://susannacey.inn.no/books-and-theses/

Nacey, S. (2013). *Metaphors in learner English*. John Benjamins Publishing Company. https://doi.org/10.1075/milcc.2

Nacey, S. (2017). Metaphor comprehension and production in a second language. In E. Semino & Z. Demjén (Eds.), *The Routledge handbook of metaphor and language* (pp. 503–415). Routledge.

Nacey, S. (2020). Development of L2 metaphorical production. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 173–198). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-009

Nacey, S. (2022). Development of metaphorical production in learner language: A longitudinal perspective. *Nordic Journal of Language Teaching and Learning*, *10*(2), 272–297. https://doi.org/10.46364/njltl.v10i2.975

Nacey, S., & Uri Jensen, B. (2019). Metaphoricity in English L2 learners' prepositions. In F. Ervas, E. Gola & M. G. Rossi (Eds.), *Metaphor in communication, science and*

education (pp. 283–304). De Gruyter Mouton. https://doi.org/10.1515/9783110549928-016

Nacey, S., Dorst, A. G., Krennmayr, T., & Reijnierse, W. G. (2019a). *Metaphor identification in multiple languages: MIPVU around the world*. John Benjamins Publishing Company. https://doi.org/10.1075/celcr.22

Nacey, S., Krennmayr, T., Dorst, A. G., & Reijnierse, W. G. (2019b). What the MIPVU protocol doesn't tell you (even though it mostly does). In S. Nacey, A. G. Dorst, T. Krennmayr & W. G. Reijnierse (Eds.), *Metaphor identification in multiple languages: MIPVU around the world* (pp. 289–312). John Benjamins Publishing Company. https://doi.org/10.1075/celcr.22.03nac

Niemeier, S. (2004). Linguistic and cultural relativity – Reconsidered for the foreign language classroom. In M. Achard & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 95–118). De Gruyter Mouton. https://doi.org/10.1515/9783110199857.95

Niemeier, S. (2017). Teaching (in) metaphors. In F. Ervas, E. Gola & M. G. Rossi (Eds.), *Metaphor in communication, science and education* (pp. 267–282). De Gruyter Mouton. https://doi.org/10.1515/9783110549928-015

O'Reilly, D., & Marsden, E. (2021). Eliciting and measuring L2 metaphoric competence: Three decades on from Low (1988). *Applied Linguistics*, *42*(1), 24–59. https://doi.org/10.1093/applin/amz066

O'Reilly, D., & Marsden, E. (2023). Elicited metaphoric competence in a second language: A construct associated with vocabulary knowledge and general proficiency? *International Review of Applied Linguistics in Language Teaching*, *61*(2), 287–327. https://doi.org/10.1515/iral-2020-0054

O'Reilly, D., & Yan, L. (2023). Playing with second language metaphor: An exploration with advanced Chinese learners of English. *Applied Linguistics*, 1–22. https://doi.org/10.1093/applin/amad067

Oxford English dictionary online. (2023). Oxford University Press. Available at https://www.oed.com/

Oxford idioms dictionary. (2006). Oxford University Press.

Özçalışkan, Ş. (2011). Acquisition of metaphor. In P. Hogan (Ed.), *Cambridge encyclopedia of language sciences* (pp. 486–488). Cambridge University Press.

Paivio, A. (1990). *Mental representations: A dual-coding approach*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Pan, M. X. (2019). The effectiveness of the conceptual metaphor approach to English idiom acquisition by young Chinese learners. *Metaphor and the Social World*, *9*(1), 59–82. https://doi.org/10.1075/msw.17024.pan

Philip, G. (2005). From concept to wording and back again: Features of learners' production of figurative language. In A. Wallington, J. Barnden, S. Glasbey, M. Lee & I. Zhang (Eds)., *Proceedings of the third interdisciplinary workshop on corpus-based approaches to figurative language* (pp. 63–79). University of Birmingham Press.

Philip, G. (2010). Drugs, traffic, and many other dirty interests: Metaphor and the language learner. In G. Low, Z. Todd, A. Deignan & L. Cameron (Eds.), *Researching and Applying metaphor in the real world* (pp. 63–79). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.26.05phi

Philip, G. (2011). *Colouring meaning: Collocation and connotation in figurative language*. John Benjamins. https://doi.org/10.1075/scl.43

Philip, G. (2023). Metaphorical SPACE in academic talk. *Journal of Pragmatics*, *210*, 1–11. https://doi.org/10.1016/j.pragma.2023.03.011

Picken, J. (2007). *Literature, metaphor and the foreign language learner*. Palgrave Macmillan.

Piquer-Píriz, A. M. (2004). Young EFL learners' understanding of some semantic extensions of the lexemes 'hand', 'mouth' and 'head'. [Unpublished doctoral dissertation]. University of Extremadura. Retrieved from https://biblioteca.unex.es/tesis/8477236577.pdf

Piquer-Píriz, A. M. (2008a). Reasoning figuratively in early EFL: Some implications for the development of vocabulary. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 233–257). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.3.219

Piquer-Píriz, A. M. (2008b). Young learners' understanding of figurative language. In M. S. Zanotto, L. Cameron & M. C. Cavalcanti (Eds.), *Confronting metaphor in use: An applied linguistic approach* (pp. 183–198). John Benjamins Publishing Company. https://doi.org/10.1075/pbns.173.13piq

Piquer-Píriz, A. M. (2010). Can people be cold and warm? Developing understanding of figurative meanings of temperature terms in early EFL. In G. Low, Z. Todd, A. Deignan & L. Cameron (Eds.), *Researching and applying metaphor in the real world* (pp. 21–34). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.26

Piquer-Píriz, A. M. (2020). Figurative language and young L2 learners. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 57–78). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-004

Piquer-Píriz, A. M., & Alejo-González, R. (2016). Applying cognitive linguistics: Identifying some current research foci (figurative language in use, constructions and typology). *Review of Cognitive Linguistics*, *14*(1), 1–20. https://doi.org/10.1075/rcl.14.1.01piq

Piquer-Píriz, A. M., & Alejo-González, R. (2018). Applying cognitive linguistics: Identifying some current research foci (figurative language in use, constructions and typology). In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Applying cognitive linguistics* (pp. 1–19). https://doi.org/10.1075/bct.99.01piq

Piquer-Píriz, A. M., & Alejo-González, R. (Eds.). (2020). *Metaphor in foreign language instruction*. De Gruyter Mouton. https://doi.org/10.1515/9783110630367

Piquer-Píriz, A. M., & Boers, F. (2019). La lingüística cognitiva y sus aplicaciones a la enseñanza de lenguas extranjeras. In I. Ibarretxe-Antuñano, T. Cadierno & A. Castañeda Castro (Eds.), *Lingüística cognitiva y español LE/L2* (pp. 52–70). Routledge.

Piquer-Píriz, A. M., & Martín-Gilete, M. (forthcoming). Bringing cognitive linguistics into real EFL classrooms: Pedagogical experience vs. experimental results. *Review of Cognitive Linguistics*.

Ponterotto, D. (1994). Metaphors we can learn by: How insights from cognitive linguistic research can improve the teaching/learning of figurative language. *English Teaching Forum*, *32*(3), 2–7.

Pourdana, N., Sahebalzamani, S., & Rajeski, J. S. (2014). Metaphorical awareness: A new horizon in vocabulary retention by Asian EFL learners. *International Journal of Applied Linguistics & English Literature,* *3*(4), 213–220. https://doi.org/10.7575/aiac.ijalel.v.3n.4p.213

Pouscoulous, N., & Tomasello, M. (2020). Early birds: Metaphor understanding in 3-year-olds. *Journal of Pragmatics*, *156*, 160–167. https://doi.org/10.1016/j.pragma.2019.05.021

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1–39. https://doi.org/10.1080/10926480709336752

Pütz, M., Niemeier, S., & Dirven, R. (Eds.). (2001a). *Applied cognitive linguistics I: Theory and language acquisition*. De Gruyter Mouton. https://doi.org/10.1515/9783110866247

Pütz, M., Niemeier, S., & Dirven, R. (Eds.). (2001b). *Applied cognitive linguistics II: Language pedagogy*. De Gruyter Mouton. https://doi.org/10.1515/9783110866254

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.1) [Computer software]. R Foundation for Statistical Computing. Available at https://www.r-project.org/

Radden, G., & Panther, K.-U. (Eds.) (2004). *Studies in linguistic motivation*. De Gruyter Mouton.

Radić-Bojanić, B. (2013). Internalization and production of metaphorical expressions with EFL students. *Romanian Journal of English Studies*, *10*(1), 135–144. https://doi.org/10.2478/rjes-2013-0011

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, *13*(4), 519–549. https://doi.org/10.1075/ijcl.13.4.06ray

Reddy, M. J. (2012). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 164–201). Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.012

Reijnierse, W. G. (2017). *The value of deliberate metaphor* [Doctoral dissertation, Universiteit van Amsterdam]. LOT. https://research.vu.nl/en/publications/the-value-of-deliberate-metaphor

Richards, J., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511667305

Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. Routledge.

Rudzka-Ostyn, B. (2003). *Word power: Phrasal verbs and compounds*. A cognitive approach. De Gruyter Mouton. https://doi.org/10.1515/9783110197235

Rundell, M. (2007). *Macmillan English dictionary* (2nd ed.). Macmillan Publishers Limited.

Saaty, R. (2016). *Teaching L2 metaphor through awareness-raising activities: Experimental studies with Saudi EFL learners*. [Unpublished doctoral dissertation]. University of Birmingham. http://etheses.bham.ac.uk/id/eprint/7061

Saaty, R. (2020). An enactment-based approach to the teaching of metaphoric expressions: A case of Saudi EFL learners. In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 263–286). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-013

Semino, E. (2008). *Metaphor in discourse*. Cambridge University Press.

Semino, E. (2019). Afterword: Some reflections on MIPVU across languages. In S. Nacey, A. G. Dorst, T. Krennmayr & W. G. Reijnierse (Eds.), *Metaphor identification in multiple*

languages: MIPVU around the world (pp. 213–321). John Benjamins Publishing Company. https://doi.org/10.1075/celcr.22.15sem

Semino, E., Demjén, Z., Hardie, A., Payne, S., & Rayson, P. E. (2018). (Eds). *Metaphor, cancer and the end of life: A corpus-based study*. Routledge.

Shokouhi, H., & Isazadeh, M. (2009). The effect of teaching conceptual and image metaphors to EFL learners. *The Open Applied Linguistics Journal*, *2*, 22–31. https://doi.org/10.2174/1874913500902010022

Skoufaki, S. (2008). Conceptual metaphoric meaning clues in two idiom presentation methods. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 101–132). De Gruyter Mouton. https://doi.org/10.1515/9783110199161.2.101

Steen, G. (2008). The paradox of metaphor: Why we need a three-dimensional model of metaphor. *Metaphor and Symbol*, *23*(4), 213–241. https://doi.org/10.1080/10926480802426753

Steen, G. (2011a). The contemporary theory of metaphor – now new and improved! *Review of Cognitive Linguistics*, *9*(1), 26–64. https://doi.org/10.1075/rcl.9.1.03ste

Steen, G. (2011b). What does “really deliberate” really mean? More thoughts on metaphor and consciousness. *Metaphor and the Social World*, *1*(1), 53–56. https://doi.org/10.1075/msw.1.1.04ste

Steen, G. (2023). *Slowing metaphor down: Elaborating Deliberate Metaphor Theory*. John Benjamins Publishing Company. https://doi.org/10.1075/celcr.26

Steen, G., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins Publishing Company. https://doi.org/10.1075/celcr.14

Szczepaniak, R., & Lew, R. (2011). The role of imagery in dictionaries of idioms. *Applied linguistics*, *32*(3), 323–347. https://doi.org/10.1093/applin/amr001

Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, *47*(2), 420–430. https://doi.org/10.1002/tesq.91

Taylor, J. R. (2003). *Linguistic categorization*. Oxford University Press.

Turner, S. (2014). *The development of metaphoric competence in French and Japanese learners of English*. [Unpublished doctoral dissertation]. University of Birmingham. https://etheses.bham.ac.uk/id/eprint/5449/

Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge. https://doi.org/10.4324/9780203876039

UCLES. (2015). *The Cambridge English scale explained*. Cambridge Assessment English. Available at https://www.cambridgeenglish.org/images/177867-the-methodology-behind-the-cambridge-english-scale.pdf

UCLES. (2018). *Cambridge English First for schools 3*. Cambridge University Press.

UCLES. (2019a). *B2 First for schools: Handbook for teachers*. Cambridge University Press. Available at https://www.cambridgeenglish.org/images/167792-b2-first-for-schools-handbook.pdf

UCLES. (2019b). *The Cambridge English scale explained: A guide to converting practice test scores to Cambridge English scale scores*. Cambridge Assessment English. Available at https://www.cambridgeenglish.org/images/210434-converting-practice-test-scores-to-cambridge-english-scale-scores.pdf

UCLES. (2020). Assessing writing for Cambridge English Qualifications: A guide for teachers. Cambridge Assessment English. Available at https://www.cambridgeenglish.org/Images/600975-teacher-guide-for-writing-b2-first-for-schools.pdf

Vasiljevic, Z. (2011). Using conceptual metaphors and L1 definitions in teaching idioms to non-native speakers. *The Journal of Asia TEFL*, *8*(3), 135–160. https://acortar.link/SWC6OE

Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, *53*(3), 547–586. https://doi.org/10.1111/1467-9922.00234

VOICE Project. (2007). Mark-up conventions. *VOICE Transcription Conventions [2.1]*. Retrieved September 27, 2022, from https://voice.acdh.oeaw.ac.at/wp-content/uploads/2021/04/VOICE-mark-up-conventions.pdf

VOICE Project. (2007). Spelling conventions. *VOICE Transcription Conventions [2.1]*. Retrieved September 27, 2022, from https://voice.acdh.oeaw.ac.at/wp-content/uploads/2021/04/VOICE-spelling-conventions.pdf

VOICE. (2021). *The Vienna-Oxford International Corpus of English* (Version VOICE 3.0 Online). Founding director: Barbara Seidlhofer; Principal investigators VOICE 3.0: Marie-Luise Pitzl, Daniel Schopper; Researchers: Angelika Breiteneder, Hans-Christian Breuer, Nora Dorn, Theresa Klimpfinger, Stefan Majewski, Ruth Osimk-Teasdale, Hannes Pirker, Marie-Luise Pitzl, Michael Radeka, Stefanie Riegler, Barbara Seidlhofer, Omar Siam, Daniel Stoxreiter. Available at https://voice3.acdh.oeaw.ac.at

Wade, T., Jongman, A., & Sereno, J. (2007) Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica*, *64*(2–3), 122–144. https://doi.org/10.1159/000107913

Wang, X., Boers, F., & Warren, P. (2020). Using literal underpinnings to help learners remember figurative idioms: Does the connection need to be crystal clear? In A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (pp. 221–240). De Gruyter Mouton. https://doi.org/10.1515/9783110630367-011

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

Wee, L. (2006). Proper names and the theory of metaphor. *Journal of Linguistics*, *42*(2), 355–371. https://doi.org/10.1017/S0022226706003926

Yasuda, S. (2010). Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners. *TESOL Quarterly*, *44*(2), 250–273. https://doi.org/10.5054/tq.2010.219945

APPENDICES

Appendix A: Research Participants' Identification and Characteristics

Table 62

Overview of Groups of Participants: Students' Identification and Speaker-Related Information

Groups of participants[45]	B2 groups	Participants		
		Students' codes[46]	Gender	Age
Experimental	GR0101	ST0101	Female	16
		ST0102	Male	15
		ST0103	Female	16
		ST0104	Female	16
		ST0105	Female	16
		ST0106	Male	14
		ST0107	Male	16
		ST0108	Female	15
		ST0109	Male	16
		ST0110	Female	14
	GR0102	ST0201	Male	17
		ST0202	Male	17
		ST0203	Female	17
		ST0204	Female	16
		ST0205	Male	17
		ST0206	Female	16
		ST0207	Male	18
		ST0208	Female	16
		ST0209	Female	16
		ST0210	Female	16
Control	GR0203	ST0301	Male	16
		ST0302	Female	16
		ST0303	Female	15
		ST0304	Male	16
		ST0305	Male	15
		ST0306	Female	16
		ST0307	Female	16
		ST0308	Female	16
		ST0309	Male	16
		ST0310	Female	16
	GR0204	ST0401	Male	16
		ST0402	Male	17
		ST0403	Female	17
		ST0404	Male	18
		ST0405	Male	16
		ST0406	Female	17

[45] Each group of participants was assigned a unique code consisting of a descriptive abbreviation ("GR" for "group"), a standardised code ("01" for experimental, "02" for control) to indicate their group assignment, and a meaningful identifier to indicate their group number in the investigation.

[46] Each student participating in the study was assigned a unique code that included a descriptive abbreviation ("ST" for "student"), a standardised code ("01" for GR0101, "02" for GR0102, "03" for GR0203, "04" for GR0204) to indicate the group they belong to, and a meaningful identifier to indicate their participant number within the group.

	ST0407	Male	17
	ST0408	Male	17
	ST0409	Female	16
	ST0410	Female	17

Appendix B: Approval of Research Ethics

Figure 5

Scanned Copy of the Research Ethics Approval

UNIVERSIDAD DE EXTREMADURA

VICERRECTORADO DE INVESTIGACIÓN, Y TRANSFERENCIA

Campus Universitario
Avdª de Elvas s/nº
06071 BADAJOZ

Tel.: 924 28 93 05
Fax: 924 27 29 83

NºRegistro: 249//2019

D. JOÃO NUNO MEIRELES DA SILVA GONÇALVES RIBEIRO, PRESIDENTE POR DELEGACIÓN DE LA COMISIÓN DE BIOÉTICA Y BIOSEGURIDAD DE LA UNIVERSIDAD DE EXTREMADURA.

INFORMA: Que una vez analizada por esta Comisión la solicitud de Proyecto de Investigación titulado "La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2", cuyo Investigador/a Principal es D/Dª Marta Martín Gilete, ha decidido por unanimidad, valorar positivamente el precitado proyecto por considerar que se ajusta a las normas éticas esenciales cumpliendo con la normativa vigente al efecto.

Y para que conste y surta los efectos oportunos firmo el presente informe en Badajoz, a 18 de diciembre de 2019.

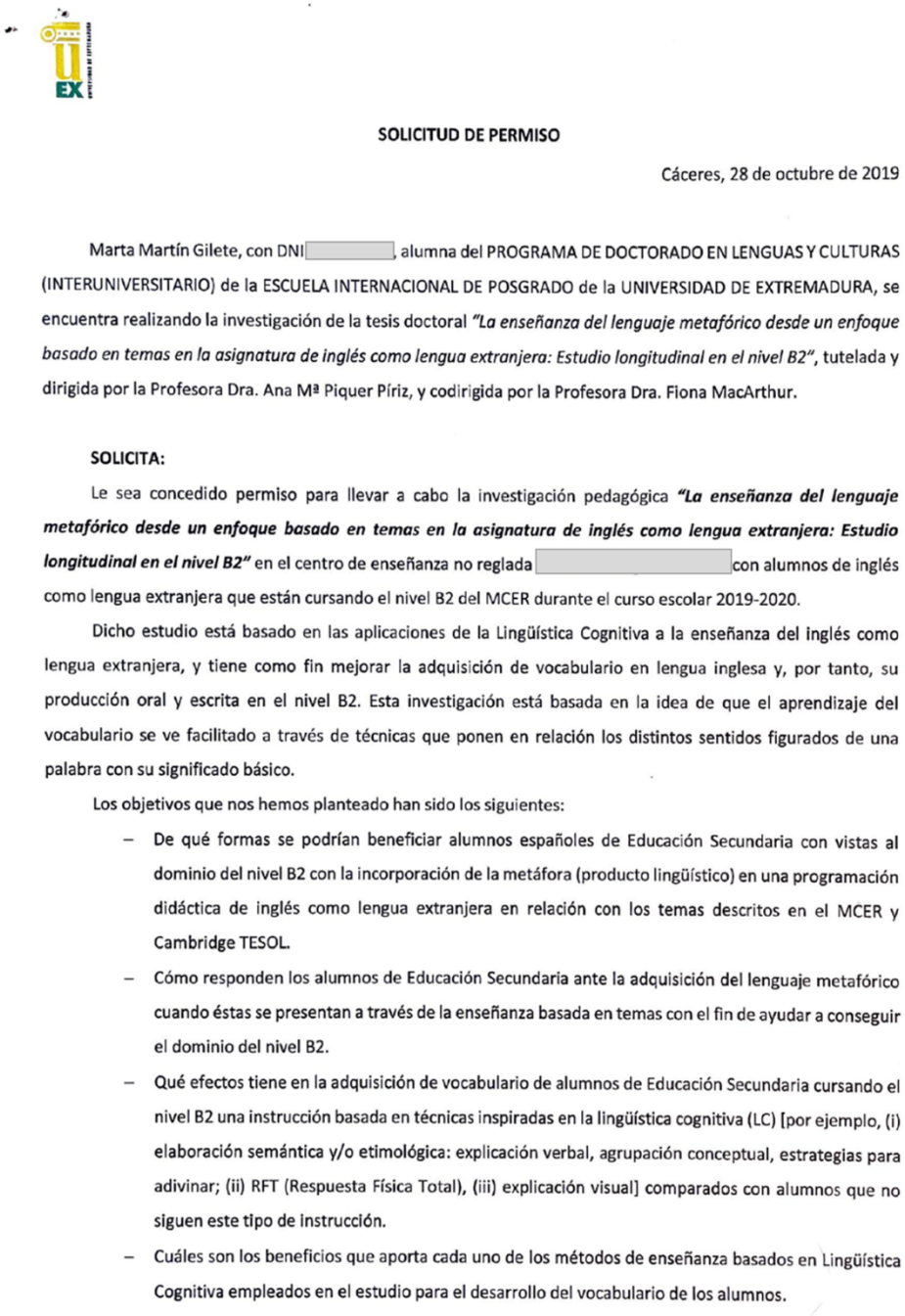
V.º B.º

Fdo.: Javier de Francisco Morcillo
Presidente por delegación de la Comisión
de Bioética y Bioseguridad

Appendix C: Research Authorisation Request Form[47]

Figure 6

Scanned Copy of the Research Authorisation Request Form: Page 1



SOLICITUD DE PERMISO

Cáceres, 28 de octubre de 2019

Marta Martín Gilete, con DNI [redacted], alumna del PROGRAMA DE DOCTORADO EN LENGUAS Y CULTURAS (INTERUNIVERSITARIO) de la ESCUELA INTERNACIONAL DE POSGRADO de la UNIVERSIDAD DE EXTREMADURA, se encuentra realizando la investigación de la tesis doctoral *"La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2"*, tutelada y dirigida por la Profesora Dra. Ana Mª Piquer Píriz, y codirigida por la Profesora Dra. Fiona MacArthur.

SOLICITA:

Le sea concedido permiso para llevar a cabo la investigación pedagógica ***"La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2"*** en el centro de enseñanza no reglada [redacted] con alumnos de inglés como lengua extranjera que están cursando el nivel B2 del MCER durante el curso escolar 2019-2020.

Dicho estudio está basado en las aplicaciones de la Lingüística Cognitiva a la enseñanza del inglés como lengua extranjera, y tiene como fin mejorar la adquisición de vocabulario en lengua inglesa y, por tanto, su producción oral y escrita en el nivel B2. Esta investigación está basada en la idea de que el aprendizaje del vocabulario se ve facilitado a través de técnicas que ponen en relación los distintos sentidos figurados de una palabra con su significado básico.

Los objetivos que nos hemos planteado han sido los siguientes:

- De qué formas se podrían beneficiar alumnos españoles de Educación Secundaria con vistas al dominio del nivel B2 con la incorporación de la metáfora (producto lingüístico) en una programación didáctica de inglés como lengua extranjera en relación con los temas descritos en el MCER y Cambridge TESOL.
- Cómo responden los alumnos de Educación Secundaria ante la adquisición del lenguaje metafórico cuando éstas se presentan a través de la enseñanza basada en temas con el fin de ayudar a conseguir el dominio del nivel B2.
- Qué efectos tiene en la adquisición de vocabulario de alumnos de Educación Secundaria cursando el nivel B2 una instrucción basada en técnicas inspiradas en la lingüística cognitiva (LC) [por ejemplo, (i) elaboración semántica y/o etimológica: explicación verbal, agrupación conceptual, estrategias para adivinar; (ii) RFT (Respuesta Física Total), (iii) explicación visual] comparados con alumnos que no siguen este tipo de instrucción.
- Cuáles son los beneficios que aporta cada uno de los métodos de enseñanza basados en Lingüística Cognitiva empleados en el estudio para el desarrollo del vocabulario de los alumnos.

[47] Sensitive information is intentionally kept confidential or concealed to prevent unauthorised access, distribution or disclosure that could potentially cause harm to individuals or organisations.

Figure 7

Scanned Copy of the Research Authorisation Request Form: Page 2

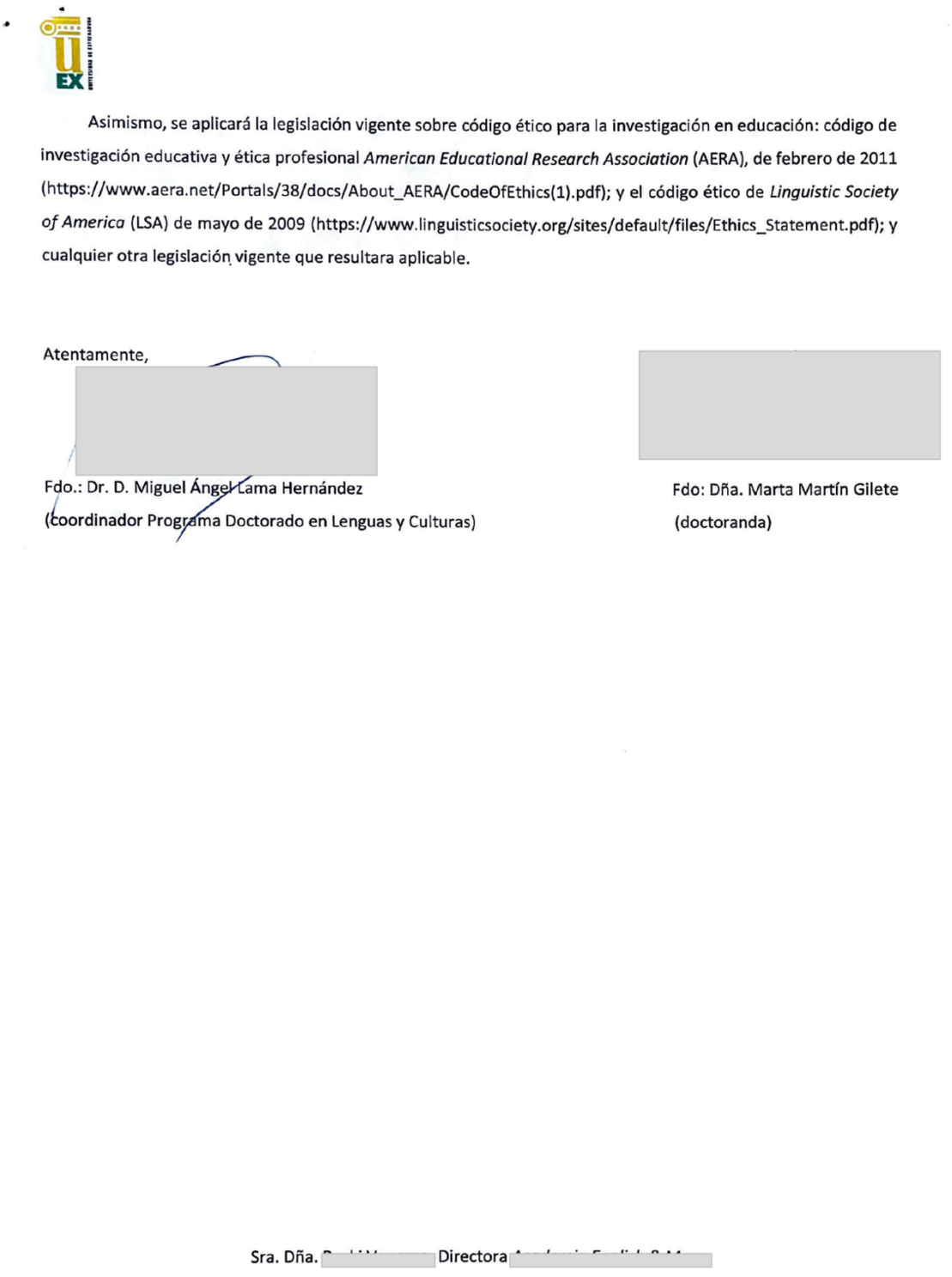
El procedimiento que el alumnado participante tendrá que realizar es llevar a cabo una prueba previa (pre-test) y una prueba posterior (post-test) para establecer los conocimientos previos y poder evaluar la potencial mejora de vocabulario empleado de forma oral y escrita en el nivel B2 de MCER. Además, se llevará a cabo la cumplimentación de cuestionarios por parte de los profesores y alumnos participantes. Asimismo, el alumnado participante en la investigación será dividido en grupos de control y experimentales, por lo que únicamente los grupos experimental serán sometidos a una instrucción con estrategias metodológicas de corte cognitivista para favorecer la adquisición de vocabulario específico en inglés. El mismo procedimiento que el alumnado tendrá que realizar será utilizado exclusivamente con finalidad de investigación sin ánimo de lucro. Al finalizar el estudio se informará al centro del resultado global del mismo si usted lo desea; también se informará a los participantes del mismo, pero nunca de su resultado personal, que se tratará con total confidencialidad. No existen riesgos asociados a la participación en el estudio y no supondrá ningún coste para los participantes.

De acuerdo con la normativa legal vigente, los resultados de las pruebas previas y posteriores realizadas, al igual que cualquier otra prueba de evaluación continua, se tratarán con total confidencialidad. Existirá un compromiso de que la confidencialidad de los datos que se puedan obtener en dicha investigación será escrupulosamente observada. Los datos se tratarán de modo leal para fines concretos de investigación y sobre la base de la ley; asimismo, los datos personales de los sujetos participantes no serán conocidos por la investigadora del estudio. El protocolo de recogida de datos será archivado, y a cada participante se le asignará una clave de tal modo que no pueda relacionarse la información obtenida con la identidad del sujeto. Se mantendrá un compromiso de no utilizar las muestras para otros estudios diferentes a los de este estudio y a no traspasar las muestras a otros posibles proyectos o equipos de investigación. De este modo, las pruebas previas y posteriores serán anonimizadas, asegurando la imposibilidad de inferir su identidad, para su estudio y potencial análisis ulterior.

Para todo lo explicado anteriormente y todo aquello que no esté previsto en el estudio, se aplicará la legislación vigente sobre protección de datos de carácter personal: artículo 18.4 de la *Constitución española* ("la ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos"); el artículo 8 de la *Carta de los Derechos Fundamentales de la Unión Europea* ("1. Toda persona tiene derecho a la protección de los datos de carácter personal que la conciernan. 2. Estos datos se tratarán de modo leal, para fines concretos y sobre la base de la ley. Toda persona tiene derecho a acceder a los datos recogidos que la conciernan y a su rectificación. 3. El respeto de estas normas quedarán sujeto al control de una autorizad independiente"); el artículo 16.1 del *Tratado de Funcionamiento de la Unión Europea* ("toda persona tiene derecho a la protección de los datos de carácter personal que le conciernan"); *Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo*, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de tatos personales y a la libre circulación de estos datos y por el que se deroga la *Directiva 95/46/CE* Reglamento general de protección de datos; *Ley Orgánica 3/2018*, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, BOE 294 de 6 de diciembre de 2018.

Figure 8

Scanned Copy of the Research Authorisation Request Form: Page 3



Asimismo, se aplicará la legislación vigente sobre código ético para la investigación en educación: código de investigación educativa y ética profesional *American Educational Research Association* (AERA), de febrero de 2011 (https://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics(1).pdf); y el código ético de *Linguistic Society of America* (LSA) de mayo de 2009 (https://www.linguisticsociety.org/sites/default/files/Ethics_Statement.pdf); y cualquier otra legislación vigente que resultara aplicable.

Atentamente,

Fdo.: Dr. D. Miguel Ángel Lama Hernández
(Coordinador Programa Doctorado en Lenguas y Culturas)

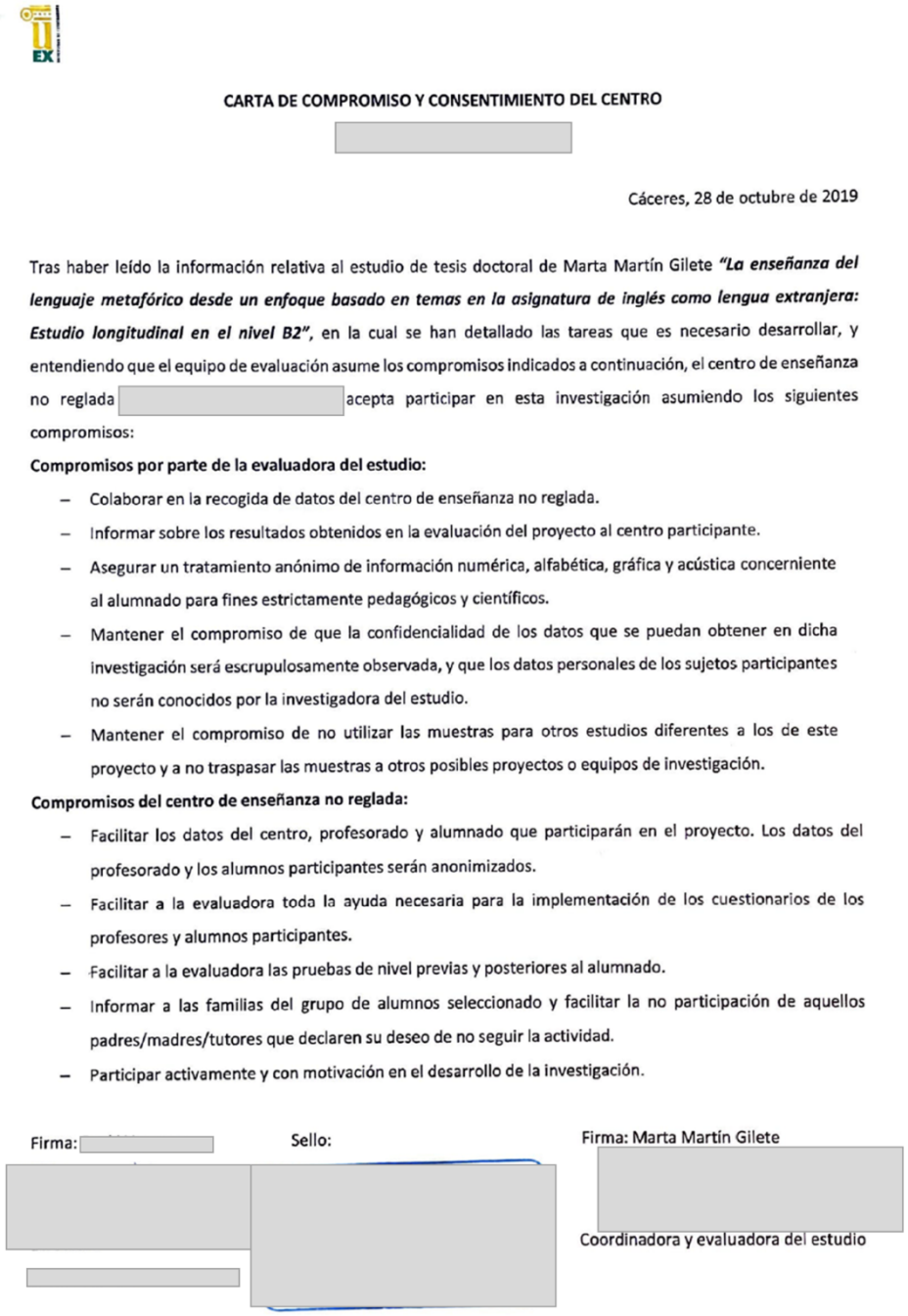
Fdo: Dña. Marta Martín Gilete
(doctoranda)

Sra. Dña. Directora

Appendix D: Research Commitment Agreement[48]

Figure 9

Scanned Copy of the Research Commitment Agreement



CARTA DE COMPROMISO Y CONSENTIMIENTO DEL CENTRO

[redacted]

Cáceres, 28 de octubre de 2019

Tras haber leído la información relativa al estudio de tesis doctoral de Marta Martín Gilete ***"La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2"***, en la cual se han detallado las tareas que es necesario desarrollar, y entendiendo que el equipo de evaluación asume los compromisos indicados a continuación, el centro de enseñanza no reglada [redacted] acepta participar en esta investigación asumiendo los siguientes compromisos:

Compromisos por parte de la evaluadora del estudio:

- Colaborar en la recogida de datos del centro de enseñanza no reglada.
- Informar sobre los resultados obtenidos en la evaluación del proyecto al centro participante.
- Asegurar un tratamiento anónimo de información numérica, alfabética, gráfica y acústica concerniente al alumnado para fines estrictamente pedagógicos y científicos.
- Mantener el compromiso de que la confidencialidad de los datos que se puedan obtener en dicha investigación será escrupulosamente observada, y que los datos personales de los sujetos participantes no serán conocidos por la investigadora del estudio.
- Mantener el compromiso de no utilizar las muestras para otros estudios diferentes a los de este proyecto y a no traspasar las muestras a otros posibles proyectos o equipos de investigación.

Compromisos del centro de enseñanza no reglada:

- Facilitar los datos del centro, profesorado y alumnado que participarán en el proyecto. Los datos del profesorado y los alumnos participantes serán anonimizados.
- Facilitar a la evaluadora toda la ayuda necesaria para la implementación de los cuestionarios de los profesores y alumnos participantes.
- Facilitar a la evaluadora las pruebas de nivel previas y posteriores al alumnado.
- Informar a las familias del grupo de alumnos seleccionado y facilitar la no participación de aquellos padres/madres/tutores que declaren su deseo de no seguir la actividad.
- Participar activamente y con motivación en el desarrollo de la investigación.

Firma: [redacted] Sello: [redacted]

Firma: Marta Martín Gilete

[redacted]

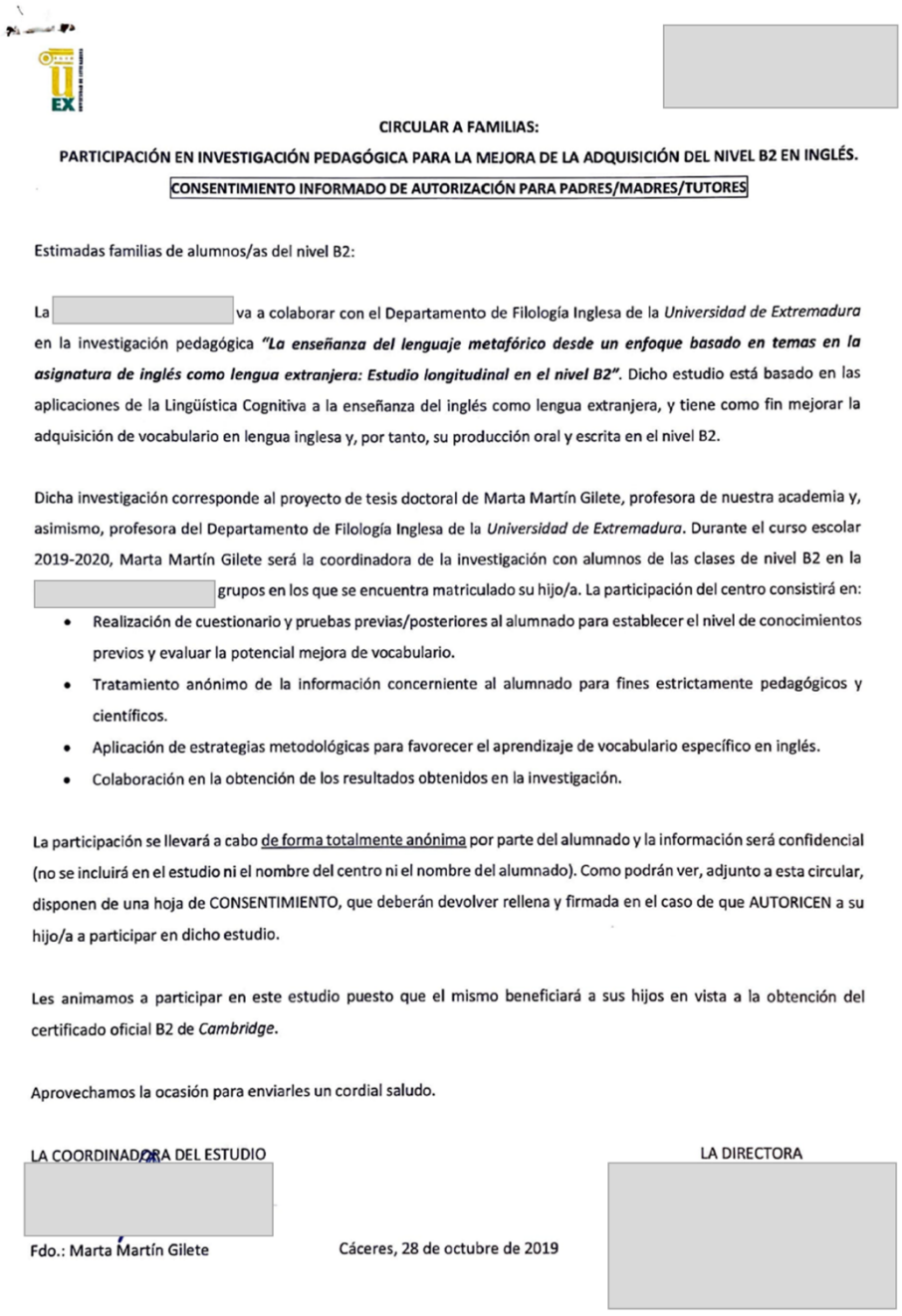
Coordinadora y evaluadora del estudio

[48] Sensitive information is intentionally kept confidential or concealed to prevent unauthorised access, distribution or disclosure that could potentially cause harm to individuals or organisations.

Appendix E: Informed Consent Form for Participants[49]

Figure 10

Scanned Copy of the Informed Consent Form for Participants: Page 1



CIRCULAR A FAMILIAS:

PARTICIPACIÓN EN INVESTIGACIÓN PEDAGÓGICA PARA LA MEJORA DE LA ADQUISICIÓN DEL NIVEL B2 EN INGLÉS.

CONSENTIMIENTO INFORMADO DE AUTORIZACIÓN PARA PADRES/MADRES/TUTORES

Estimadas familias de alumnos/as del nivel B2:

La [redacted] va a colaborar con el Departamento de Filología Inglesa de la *Universidad de Extremadura* en la investigación pedagógica ***"La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2"***. Dicho estudio está basado en las aplicaciones de la Lingüística Cognitiva a la enseñanza del inglés como lengua extranjera, y tiene como fin mejorar la adquisición de vocabulario en lengua inglesa y, por tanto, su producción oral y escrita en el nivel B2.

Dicha investigación corresponde al proyecto de tesis doctoral de Marta Martín Gilete, profesora de nuestra academia y, asimismo, profesora del Departamento de Filología Inglesa de la *Universidad de Extremadura*. Durante el curso escolar 2019-2020, Marta Martín Gilete será la coordinadora de la investigación con alumnos de las clases de nivel B2 en la [redacted] grupos en los que se encuentra matriculado su hijo/a. La participación del centro consistirá en:

- Realización de cuestionario y pruebas previas/posteriores al alumnado para establecer el nivel de conocimientos previos y evaluar la potencial mejora de vocabulario.
- Tratamiento anónimo de la información concerniente al alumnado para fines estrictamente pedagógicos y científicos.
- Aplicación de estrategias metodológicas para favorecer el aprendizaje de vocabulario específico en inglés.
- Colaboración en la obtención de los resultados obtenidos en la investigación.

La participación se llevará a cabo de forma totalmente anónima por parte del alumnado y la información será confidencial (no se incluirá en el estudio ni el nombre del centro ni el nombre del alumnado). Como podrán ver, adjunto a esta circular, disponen de una hoja de CONSENTIMIENTO, que deberán devolver rellena y firmada en el caso de que AUTORICEN a su hijo/a a participar en dicho estudio.

Les animamos a participar en este estudio puesto que el mismo beneficiará a sus hijos en vista a la obtención del certificado oficial B2 de *Cambridge*.

Aprovechamos la ocasión para enviarles un cordial saludo.

LA COORDINADORA DEL ESTUDIO

[redacted]

Fdo.: Marta Martín Gilete

Cáceres, 28 de octubre de 2019

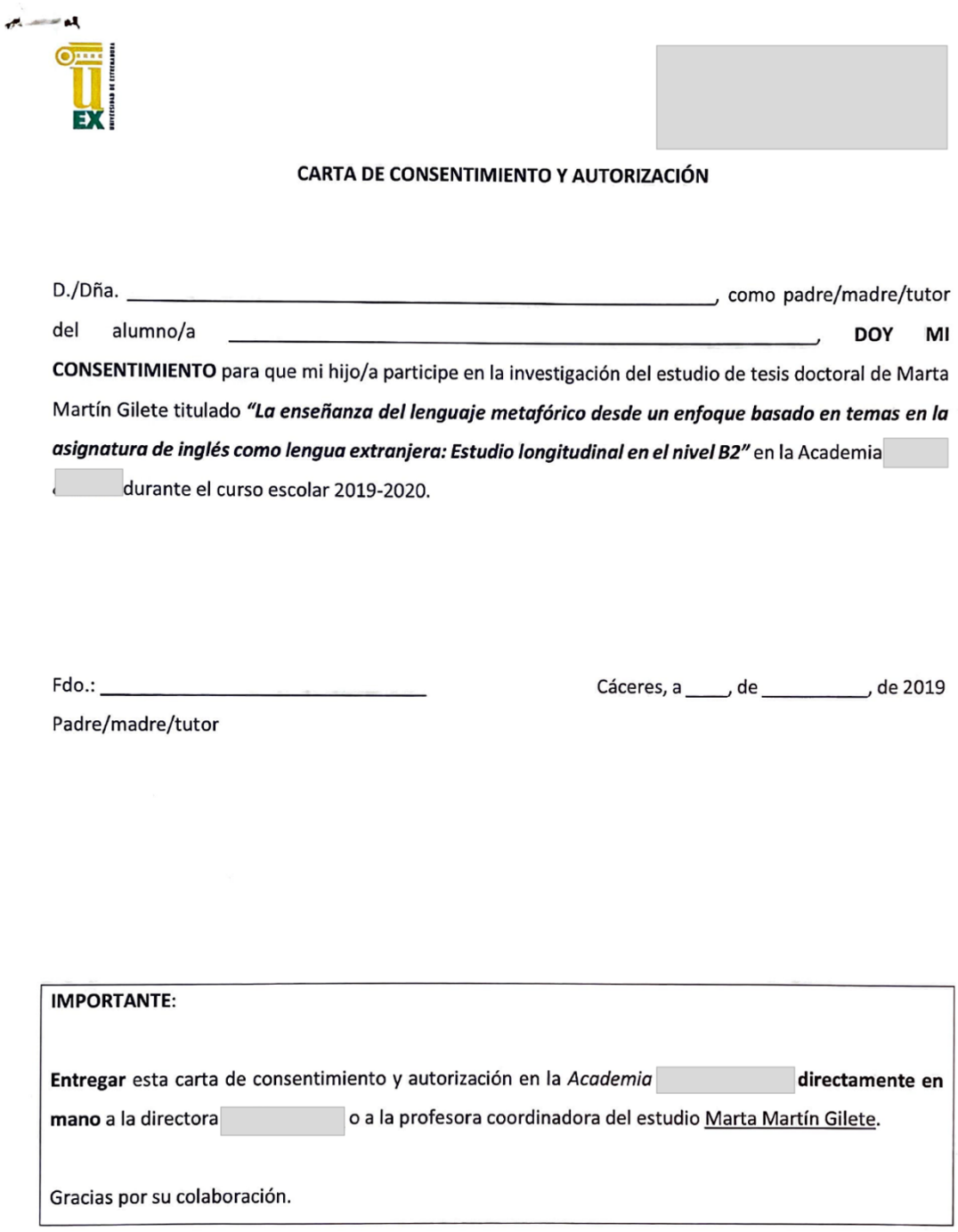
LA DIRECTORA

[redacted]

[49] Sensitive information is intentionally kept confidential or concealed to prevent unauthorised access, distribution or disclosure that could potentially cause harm to individuals or organisations.

Figure 11

Scanned Copy of the Informed Consent Form for Participants: Page 2



CARTA DE CONSENTIMIENTO Y AUTORIZACIÓN

D./Dña. _____, como padre/madre/tutor del alumno/a _____, **DOY MI CONSENTIMIENTO** para que mi hijo/a participe en la investigación del estudio de tesis doctoral de Marta Martín Gilete titulado ***"La enseñanza del lenguaje metafórico desde un enfoque basado en temas en la asignatura de inglés como lengua extranjera: Estudio longitudinal en el nivel B2"*** en la Academia durante el curso escolar 2019-2020.

Fdo.: _____ Cáceres, a ____, de _____, de 2019

Padre/madre/tutor

IMPORTANTE:

Entregar esta carta de consentimiento y autorización en la *Academia* **directamente en mano** a la directora o a la profesora coordinadora del estudio Marta Martín Gilete.

Gracias por su colaboración.

Appendix F:

Textbook Used in the L2 Classroom to Prepare Participants for Achieving B2 Level

Table 63

A Map of the Textbook Units Used for B2 Training

Unit title	Topic	Reference
Unit 1. A family affair	Home and daily routine	pp. 8–17
Unit 2. Leisure and pleasure	Hobbies and entertainment	pp. 18–27
Unit 3. Happy holidays?	Holidays and travel	pp. 30–39
Unit 4. Food, glorious food	Food and drink	pp. 40–49
Unit 5. Study time	Education and study	pp. 52–61
Unit 6. My first job	Work and relationships	pp. 62–71
Unit 7. High adventure	Sport and adventure	pp. 74–83
Unit 8. Dream of the stars	Careers and aspirations	pp. 84–93
Unit 9. Secrets of the mind	Happiness and the mind	pp. 96–105
Unit 10. Spend, spend, spend?	Shopping and leisure	pp. 106–115
Unit 11. Medical matters	Health and fitness	pp. 118–127
Unit 12. Animal kingdom	Human-animal relationships	pp. 128–137
Unit 13. House space	Housing and places	pp. 140–149
Unit 14. Fiesta!	Celebrations and special occasions	pp. 150–159

Source: Own elaboration from data extracted from the textbook used in class for B2 level preparation (Brook-Hart, 2014).

Appendix G: Collection of Input Texts from the Textbook and the Testing Measures

Table 64

Input Texts Collected from the Textbook: Unit 8 "Dream of the Stars" (Careers and Aspirations)

Input source	Text code[50]	Text title	Type of text	Exam paper	Exam task	No. of words
Written comprehension	TB8.WINPUT.R1	"YouTube millionaire celebrities"	Adapted article	*Reading*	Part 1	166
	TB8.WINPUT.R7	"Five young actors"	Adapted article	*Reading*	Part 7	773
	TB8.WINPUT.W1	"Pros and cons of being famous"	Essay question	*Writing*	Part 1	25
	TB8.WINPUT.W2	"TV talent contests"	Article question	*Writing*	Part 2	28
Oral comprehension	TB8.OINPUT.L2	"Ten minutes of fame"	Monologue	*Listening*	Part 2	810
	TB8.OINPUT.S4	"Talking about the dream of the stars"	Prompt questions	*Speaking*	Part 4	91

Source: Own elaboration from data extracted from Unit 8 of the textbook used in class (Brook-Hart, 2014, pp. 84–93).

[50] Each input text was assigned a unique code that consisted of several elements, including the textbook ("TB") and unit number abbreviation, the type of input discourse ("WINPUT" for written input, "OINPUT" for oral input), the corresponding exam paper abbreviation ("R" for *Reading and Use of English*, "W" for *Writing*, "L" for *Listening*, "S" for *Speaking*), and the task number of the exam paper.

Table 65

Input Texts Collected from the Textbook: Unit 9 "Secrets of the Mind" (Happiness and the Mind)

Input source	Text code	Text title	Type of text	Exam paper	Exam task	No. of words
Written comprehension	TB9.WINPUT.R5	"The secrets of happiness"	Adapted article	*Reading*	Part 5	775
	TB9.WINPUT.W1	"Pursuit of happiness"	Essay question	*Writing*	Part 1	16
	TB9.WINPUT.W2	"Childhood memories"	Article question	*Writing*	Part 2	20
Oral comprehension	TB9.OINPUT.L1	Untitled	Unrelated short texts	*Listening*	Part 1	1163
	TB9.OINPUT.S2	"Walking on air"	Prompt questions	*Speaking*	Part 2	68

Source: Own elaboration from data extracted from Unit 9 of the textbook used in class (Brook-Hart, 2014, pp. 96–105).

Table 66

Input Texts Collected from the Textbook: Unit 10 "Spend, Spend, Spend?" (Shopping and Leisure)

Input source	Text code	Text title	Type of text	Exam paper	Exam task	No. of words
Written comprehension	TB10.WINPUT.R2	"Shopping online vs. shopping locally"	Article	*Use of English*	Part 2	151
	TB10.WINPUT.R5	"My greatest influence"	Article	*Reading*	Part 5	850
	TB10.WINPUT.W1	"Shopping locally vs. shopping online"	Essay question	*Writing*	Part 1	23
	TB10.WINPUT.W2	"How to enjoy yourself"	Review question	*Writing*	Part 2	48
Oral comprehension	TB10.OINPUT.L4	Untitled	Interview	*Listening*	Part 4	796
	TB10.OINPUT.S1	"Spending money like water"	Prompt questions	*Speaking*	Part 1	42

Source: Own elaboration from data extracted from Unit 10 of the textbook used in class (Brook-Hart, 2014, pp. 106–115).

Table 67

Input Texts Collected from the Textbook: Unit 11 "Medical Matters" (Health and Fitness)

Input source	Text code	Text title	Type of text	Exam paper	Exam task	No. of words
Written comprehension	TB11.WINPUT.R3	"Is there a doctor on board?"	Article	*Use of English*	Part 3	147
	TB11.WINPUT.R6	"What's it like to study medicine?"	Adapted article	*Reading*	Part 6	540
	TB11.WINPUT.W1	"Healthy habits"	Essay question	*Writing*	Part 1	28
	TB11.WINPUT.W2	"Staying fit"	Article question	*Writing*	Part 2	28
Oral comprehension	TB11.OINPUT.L3	Untitled	Unrelated short monologues	*Listening*	Part 3	535
	TB11.OINPUT.S2	"Facing your problems"	Prompt questions	*Speaking*	Part 2	77

Source: Own elaboration from data extracted from Unit 11 of the textbook used in class (Brook-Hart, 2014, pp. 118–127).

Table 68

Input Texts Collected from the Testing Measures: Pre-Test and Post-Test

Testing measure	Input source	Text code	Type of text	Exam paper	Exam task	No. of words
Pre-test	Written comprehension	PREB2.WINPUT.W1	Essay question	*Writing*	Part 1	21
	Oral comprehension	PREB2.OINPUT.S4	Prompt questions	*Speaking*	Part 4	121
Post-test	Written comprehension	POSTB2.WINPUT.W1	Essay question	*Writing*	Part 1	22
	Oral comprehension	POSTB2.OINPUT.S4	Prompt questions	*Speaking*	Part 4	109

Source: Own elaboration from data extracted from the B2 tests used for the testing measures (UCLES, 2018, pp. 51–106).

Appendix H: Objectives and Format of the *B2 First for Schools* Examination

Reading and Use of English Test

The *Reading and Use of English* test consists of seven parts with seven different task types, resulting in a total of 52 questions with a weight of 40% toward the overall score. Although the *Reading* and *Use of English* components are combined in a single paper, they are evaluated independently. Test-takers are allowed 75 minutes to complete the entire test.

The *Reading and Use of English* test is widely recognised for its ability to evaluate L2 learners' reading skills, as well as their knowledge and use of grammar and vocabulary, through various texts that are commonly found in sources familiar to school-aged students. These texts may include magazine articles, fiction, or advertisements.

The exam paper is divided into two main sections, with Parts 1 to 4 focusing on texts that require applying lexico-grammatical knowledge through accompanying tasks. Parts 5 to 7 present a range of texts that are accompanied by reading comprehension tasks. Table 69 below provides an overview of the tasks included in the exam paper.

Table 69

Reading and Use of English Tasks in the B2 First for Schools Examination

Part	Task types	Focus
1	Multiple-choice cloze	Lexico-grammatical knowledge.
2	Open cloze	Lexico-grammatical knowledge.
3	Word formation	Lexico-grammatical knowledge.
4	Key word transformation	Lexico-grammatical knowledge.
5	Multiple choice	Understanding the details of a text, including the expression of opinion, attitude, purpose, main idea, detail, tone, implication, and gist. Ability to recognise meaning from context and follow text organisation features, such as exemplification, comparison, and reference.
6	Gapped text	Understanding the text structure (cohesion and coherence). Ability to follow the development of a long text.
7	Multiple matching	Finding specific information and detail and recognising opinion and attitude in a long text or group of short texts.

Source: Own elaboration from data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, pp. 8–9).

Writing Test

The *Writing* test accounts for 20% of the overall score and requires L2 learners to produce two different types of texts in English, each ranging from 140 to 190 words. As is well-known, this test evaluates L2 learners' ability to write effectively on topics that are relevant and interesting to teenage students. The time limit for this paper is 80 minutes.

Part 1 is a compulsory task requiring students to write an argumentative essay expressing their opinion on a given topic and supporting it with relevant reasons. In part 2, students can choose one question from a set of three. These questions may involve writing an article, email/letter, review, story, or essay in response to a set text question. Students may be asked to advise, compare, describe, explain, express opinions, or make recommendations.

Listening Test

The *Listening* test assesses L2 learners' comprehension of spoken English in various real-life contexts, such as radio programmes, podcasts, and informal conversations. This well-known paper comprises four parts, with a total of 30 questions that account for 20% of the overall score (see Table 70 below). Each exam part features one or more recorded texts and their corresponding comprehension questions. Test-takers can listen to each recording twice and are given approximately 40 minutes to complete the exam paper.

Table 70

Listening Tasks in the B2 First for Schools Examination

Part	Task types	Focus
1	Multiple-choice	Listening for gist, detail, function, purpose, attitude, opinion, genre, agreement, etc. in a series of unrelated short texts.
2	Sentence completion	Listening and finding specific information and stated opinion, from a single long text, and produce written answers by completing gapped sentences.
3	Multiple matching	Listening for general gist, purpose, feeling, main points, and detail; and matching an option to the correct speaker.
4	Multiple choice	Listening for opinion, attitude, gist, main idea, specific information.

Source: Own elaboration from data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, p. 53).

Speaking Test

The *Speaking* test comprises four parts designed to reflect authentic communication and counts 20% of the overall score. It is carried out in pairs for approximately 14 minutes,[51] with two examiners present to ensure more reliable marking. One examiner serves as the interlocutor, conducting the test and managing the interaction by overseeing questions or providing cues. The other examiner acts as the assessor, solely responsible for assessing the L2 learners' individual performance without intervening in the conversation.

The *Speaking* test is highly regarded for its capacity to assess L2 learners' ability to produce spontaneous oral language in a variety of contexts, such as conversing with the examiner, interacting with their partner, or discussing a given topic. As shown in table 71 below, this paper assesses students' proficiency in *Speaking* through four tasks that evaluate their ability to interact with other speakers, organise their ideas, exhibit accurate pronunciation, and use vocabulary and grammar effectively for communication at B2 level.

Table 71

Speaking Tasks in the B2 First for Schools Examination

Part	Task types	Focus
1	Interview with the examiner	Using social and interactional language by giving information about themselves and expressing opinions about various topics.
2	Individual long turn	Producing an extended piece of discourse on their own while comparing, describing, and expressing opinions.
3	Collaborative task	Engaging in a discussion and working towards a negotiated outcome of the task set while exchanging ideas, expressing, and justifying opinions, agreeing and/or disagreeing, suggesting, speculating, evaluating, reaching a decision through negotiation, etc.
4	Discussion	Engaging in a discussion based on the topic of the collaborative task in Part 3 while expressing and justifying opinions, agreeing, and/or disagreeing.

Source: Own elaboration from data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, p. 72).

[51] L2 learners may occasionally be tested in groups of three due to practical considerations. In such cases, the *Speaking* test duration may be extended up to 20 minutes, ensuring that each student has ample time to complete their tasks thoroughly.

Appendix I: Assessment Criteria for Measuring Performance in *B2 First for Schools*

The overall communicative language ability at B2 level is assessed by assigning equal weight to each exam paper (UCLES, 2015). After calculating the overall scores for each component, they are converted into *Cambridge English Scale* scores and are assigned a CEFR level (UCLES, 2019b). The specific *Cambridge English Scale* scores for each B2 exam paper are outlined in Table 72.

Table 72

The Cambridge English Scale Scores for B2 First for Schools

B2 exam paper	Exam paper score	*Cambridge English Scale* score[52]	CEFR level
Reading	37–42	180–190	C1 level
	24–36	160–179	B2 level
	16–23	140–159	B1 level
	10–15	122–139	not reported
Use of English	24–28	180–190	C1 level
	18–23	160–179	B2 level
	11–17	140–159	B1 level
	7–10	122–139	not reported
Writing	34–40	180–190	C1 level
	24–33	160–179	B2 level
	16–23	140–159	B1 level
	10–15	122–139	not reported
Listening	27–30	180–190	C1 level
	18–26	160–179	B2 level
	12–17	140–159	B1 level
	8–11	122–139	not reported
Speaking	54–60	180–190	C1 level
	36–53	160–179	B2 level
	24–35	140–159	B1 level
	14–23	122–139	not reported

Source: Own elaboration from data extracted from *The Cambridge English scale explained: A guide to converting practice test scores to Cambridge English scale scores* (UCLES, 2019a, pp. 7–52).

[52] The *B2 First for Schools* examination only reports scores above 140.

Assessment of Comprehension Tests

The *Reading*, *Use of English* and *Listening* exam papers use a system of objective grading where each item is marked as either correct or incorrect (UCLES, 2015). These exam papers are linked to the CEFR thresholds through a process of standards-setting (Rasch,1980), similar to how the *Cambridge English Scale* is aligned to the CEFR. The *Reading and Use of English* test has a total of 70 possible marks, while the *Listening* test has 30 possible marks. Table 73 outlines the marking criteria used for the *Reading*, *Use of English*, and *Listening* tests.

Table 73

Assessment Criteria for B2 First for Schools: Reading, Use of English, and Listening Tests

Exam paper	Part	No. questions	No. of possible marks
Reading	1	8	8
	5	6	12
	6	6	12
	7	10	10
Use of English	2	8	8
	3	8	8
	4	6	12
Listening	1	8	8
	2	10	10
	3	5	5
	4	7	7

Source: Own elaboration from data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, pp. 7–52).

Individual performance for each component is calculated as follows. For the *Reading and Use of English* paper, the test score of each participant is determined by combining the marks received for correct answers in the seven parts of the exam paper. For the *Listening* paper, the test score of each participant was calculated by combining the marks received for correct answers in the four parts of the test.

Assessment of Production Tests

The *Writing* and *Speaking* tests are evaluated using analytic scales that encompass various assessment criteria explicitly linked to the CEFR. This approach guarantees that the marking process for these components is objective and consistent. Table 74 provides an overview of the assessment criteria for grading the *Writing* and *Speaking* tests.

Table 74

Assessment Criteria for B2 First for Schools: Writing and Speaking Tests

Exam paper	Criteria	No. of possible marks	Descriptors
Writing	Content	5	Task fulfillment.
	Communicative achievement	5	Appropriate register and writing for the task.
	Organisation	5	Coherence and cohesion.
	Language	5	Range and accuracy of vocabulary and grammar.
Speaking	Grammar and vocabulary	10	Range and accuracy of vocabulary and grammar.
	Discourse management	10	Length appropriateness, hesitation and repetition, pertinence, and use of discourse markers.
	Pronunciation	10	Intonation, word stress, and articulation and accuracy of sounds.
	Interactive communication	10	Discussion management, reciprocity, and support.
	Global achievement	20	Overall communication performance.

Source: Own elaboration from data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, pp. 27–71).

Assessment of Writing Test. The *Writing* paper is evaluated on a scale of 0 to 40 marks, with equal weight assigned to each piece of writing. The assessment of each piece is based on four criteria: content, communicative achievement (comm. achv.), organisation, and language. A score from 0 (lowest) to 5 (highest) is given to each criterion. The marks obtained in the two questions are then combined to calculate the overall score achieved. The band descriptors provide clear guidance on the expected levels of performance (UCLES, 2020). A score band of 3 or higher in the *Writing* test demonstrates B2 level of English language proficiency, while bands 1 and 2 correspond to B1 level. These descriptors are useful in

identifying areas of strength and weakness for each of the four assessment criteria. Table 75 provides a detailed description of the band descriptors used to grade the *Writing* paper.

Table 75

Assessment Scale for B2 First for Schools: Writing Test

Mark	Content	Communicative achievement	Organisation	Language
5	All content is relevant to the task. Target reader is fully informed.	Uses the conventions of the communicative task effectively to hold the target reader's attention and communicate straightforward and complex ideas, as appropriate.	Text is well organised and coherent, using a variety of cohesive devices and organisational patterns to generally good effect.	Uses a range of vocabulary, including less common lexis, appropriately. Uses a range of simple and complex grammatical forms with control and flexibility. Occasional errors may be present but do not impede communication.
4	Performance shares features of Bands 3 and 5.			
3	Minor irrelevances and/or omissions may be present. Target reader is on the whole informed.	Uses the conventions of the communicative task to hold the target reader's attention and communicate straightforward ideas.	Text is generally well organised and coherent, using a variety of linking words and cohesive devices.	Uses a range of everyday vocabulary appropriately, with occasional inappropriate use of less common lexis. Uses a range of simple and some complex grammatical forms with a good degree of control. Errors do not impede communication.
2	Performance shares features of Bands 1 and 3.			
1	Irrelevances and misinterpretation of task may be present. Target reader is minimally informed.	Uses the conventions of the communicative task in generally appropriate ways to communicate straightforward ideas.	Text is connected and coherent, using basic linking words and a limited number of cohesive devices.	Uses everyday vocabulary generally appropriately, while occasionally overusing certain lexis. Uses simple grammatical forms with a good degree of control. While errors are noticeable, meaning can still be determined.
0	Content is totally irrelevant. Target reader is not informed.	Performance below Band 1.		

Source: Data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, p. 34).

Assessment of Speaking Test. The *Speaking* paper is assessed out of 60 marks. Despite taking the exam in pairs, each L2 learner is evaluated based on their individual performance. The examiner who acts as the assessor grades the learner's individual performance according to a specific *Speaking* assessment scale, which comprises four criteria: grammar and vocabulary, discourse management (discourse mgmt.), pronunciation (pron.), and interactive communication (interactive comm.). Each criterion is marked from 0 (lowest) to 5 (highest) using the descriptors presented in Table 76.

The examiner who acts as the interlocutor provides a mark on the global achievement scale from 0 (lowest) to 5 (highest) for the overall performance. The test score of each participant is determined by doubling the 0–5 marks obtained for each of the assessment criteria. Next, the 0–5 mark received for global achievement is multiplied by 4. Finally, marks for all criteria are combined to calculate the overall score.

Table 76

Assessment Scale for B2 First for Schools: Speaking Test

Mark	Grammar and vocabulary	Discourse management	Pronunciation	Interactive communication	Global achievement
5	Shows a good degree of control of a range of simple and some complex grammatical forms. Uses a range of appropriate vocabulary to give and exchange views on a wide range of familiar topics.	Produces extended stretches of language with very little hesitation. Contributions are relevant and there is a clear organisation of ideas. Uses a range of cohesive devices and discourse markers.	Is intelligible. Intonation is appropriate. Sentence and word stress is accurately placed. Individual sounds are articulated clearly.	Initiates and responds appropriately, linking contributions to those of other speakers. Maintains and develops the interaction and negotiates towards an outcome.	Handles communication on a range of familiar topics, with very little hesitation. Uses accurate and appropriate linguistic resources to express ideas and produce extended discourse that is generally coherent.
4	Performance shares features of Bands 3 and 5.				
3	Shows a good degree of control of simple grammatical	Produces extended stretches of language	Is intelligible. Intonation is generally appropriate.	Initiates and responds appropriately. Maintains and develops the	Handles communication on familiar topics, despite some hesitation.

	forms, and attempts some complex grammatical forms. Uses a range of appropriate vocabulary to give and exchange views on a range of familiar topics.	despite some hesitation. Contributions are relevant and there is very little repetition. Uses a range of cohesive devices.	Sentence and word stress is generally accurately placed. Individual sounds are generally articulated clearly.	interaction and negotiates towards an outcome with very little support.	Organises extended discourse but occasionally produces utterances that lack coherence, and some inaccuracies and inappropriate usage occur.
2	Performance shares features of Bands 1 and 3.				
1	Shows a good degree of control of simple grammatical forms. Uses a range of appropriate vocabulary when talking about everyday situations.	Produces responses which are extended beyond short phrases, despite hesitation. Contributions are mostly relevant, despite some repetition. Uses basic cohesive devices.	Is mostly intelligible, and has some control of phonological features at both utterance and word levels.	Initiates and responds appropriately. Keeps the interaction going with very little prompting and support.	Handles communication in everyday situations, despite hesitation. Constructs longer utterances but is not able to use complex language except in well-rehearsed utterances.
0	Performance below Band 1.				

Source: Data extracted from *B2 First for Schools: Handbook for teachers* (UCLES, 2019a, p. 82).

Appendix J: Collection of Output Texts from the Testing Measures

Table 77

Output Texts Collected from the Testing Measures: Written Discourse

Participants	Written output texts[53]			
	Pre-test text codes	No. of words	Post-test text codes	No. of words
ST0101	PREB2.WOUTPUT.W1.ST0101	192	POSTB2.WOUTPUT.W1.ST0101	201
ST0102	PREB2.WOUTPUT.W1.ST0102	181	POSTB2.WOUTPUT.W1.ST0102	170
ST0103	PREB2.WOUTPUT.W1.ST0103	182	POSTB2.WOUTPUT.W1.ST0103	177
ST0104	PREB2.WOUTPUT.W1.ST0104	179	POSTB2.WOUTPUT.W1.ST0104	192
ST0105	PREB2.WOUTPUT.W1.ST0105	175	POSTB2.WOUTPUT.W1.ST0105	241
ST0106	PREB2.WOUTPUT.W1.ST0106	169	POSTB2.WOUTPUT.W1.ST0106	164
ST0107	PREB2.WOUTPUT.W1.ST0107	220	POSTB2.WOUTPUT.W1.ST0107	167
ST0108	PREB2.WOUTPUT.W1.ST0108	149	POSTB2.WOUTPUT.W1.ST0108	176
ST0109	PREB2.WOUTPUT.W1.ST0109	174	POSTB2.WOUTPUT.W1.ST0109	162
ST0110	PREB2.WOUTPUT.W1.ST0110	169	POSTB2.WOUTPUT.W1.ST0110	193
ST0201	PREB2.WOUTPUT.W1.ST0201	126	POSTB2.WOUTPUT.W1.ST0201	171
ST0202	PREB2.WOUTPUT.W1.ST0202	179	POSTB2.WOUTPUT.W1.ST0202	170
ST0203	PREB2.WOUTPUT.W1.ST0203	187	POSTB2.WOUTPUT.W1.ST0203	177
ST0204	PREB2.WOUTPUT.W1.ST0204	170	POSTB2.WOUTPUT.W1.ST0204	209
ST0205	PREB2.WOUTPUT.W1.ST0205	176	POSTB2.WOUTPUT.W1.ST0205	205
ST0206	PREB2.WOUTPUT.W1.ST0206	162	POSTB2.WOUTPUT.W1.ST0206	181
ST0207	PREB2.WOUTPUT.W1.ST0207	206	POSTB2.WOUTPUT.W1.ST0207	218
ST0208	PREB2.WOUTPUT.W1.ST0208	170	POSTB2.WOUTPUT.W1.ST0208	252
ST0209	PREB2.WOUTPUT.W1.ST0209	175	POSTB2.WOUTPUT.W1.ST0209	186
ST0210	PREB2.WOUTPUT.W1.ST0210	157	POSTB2.WOUTPUT.W1.ST0210	192
ST0301	PREB2.WOUTPUT.W1.ST0301	139	POSTB2.WOUTPUT.W1.ST0301	159
ST0302	PREB2.WOUTPUT.W1.ST0302	166	POSTB2.WOUTPUT.W1.ST0302	242
ST0303	PREB2.WOUTPUT.W1.ST0303	148	POSTB2.WOUTPUT.W1.ST0303	148
ST0304	PREB2.WOUTPUT.W1.ST0304	115	POSTB2.WOUTPUT.W1.ST0304	124
ST0305	PREB2.WOUTPUT.W1.ST0305	151	POSTB2.WOUTPUT.W1.ST0305	178
ST0306	PREB2.WOUTPUT.W1.ST0306	172	POSTB2.WOUTPUT.W1.ST0306	213
ST0307	PREB2.WOUTPUT.W1.ST0307	175	POSTB2.WOUTPUT.W1.ST0307	203
ST0308	PREB2.WOUTPUT.W1.ST0308	158	POSTB2.WOUTPUT.W1.ST0308	183
ST0309	PREB2.WOUTPUT.W1.ST0309	124	POSTB2.WOUTPUT.W1.ST0309	171
ST0310	PREB2.WOUTPUT.W1.ST0310	171	POSTB2.WOUTPUT.W1.ST0310	203
ST0401	PREB2.WOUTPUT.W1.ST0401	221	POSTB2.WOUTPUT.W1.ST0401	227
ST0402	PREB2.WOUTPUT.W1.ST0402	132	POSTB2.WOUTPUT.W1.ST0402	227
ST0403	PREB2.WOUTPUT.W1.ST0403	177	POSTB2.WOUTPUT.W1.ST0403	215
ST0404	PREB2.WOUTPUT.W1.ST0404	190	POSTB2.WOUTPUT.W1.ST0404	158
ST0405	PREB2.WOUTPUT.W1.ST0405	156	POSTB2.WOUTPUT.W1.ST0405	180
ST0406	PREB2.WOUTPUT.W1.ST0406	170	POSTB2.WOUTPUT.W1.ST0406	179
ST0407	PREB2.WOUTPUT.W1.ST0407	225	POSTB2.WOUTPUT.W1.ST0407	199
ST0408	PREB2.WOUTPUT.W1.ST0408	183	POSTB2.WOUTPUT.W1.ST0408	178
ST0409	PREB2.WOUTPUT.W1.ST0409	171	POSTB2.WOUTPUT.W1.ST0409	213
ST0410	PREB2.WOUTPUT.W1.ST0410	155	POSTB2.WOUTPUT.W1.ST0410	170

[53] Each written output text was assigned a unique code consisting of different elements: the abbreviation of testing measure ("PRE" for pre-test, "POST" for post-test) followed by "B2" to indicate the proficiency level, the code for the type of output discourse ("WOUTPUT" for written output), the task completed in the test ("W1" for *Writing* Part 1), and the corresponding student's code.

Table 78

Output Texts Collected from the Testing Measures: Oral Discourse

Participants	Oral output texts[54]			
	Pre-test text codes	No. of words	Post-test text codes	No. of words
ST0101	PREB2.SOUTPUT.S4.ST0101	138	POSTB2.SOUTPUT.S4.ST0101	208
ST0102	PREB2.SOUTPUT.S4.ST0102	281	POSTB2.SOUTPUT.S4.ST0102	343
ST0103	PREB2.SOUTPUT.S4.ST0103	279	POSTB2.SOUTPUT.S4.ST0103	272
ST0104	PREB2.SOUTPUT.S4.ST0104	187	POSTB2.SOUTPUT.S4.ST0104	187
ST0105	PREB2.SOUTPUT.S4.ST0105	412	POSTB2.SOUTPUT.S4.ST0105	312
ST0106	PREB2.SOUTPUT.S4.ST0106	150	POSTB2.SOUTPUT.S4.ST0106	190
ST0107	PREB2.SOUTPUT.S4.ST0107	274	POSTB2.SOUTPUT.S4.ST0107	230
ST0108	PREB2.SOUTPUT.S4.ST0108	283	POSTB2.SOUTPUT.S4.ST0108	247
ST0109	PREB2.SOUTPUT.S4.ST0109	225	POSTB2.SOUTPUT.S4.ST0109	244
ST0110	PREB2.SOUTPUT.S4.ST0110	162	POSTB2.SOUTPUT.S4.ST0110	160
ST0201	PREB2.SOUTPUT.S4.ST0201	247	POSTB2.SOUTPUT.S4.ST0201	258
ST0202	PREB2.SOUTPUT.S4.ST0202	158	POSTB2.SOUTPUT.S4.ST0202	124
ST0203	PREB2.SOUTPUT.S4.ST0203	301	POSTB2.SOUTPUT.S4.ST0203	280
ST0204	PREB2.SOUTPUT.S4.ST0204	191	POSTB2.SOUTPUT.S4.ST0204	265
ST0205	PREB2.SOUTPUT.S4.ST0205	178	POSTB2.SOUTPUT.S4.ST0205	108
ST0206	PREB2.SOUTPUT.S4.ST0206	225	POSTB2.SOUTPUT.S4.ST0206	195
ST0207	PREB2.SOUTPUT.S4.ST0207	155	POSTB2.SOUTPUT.S4.ST0207	126
ST0208	PREB2.SOUTPUT.S4.ST0208	354	POSTB2.SOUTPUT.S4.ST0208	206
ST0209	PREB2.SOUTPUT.S4.ST0209	175	POSTB2.SOUTPUT.S4.ST0209	146
ST0210	PREB2.SOUTPUT.S4.ST0210	252	POSTB2.SOUTPUT.S4.ST0210	134
ST0301	PREB2.SOUTPUT.S4.ST0301	154	POSTB2.SOUTPUT.S4.ST0301	137
ST0302	PREB2.SOUTPUT.S4.ST0302	143	POSTB2.SOUTPUT.S4.ST0302	105
ST0303	PREB2.SOUTPUT.S4.ST0303	203	POSTB2.SOUTPUT.S4.ST0303	151
ST0304	PREB2.SOUTPUT.S4.ST0304	183	POSTB2.SOUTPUT.S4.ST0304	135
ST0305	PREB2.SOUTPUT.S4.ST0305	163	POSTB2.SOUTPUT.S4.ST0305	121
ST0306	PREB2.SOUTPUT.S4.ST0306	207	POSTB2.SOUTPUT.S4.ST0306	172
ST0307	PREB2.SOUTPUT.S4.ST0307	153	POSTB2.SOUTPUT.S4.ST0307	170
ST0308	PREB2.SOUTPUT.S4.ST0308	213	POSTB2.SOUTPUT.S4.ST0308	148
ST0309	PREB2.SOUTPUT.S4.ST0309	201	POSTB2.SOUTPUT.S4.ST0309	112
ST0310	PREB2.SOUTPUT.S4.ST0310	230	POSTB2.SOUTPUT.S4.ST0310	147
ST0401	PREB2.SOUTPUT.S4.ST0401	171	POSTB2.SOUTPUT.S4.ST0401	212
ST0402	PREB2.SOUTPUT.S4.ST0402	162	POSTB2.SOUTPUT.S4.ST0402	141
ST0403	PREB2.SOUTPUT.S4.ST0403	218	POSTB2.SOUTPUT.S4.ST0403	225
ST0404	PREB2.SOUTPUT.S4.ST0404	219	POSTB2.SOUTPUT.S4.ST0404	152
ST0405	PREB2.SOUTPUT.S4.ST0405	182	POSTB2.SOUTPUT.S4.ST0405	176
ST0406	PREB2.SOUTPUT.S4.ST0406	234	POSTB2.SOUTPUT.S4.ST0406	238
ST0407	PREB2.SOUTPUT.S4.ST0407	217	POSTB2.SOUTPUT.S4.ST0407	169
ST0408	PREB2.SOUTPUT.S4.ST0408	160	POSTB2.SOUTPUT.S4.ST0408	132
ST0409	PREB2.SOUTPUT.S4.ST0409	256	POSTB2.SOUTPUT.S4.ST0409	128
ST0410	PREB2.SOUTPUT.S4.ST0410	190	POSTB2.SOUTPUT.S4.ST0410	117

[54] Each oral output text was assigned a unique code comprised different components: the abbreviation of testing measure ("PRE" for pre-test, "POST" for post-test) followed by "B2" to indicate the proficiency level, the code for the type of output discourse ("SOUTPUT" for oral output), the task completed in the test ("S4" for *Speaking* Part 4), and the corresponding student's code.

Appendix K: Adapted Version of the VOICE (2021) Transcription Method

Figure 12

An Example Transcript of Learner Oral Discourse

<DISCOURSE: OUTPUT MODE: ORAL PROVENANCE: LEARNER PRODUCTION TEXT SOURCE: B2 FIRST PRE-TEST TEST: SPEAKING TASK: DISCUSSION GROUP: GR0102 PARTICIPANT: ST0206>
<{Discussion on helping children with homework}>
<{Do your parents give you a lot of help with your homework? What do they help with?}>
me too because my parents work and in the afternoon they have to do different things so they don't have time to help me (.)
<{Some people say that children shouldn't do any homework during the school week. Do you agree?}>
I'm not agree with this opinion because I believe that is very important to do homeworks because the the things that you learn at mornings in the school then you have to practise at home (.)
I'm totally agree with you because from my point of view I think that for example during the week of exams teachers shouldn't have to send homework because students have to study in the afternoon and they don't have time (.)
<{Do you think it would be a good idea for students to do homework online so that the teachers can see what they've done?}>
I in my view it could have a good part and also a bad part because I think that students are it's easier to do in a notebook because you have your books but also you can do online and is it has different advantages for example you can search information on internet (.)
I'm totally agree with you because I think that technology is the future so childrens have to practise with doing exercise online (.)
<{Do you think it's true that giving children prizes is the best way to encourage them to work harder?}>
I couldn't agree more I think that do homeworks and study is an obligation and you have to do it but for example there are children that have more diffi- have more and for example if you don't study and you know that you have that if you study you will have a little prize probably you enjoy what you study (.)
</DISCOURSE: OUTPUT MODE: SPEAKING PROVENANCE: LEARNER PRODUCTION TEXT SOURCE: B2 FIRST PRE-TEST TEST: SPEAKING TASK: DISCUSSION GROUP: GR0102 PARTICIPANT: ST0206>

Source: Example extracted from the transcript of ST0206's pre-test oral production.

Appendix L: CL-Inspired Teaching Materials for B2 Level EFL Instruction[55]

Figure 13

A Pictorial Example of the CM CAREER IS A BUILDING (Unit 8 "Dream of the Stars")

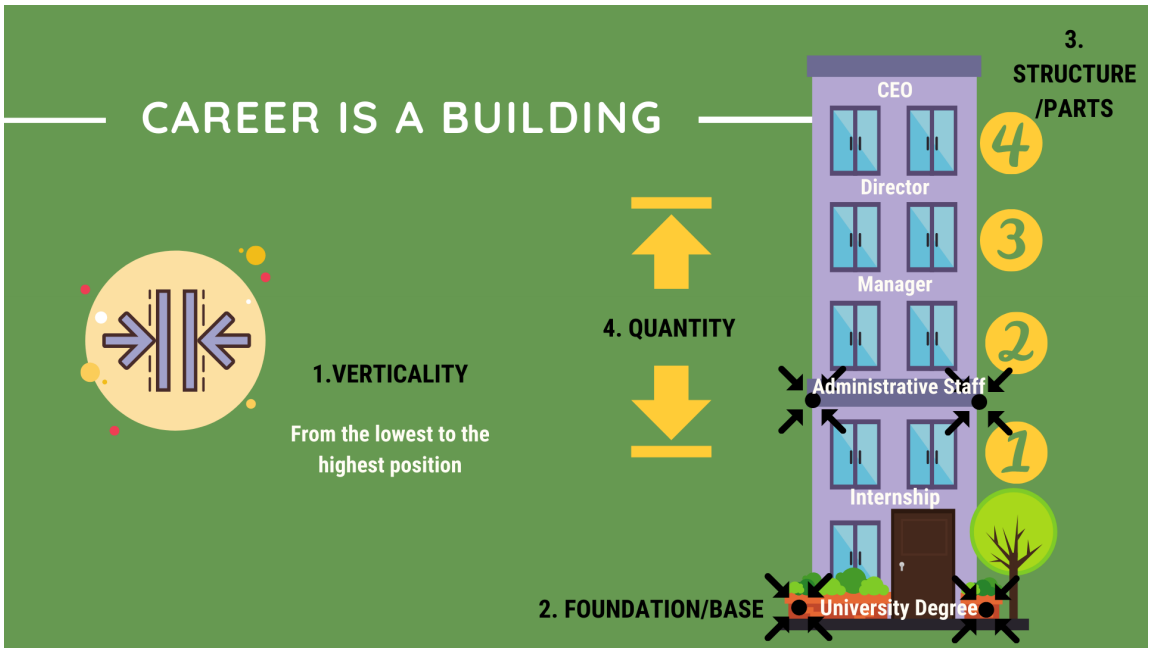


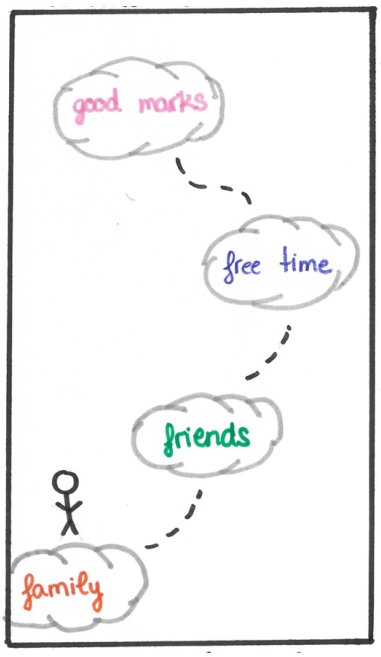
Figure 14

A Pictorial Example of the CM HAPPINESS IS VERTICALITY (Unit 9 "Secrets of the Mind")

[55] The CL-inspired teaching materials included in this PhD dissertation are licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Figure 15

Learners' Meaning-Meaning Connections of "Walking on Air" (Unit 9 "Secrets of the Mind"): Example 1



Source: Drawing created by experimental group participant ST0103.

Figure 16

Learners' Meaning-Meaning Connections of "Walking on Air" (Unit 9 "Secrets of the Mind"): Example 2

Source: Drawing created by experimental group participant ST0206.

Figure 17

Example Material Used to Introduce Metalanguage: Metaphors

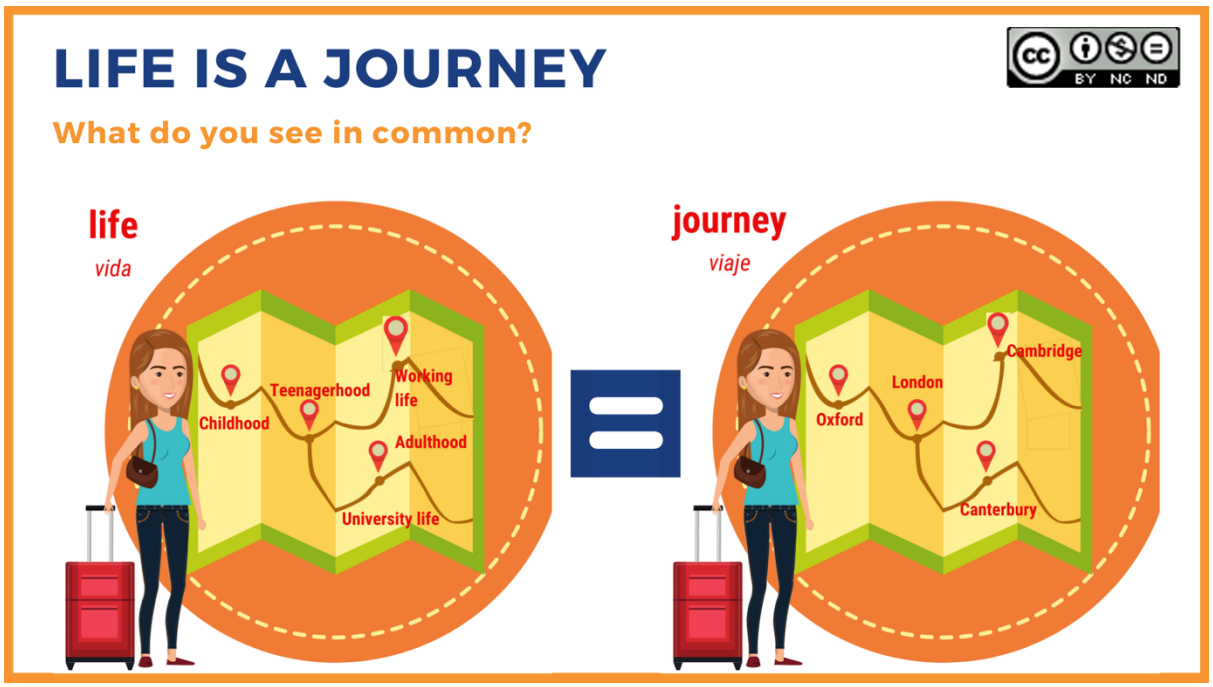


Figure 18

Example Material Used to Introduce Metalanguage: Concrete vs. Abstract Concepts

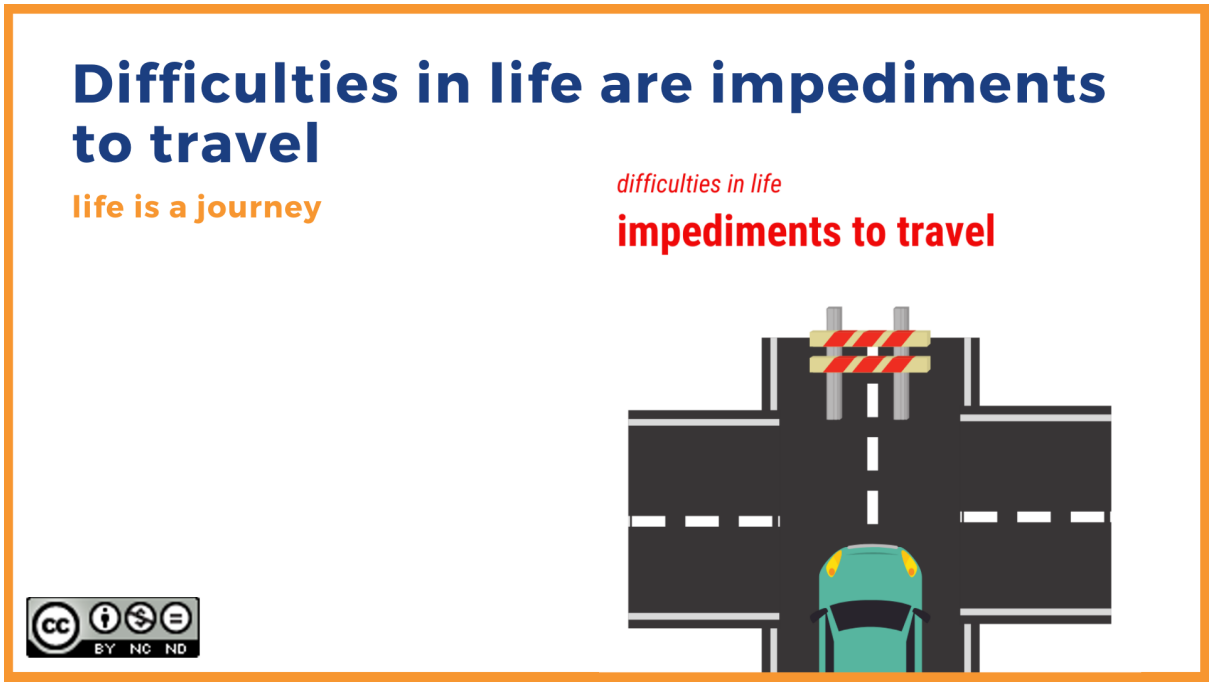


Figure 19

Example Material Used to Introduce Metalanguage: Basic Meanings

Figure 20

Example Material Used to Introduce Metalanguage: Extended Meanings

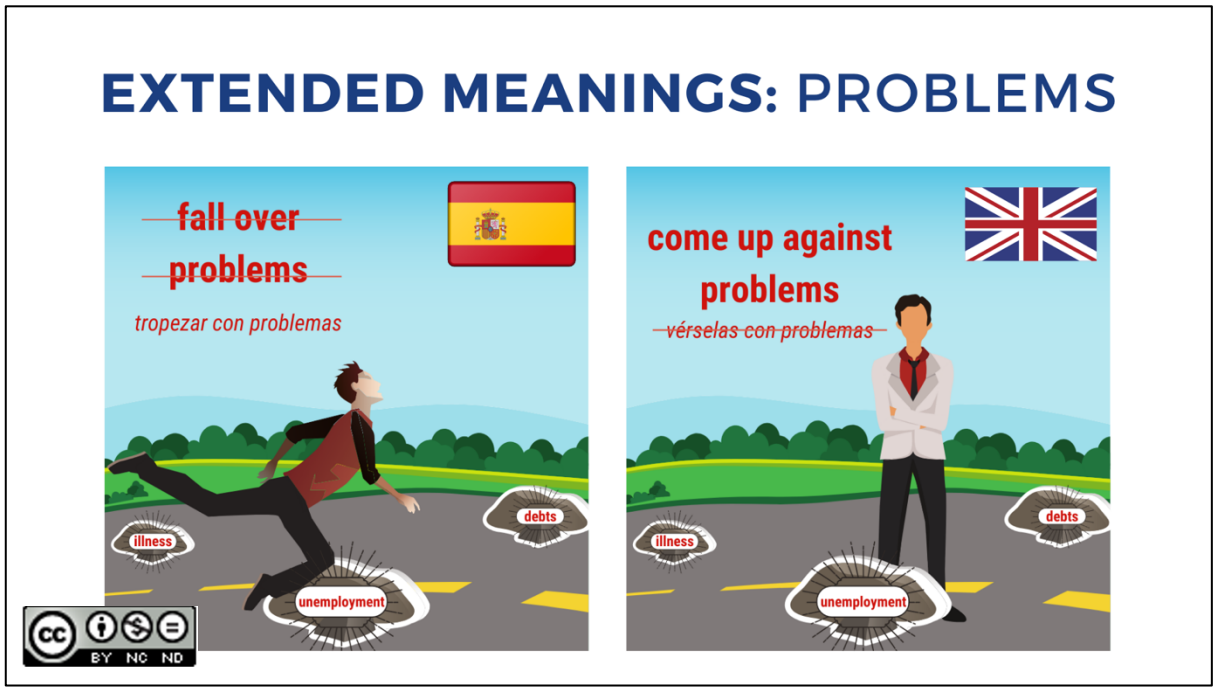
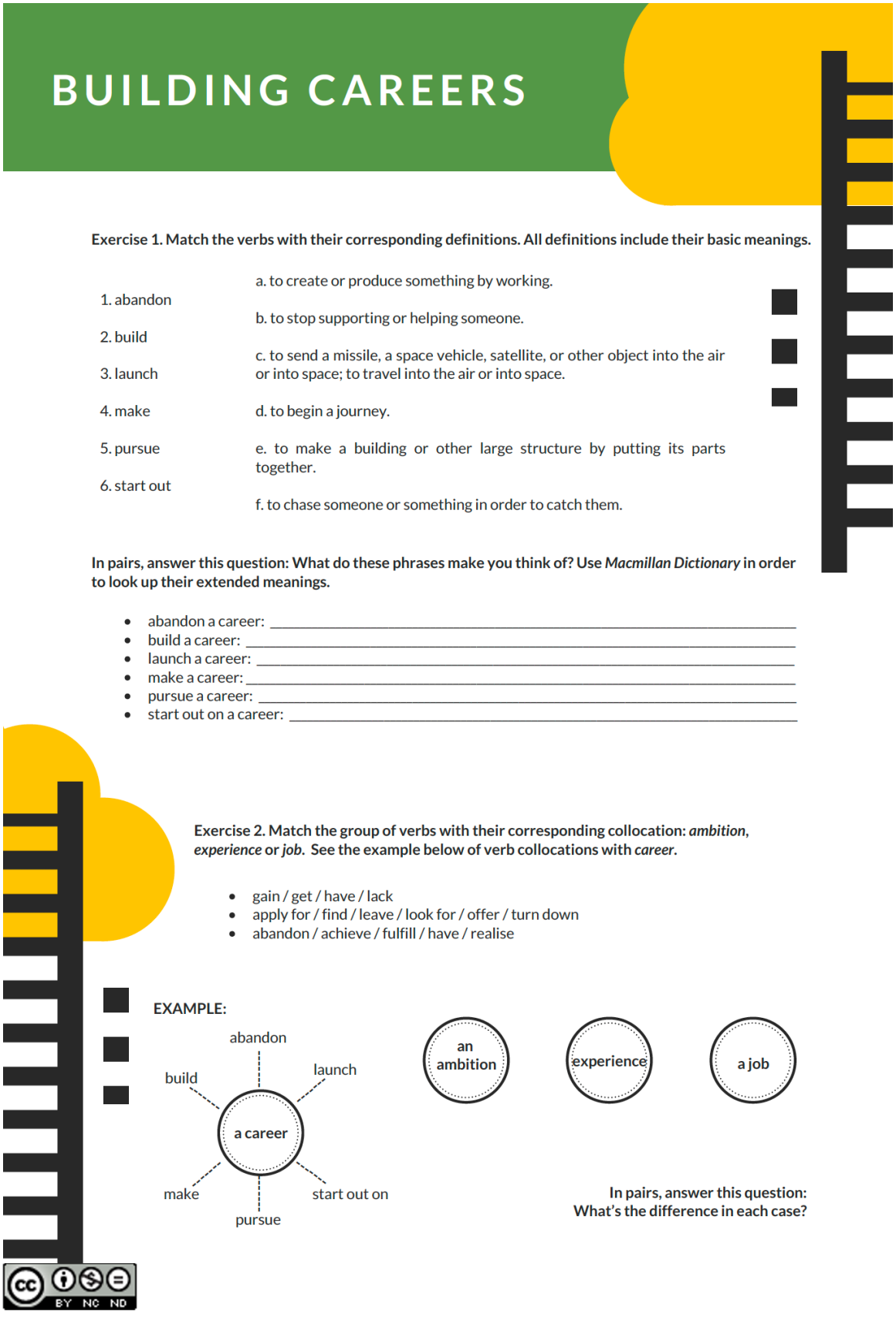


Figure 21

Adaptation of Proposals for CL-Oriented Pedagogies: Learning Collocations with "Ambition", "Experience", and "Job" in Unit 8 "Dream of the Stars" (Page 1)



BUILDING CAREERS

Exercise 1. Match the verbs with their corresponding definitions. All definitions include their basic meanings.

1. abandon	a. to create or produce something by working.
2. build	b. to stop supporting or helping someone.
3. launch	c. to send a missile, a space vehicle, satellite, or other object into the air or into space; to travel into the air or into space.
4. make	d. to begin a journey.
5. pursue	e. to make a building or other large structure by putting its parts together.
6. start out	f. to chase someone or something in order to catch them.

In pairs, answer this question: What do these phrases make you think of? Use *Macmillan Dictionary* in order to look up their extended meanings.

- abandon a career: _____
- build a career: _____
- launch a career: _____
- make a career: _____
- pursue a career: _____
- start out on a career: _____

Exercise 2. Match the group of verbs with their corresponding collocation: *ambition*, *experience* or *job*. See the example below of verb collocations with *career*.

- gain / get / have / lack
- apply for / find / leave / look for / offer / turn down
- abandon / achieve / fulfill / have / realise

EXAMPLE:

an ambition

experience

a job

In pairs, answer this question: What's the difference in each case?

Figure 22

Adaptation of Proposals for CL-Oriented Pedagogies: Learning Collocations with "Ambition", "Experience", and "Job" in Unit 8 "Dream of the Stars" (Page 2)

Exercise 3. Complete the story by writing a verb from Exercise 2 in the correct form in the gaps. For some gaps, more than one answer may be possible.

I've always enjoyed performing in front of people and I'd like to **(1)** _____ a career as an actor. If I could **(2)** _____ my first ambition of going to drama school, I'd **(3)** _____ the knowledge and experience which is needed if I'm going to **(4)** _____ a job in the theatre. Acting is a very competitive profession, and you have to **(5)** _____ your career step by step until, hopefully, a well-known director recognizes your talent and **(6)** _____ you a job which really **(7)** _____ your career on the stage.

Exercise 4. Read the following phrases and explain their analogous meaning with a 'building'.

- to enter a profession: _____
- to make your way in the world: _____
- to be a step ahead of your workmates: _____

Source: Exercises 2 and 3 from Unit 8, "Dream of the stars" (Brook-Hart, 2014, p. 86), were adapted to the CL-inspired activity.

Figure 23

Example Material 1 Used in Teaching Treatment – Phase 1: Activity 1 "The Journey of Learning English"

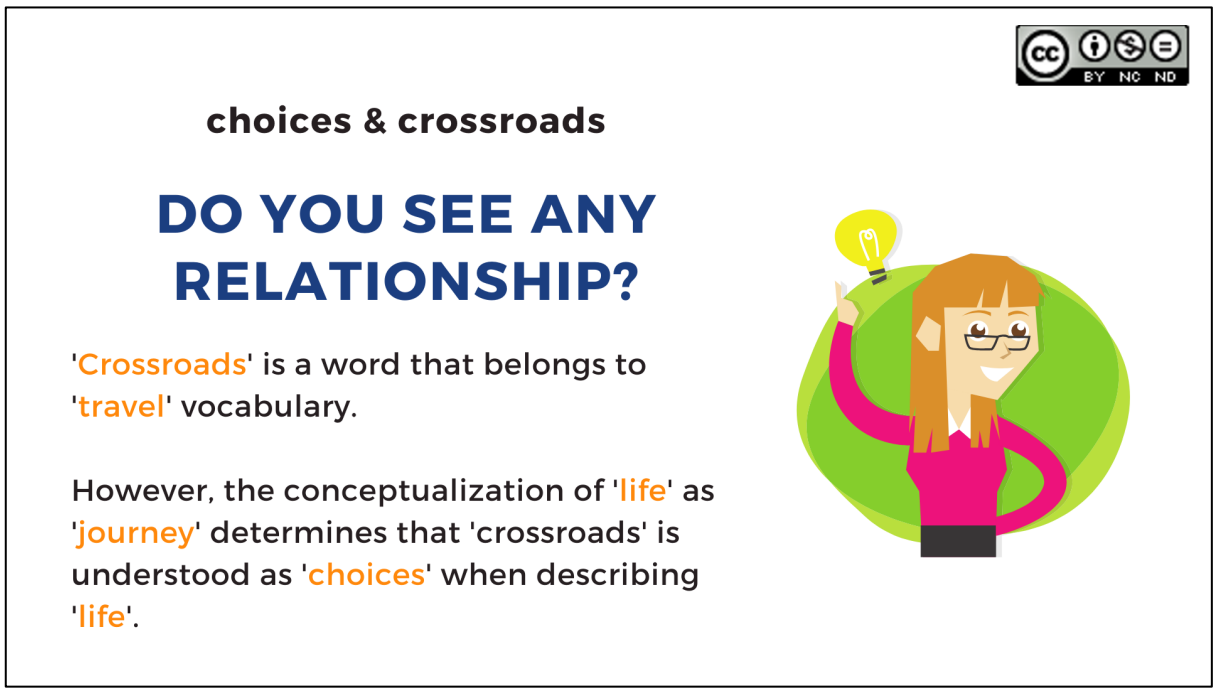


Figure 24

Example Material 2 Used in Teaching Treatment – Phase 1: Activity 1 "The Journey of Learning English"

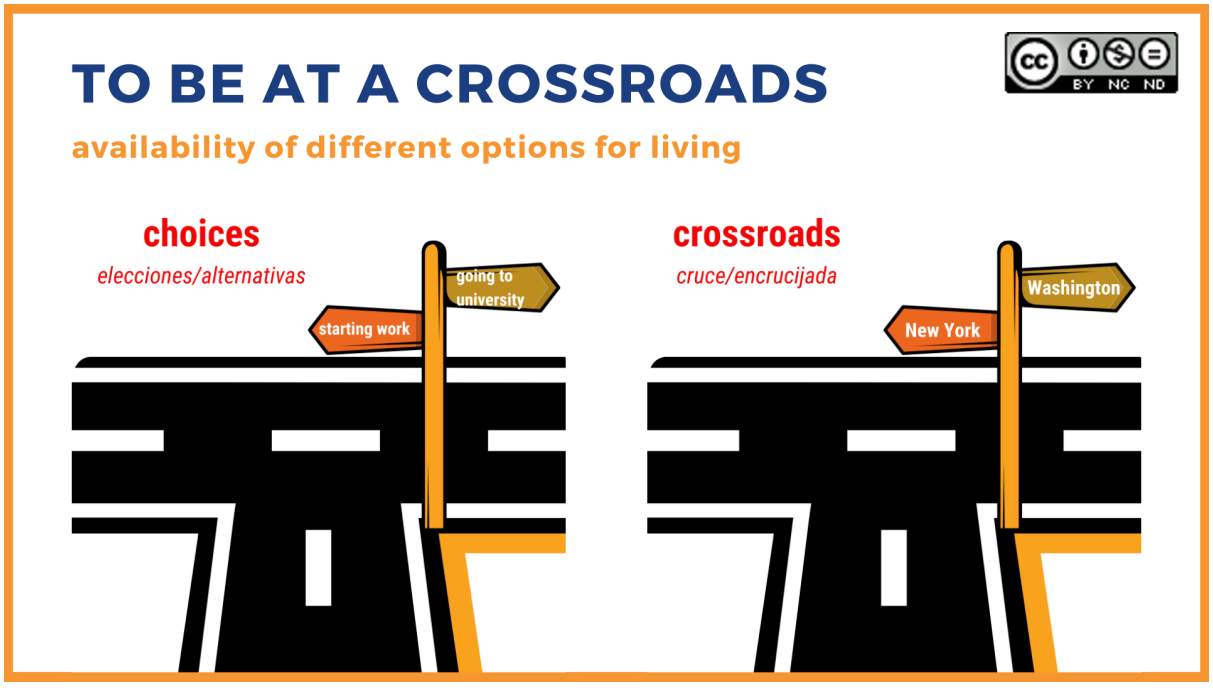


Figure 25

Example Material Used in Teaching Treatment – Phase 1: Activity 2 "Life is a Journey"

LIFE IS A JOURNEY

TASK 1: Match the vocabulary of 'travelling' with the definitions of their analogous meaning in the topic of 'life'. Do you agree with your partner?

1. to go nowhere
2. to arrive
3. to go through
4. to be at a crossroads
5. to move on
6. to pass away
7. to be over the hill
8. to go downhill
9. to be lost
10. to cross the bridge when it comes to it
11. to miss the boat
12. to start off on the wrong foot
13. to take an unexpected direction
14. to spin your wheels
15. to reach a dead-end

A. to be no longer young and therefore unable to do things that you could do in the past.
B. not to make a successful start.
C. to get worse.
D. to be born.
E. to not worry about a possible future problem but will deal with it if it happens.
F. not feeling confident or relaxed because you are in a new situation, for example with a group of people who you do not know.
G. to be in a point during the development of something when you have to make an important decision about what to do next.
H. to change surprisingly when you do not expect it at all, or you expect it to happen in a different way.
I. to waste time doing things that achieve nothing.
J. to fail to achieve something, or to fail to be successful.
K. to fail to take advantage of an opportunity by not taking action quickly enough.
L. to be in a situation in which no further progress is possible.
M. to die.
N. to start to continue with your life after you have dealt successfully with a bad experience.
O. to perform a set of actions that you regularly perform.

TASK 2: In pairs, group the vocabulary of 'travelling' according to their meaning in the topic of 'lif'e: problems, decisions or progress. Is there any rule or pattern?

PROBLEMS **DECISIONS** **PROGRESS**

TASK 3: In groups, stop by the different stations in the journey of life (past, present and future) and answer the given questions.

Figure 26

Example Material Used in Teaching Treatment – Phase 1: Activity 3 "The Journey of My Life" (Part 1)

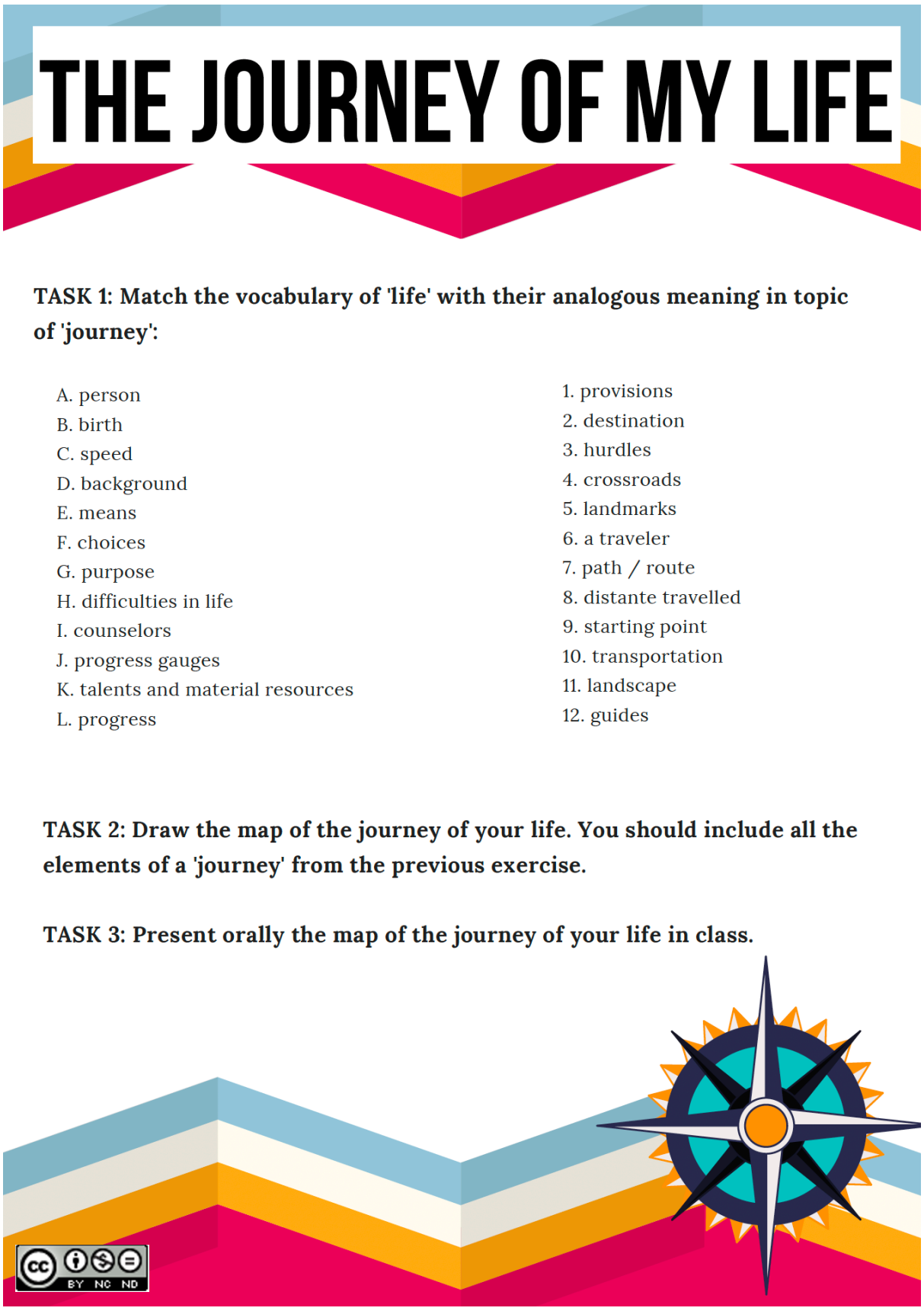
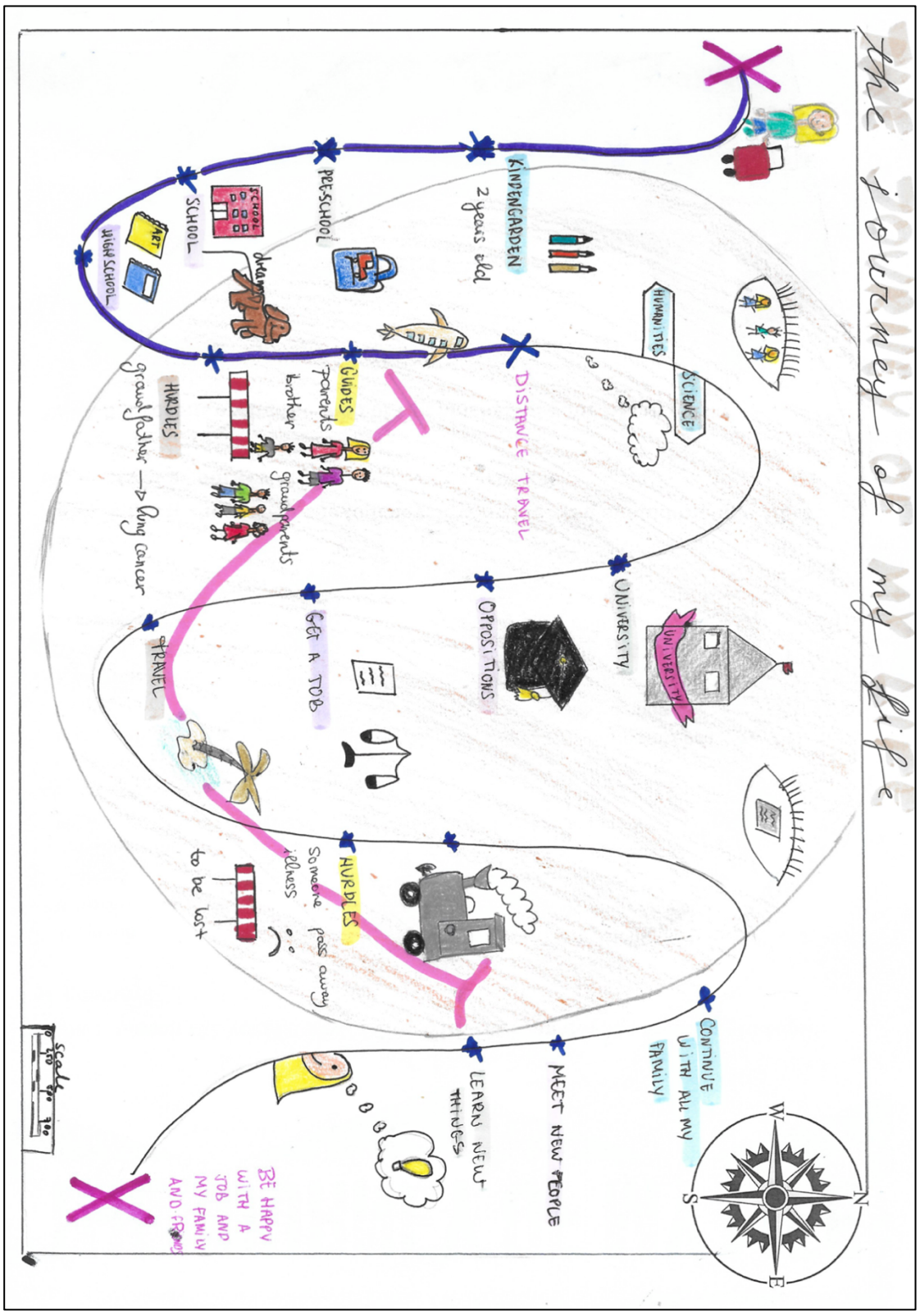


Figure 27

Example Material Used in Teaching Treatment – Phase 1: Activity 3 "The Journey of My Life" (Part 2)



Source: Drawing created by experimental group participant ST0206.

Figure 28

Example Material Used in Teaching Treatment – Phase 1: Activity 4 "Is it Worth Learning a Foreign Language?"

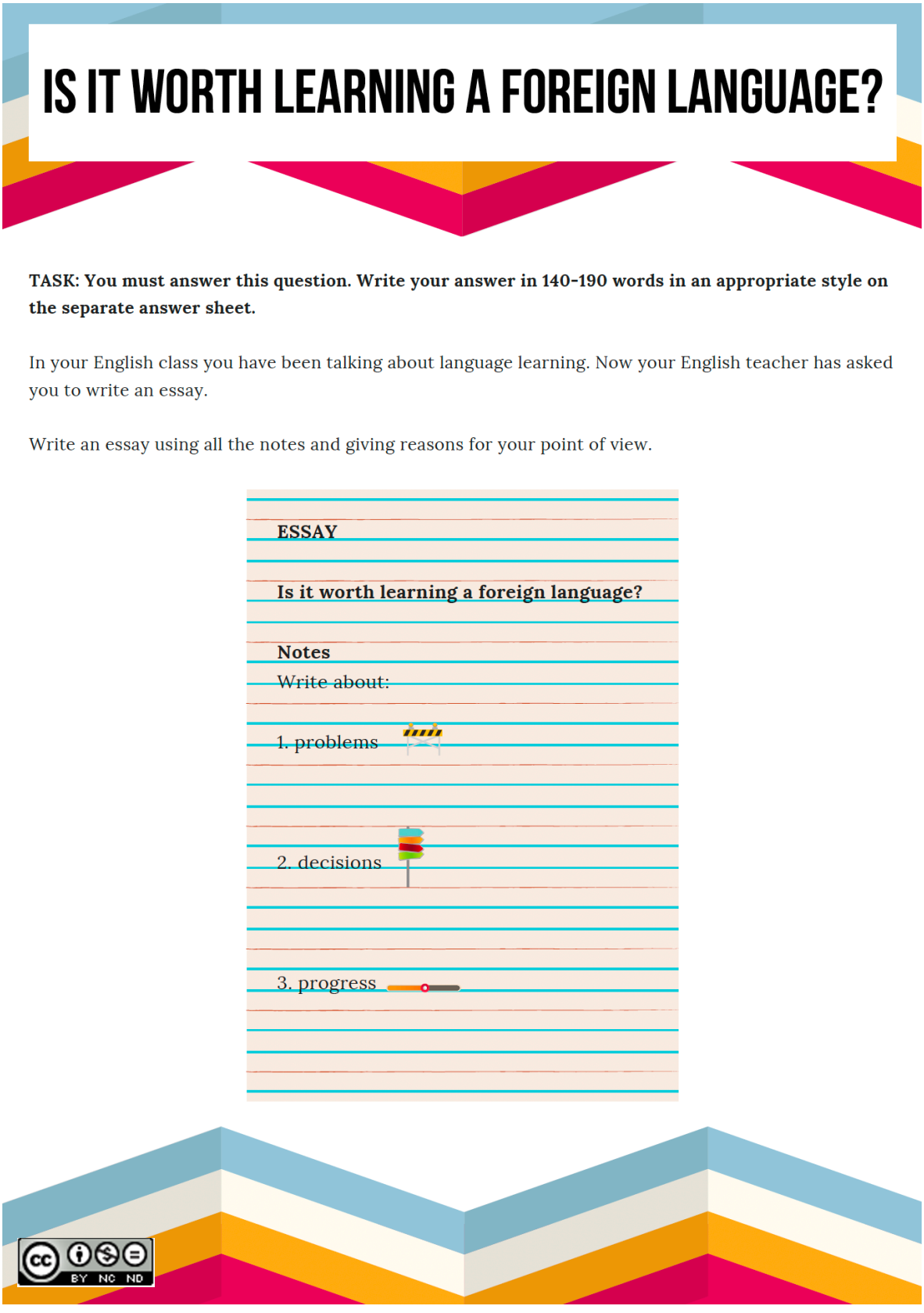


Figure 29

A Pictorial Example of the CM TIME IS MONEY *(Unit 10 "Spend, Spend, Spend?")*

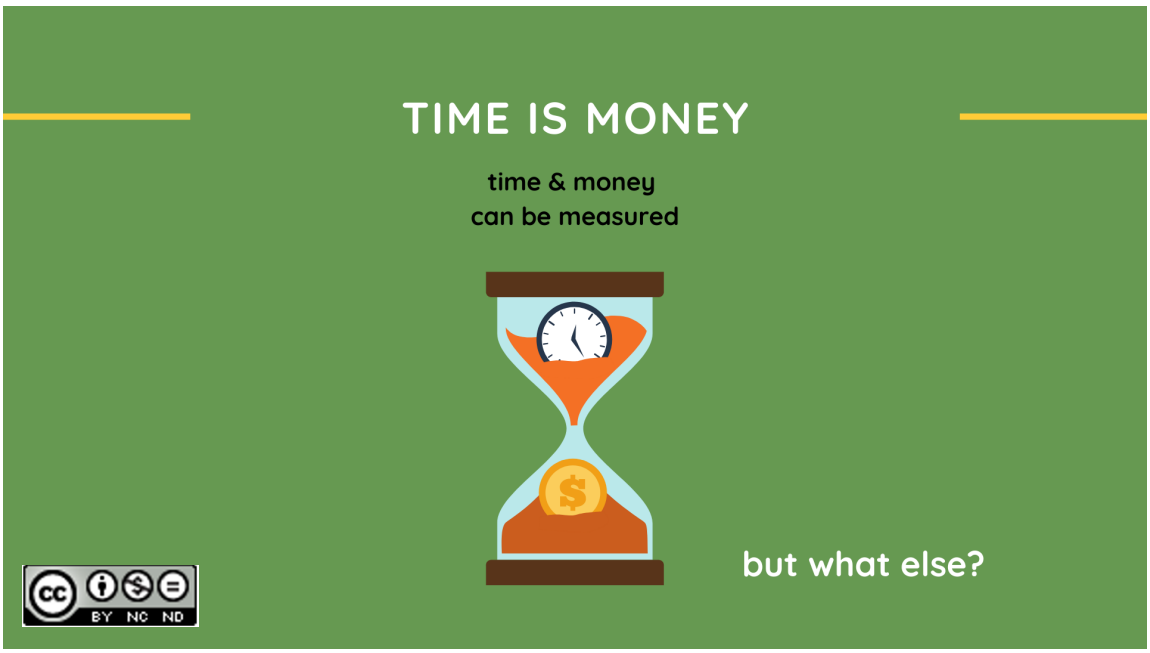


Figure 30

Example Material Used to Practice the Idiom "To Feel under the Weather" (Unit 11 "Medical Matters")

Source: Idioms from Unit 11, "Medical Matters" (Brook-Hart, 2014, p. 121), were adapted to the CL-inspired activity.

Figure 31

Example of Follow-up Written Task (Essay) in Unit 9 "Secrets of the Mind"

Source: The written task from Unit 9, "Secrets of the Mind" (Brook-Hart, 2014, pp. 96–105), was adapted to the follow-up essay.

Figure 32

Photograph of the Journey Analogy to Enhance Metaphor Awareness in the L2 Classroom

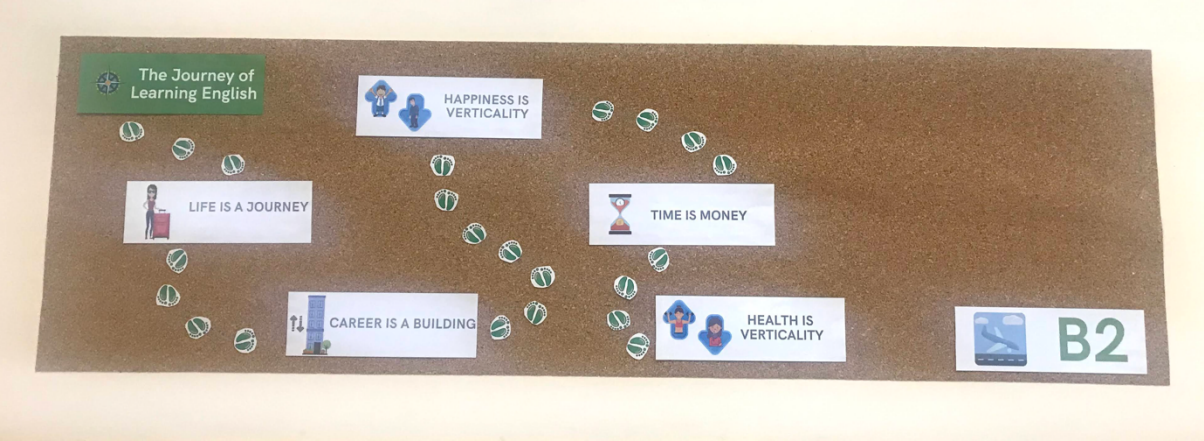
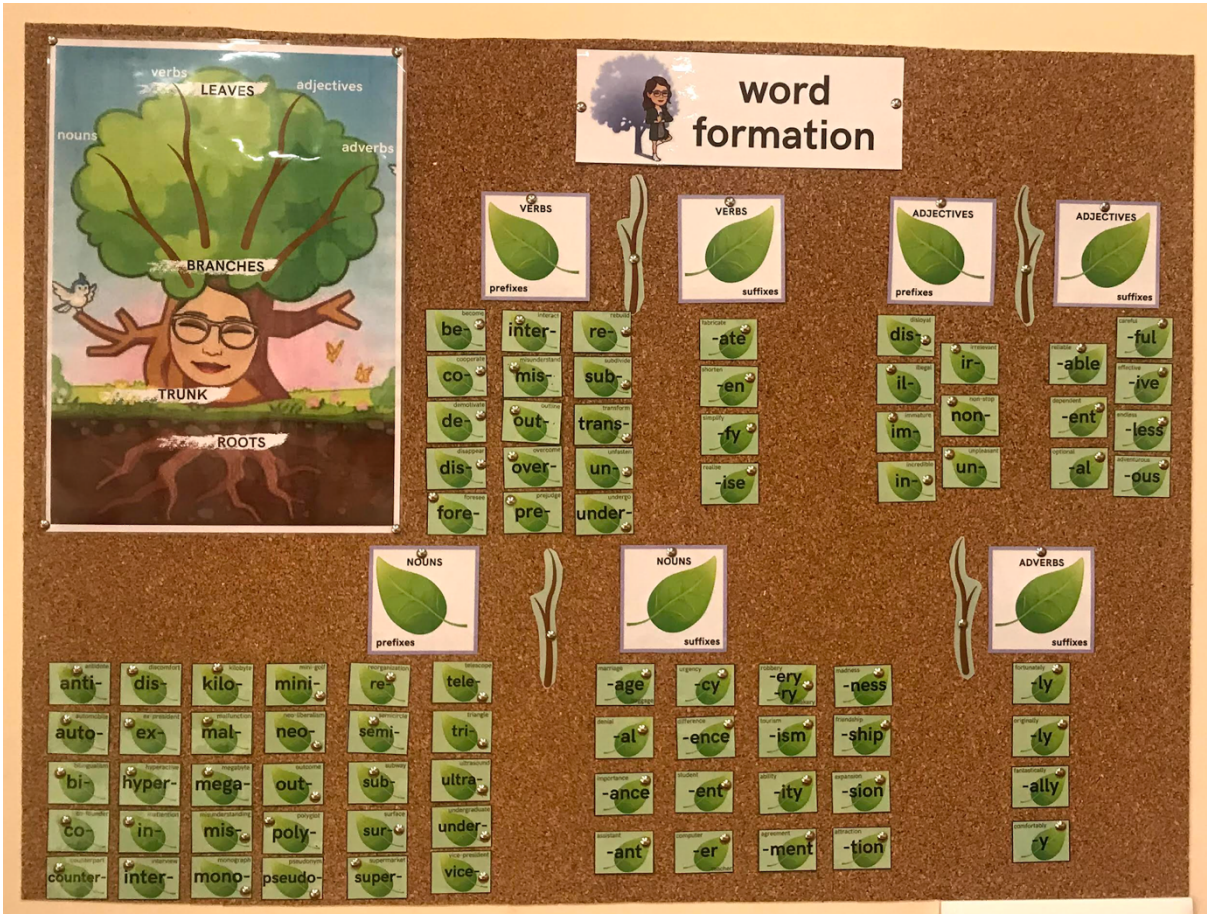


Figure 33

Photograph of the Tree Analogy Used to Support Learners with Word Formation Practice



Appendix M: Metaphor Identification Procedure

Figure 34

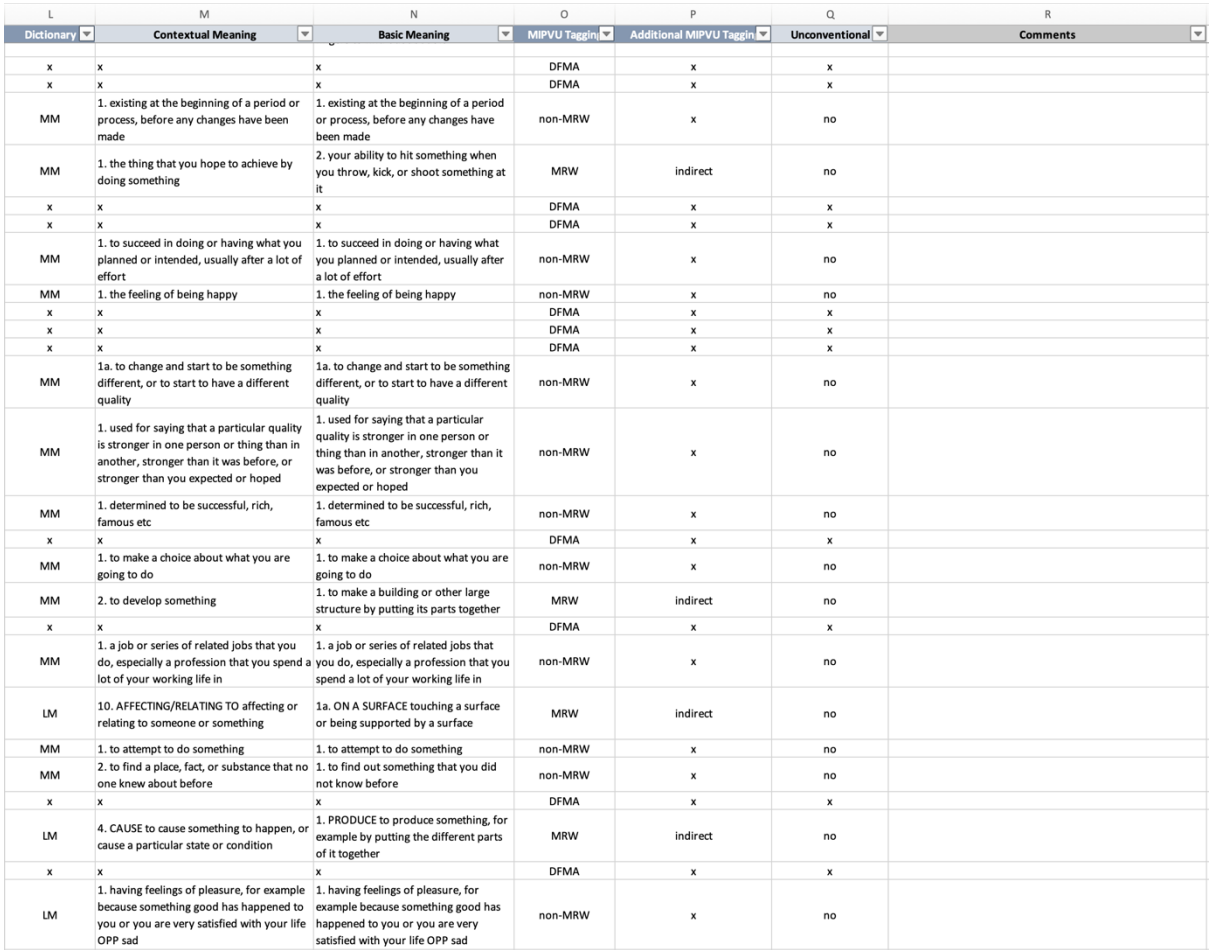
Screenshot of the Excel Spreadsheet Used as the Data Annotation Tool: Part 1

A	B	C	D	E	F	G	H	I	J	K
ID	Text ID	Lexical Unit	Lemma	LU Codes	Part of Speech	Error	Correct LU	Ortographic Error	Irregular Construction	Metaphor Analysis
2146	7	Although	although	word	conjunction	no	x	x	x	x
2147	7	my	my	word	determiner	no	x	x	x	x
2148	7	original	original	word	adjective	no	x	x	x	yes
2149	7	aim	aim	word	noun	no	x	x	x	yes
2150	7	had	have	word	verb	no	x	x	x	x
2151	7	been	be	word	verb	no	x	x	x	x
2152	7	to_achieve	achieve	word	verb	no	x	x	x	yes
2153	7	happiness	happiness	word	noun	no	x	x	x	yes
2154	7	for	for	word	preposition	no	x	x	x	x
2155	7	myself	myself	word	pronoun	no	x	x	x	x
2156	7	I	I	word	pronoun	no	x	x	x	x
2157	7	became	become	word	verb	no	x	x	x	yes
2158	7	more	more	word	adverb	no	x	x	x	yes
2159	7	ambitious	ambitious	word	adjective	no	x	x	x	yes
2160	7	I	I	word	pronoun	no	x	x	x	x
2161	7	decided	decide	word	verb	no	x	x	x	yes
2162	7	to_build	build	word	verb	no	x	x	x	yes
2163	7	my	my	word	determiner	no	x	x	x	x
2164	7	career	career	word	noun	no	x	x	x	yes
2165	7	on	on	word	preposition	no	x	x	x	yes
2166	7	trying	try	word	verb	no	x	x	x	yes
2167	7	to_discover	discover	word	verb	no	x	x	x	yes
2168	7	what	what	word	pronoun	no	x	x	x	x
2169	7	made	make	word	verb	no	x	x	x	yes
2170	7	others	other	word	pronoun	no	x	x	x	x
2171	7	happy	happy	word	adjective	no	x	x	x	yes

Source: This *Excel* spreadsheet screenshot shows a sentence extracted from a written input text (TB9.WINPUT.R5) of the *Reading* comprehension task titled *The secrets of happiness* in the textbook unit “Secrets of the Mind” (Brook-Hart, 2014, pp. 96–97).

Figure 35

Screenshot of the Excel Spreadsheet Used as the Data Annotation Tool: Part 2



L	M	N	O	P	Q	R
Dictionary	Contextual Meaning	Basic Meaning	MIPVU Taggin	Additional MIPVU Taggin	Unconventional	Comments
x	x	x	DFMA	x	x	
x	x	x	DFMA	x	x	
MM	1. existing at the beginning of a period or process, before any changes have been made	1. existing at the beginning of a period or process, before any changes have been made	non-MRW	x	no	
MM	1. the thing that you hope to achieve by doing something	2. your ability to hit something when you throw, kick, or shoot something at it	MRW	indirect	no	
x	x	x	DFMA	x	x	
x	x	x	DFMA	x	x	
MM	1. to succeed in doing or having what you planned or intended, usually after a lot of effort	1. to succeed in doing or having what you planned or intended, usually after a lot of effort	non-MRW	x	no	
MM	1. the feeling of being happy	1. the feeling of being happy	non-MRW	x	no	
x	x	x	DFMA	x	x	
x	x	x	DFMA	x	x	
x	x	x	DFMA	x	x	
MM	1a. to change and start to be something different, or to start to have a different quality	1a. to change and start to be something different, or to start to have a different quality	non-MRW	x	no	
MM	1. used for saying that a particular quality is stronger in one person or thing than in another, stronger than it was before, or stronger than you expected or hoped	1. used for saying that a particular quality is stronger in one person or thing than in another, stronger than it was before, or stronger than you expected or hoped	non-MRW	x	no	
MM	1. determined to be successful, rich, famous etc	1. determined to be successful, rich, famous etc	non-MRW	x	no	
x	x	x	DFMA	x	x	
MM	1. to make a choice about what you are going to do	1. to make a choice about what you are going to do	non-MRW	x	no	
MM	2. to develop something	1. to make a building or other large structure by putting its parts together	MRW	indirect	no	
x	x	x	DFMA	x	x	
MM	1. a job or series of related jobs that you do, especially a profession that you spend a lot of your working life in	1. a job or series of related jobs that you do, especially a profession that you spend a lot of your working life in	non-MRW	x	no	
LM	10. AFFECTING/RELATING TO affecting or relating to someone or something	1a. ON A SURFACE touching a surface or being supported by a surface	MRW	indirect	no	
MM	1. to attempt to do something	1. to attempt to do something	non-MRW	x	no	
MM	2. to find a place, fact, or substance that no one knew about before	1. to find out something that you did not know before	non-MRW	x	no	
x	x	x	DFMA	x	x	
LM	4. CAUSE to cause something to happen, or cause a particular state or condition	1. PRODUCE to produce something, for example by putting the different parts of it together	MRW	indirect	no	
x	x	x	DFMA	x	x	
LM	1. having feelings of pleasure, for example because something good has happened to you or you are very satisfied with your life OPP sad	1. having feelings of pleasure, for example because something good has happened to you or you are very satisfied with your life OPP sad	non-MRW	x	no	

Each column's purpose in the *Excel* spreadsheet employed for analysing metaphors is explained as follows:

- Column A was labelled "ID" and employed to assign a unique identification number to each LU for control purposes.
- Column B was labelled "Text ID" and used to store the assigned identification number for the text being analysed for metaphor identification.
- Column C was labelled "Lexical Unit" and contained the LUs transferred from the text being analysed for metaphor identification.
- Column D was labelled "Lemmas" and contained the lemmas related to the LUs identified in Column C.

- Column E was labelled “LU Codes” and used to record the classification of the different types of LUs in terms of their lexical demarcation. While most words were assigned the basic label of “word”, other codes were applied in special cases for LU determination, such as “proper noun”, “compound”, “phrasal verb”, “polyword” or “x” (used for LUs that could not be tagged within the previous selection range, e.g., truncated words in discourse).
- Column F was labelled “Part of Speech” and used to record the word class information for each LU in the analysis. The PoS tags used included “noun”, “verb”, “adjective, “adverb”, “preposition”, “conjunction”, “pronoun”, “determiner”, “interjection,” and “x” (used for LUs that could not be tagged within the previous selection range, e.g., abbreviations).
- Column G was labelled “Error” and was used to indicate whether the intended LU was used correctly in context (“yes”) or incorrectly (“no”).
- Column H was labelled “Correct LU” and used to record possible correct options for the LU in case it was incorrectly produced in context.
- Column I was labelled “Orthographic Error” and used to indicate the type of orthographic error resulting from grammar or spelling errors made in the production of the LU. The code “x” was entered if no error was found in the intended LU.
- Column J was labelled “Irregular Construction” (column J) and used to record information about the type of non-native-like phraseology used in the intended LU.
- Column K was labelled “Metaphor Analysis” and used to indicate whether the LU was included in the metaphor identification (“yes”) or excluded (“no”).
- Column L was labelled “Dictionary” and used to record the dictionary consulted to support decisions about metaphoricity in the analysis: *Macmillan Dictionary* online (MM), *Longman Dictionary of Contemporary English* online (LM), *Cambridge Dictionary* online (CAMD), and *Oxford English* online (OED).

- Column M was labelled “Contextual Meaning” and used to record the contextual sense of the LU in the text.
- Column N was labelled “Basic Meaning” and used to record the basic sense of the LU in the text.
- Column O was labelled “MIPVU Tagging” and used to indicate the possible tags for decisions about metaphor status in the analysis: “non-MRW” (non-Metaphor-Related Word), “MRW” (Metaphor-Related Word), “WiDLii” (When in Doubt, Leave it in), “MFlag” (Metaphor Flag), and “DFMA” (Discarded for Metaphor Analysis).
- Column P was labelled “Additional MIPVU Tagging” and used to further identify metaphorically used LUs as “indirect” or “direct”. The code “x” was selected when a LU was identified as “non-MRW”, “WiDLii”, “MFlag” or “DFMA”.
- Column Q was labelled “Unconventional” and was used to indicate whether the intended LU was conventionally used (“yes”) or not (“no”).
- Column R was labelled “Comments” and used to record comments on the metaphor analysis, or to note other figurative uses such as metonymy or possible personification.

Figure 36

Screenshot of the Excel Spreadsheet Used to Collect Basic Meanings

A	B	C	D	E
Term	Dictionary	PoS	Entry	Basic meaning
aback	LM	adverb	1	1. to be very surprised or shocked by something
abandon	MM	verb	1	3. to leave something in a place, especially because you are in a hurry or are trying to escape
abandoned	MM	adjective	1	1. left empty or no longer used
ability	MM	noun	1	1. the fact of being able to do something
able	LM	adjective	1	2. clever or good at doing something
about	LM	preposition	1	3. in many different directions within a particular place, or in different parts of a place SYN around, round
about	LM	adverb	2	2. British English in many different directions within a place or in different parts of a place SYN around
above	LM	preposition	1	1. in a higher position than something else SYN over OPP below
above	LM	adverb	1	1. in a higher position than something else SYN over OPP below
above	LM	adjective	2	1. used in a piece of writing to refer to something mentioned in an earlier part of the same piece of writing
abroad	MM	adverb	1	1. in or to a foreign country
absolutely	MM	adverb	1	1. completely: used for emphasis
abuse	MM	verb	1	2. to use something in a bad, dishonest, or harmful way
academic	MM	adjective	1	1. relating to education, especially education in colleges and universities
academy	MM	noun	1	1. a school or college that teaches a particular subject or skill
accept	MM	verb	1	1. to take something that someone gives you
access	MM	noun	1	2. the means by which you get to a place
accident	LM	noun	1	3. a situation in which someone is injured or something is damaged without anyone intending them to be
accomplish	MM	verb	1	1. to succeed in doing something, especially something that you have been trying to do for a period of time
account	LM	noun	1	1. a written or spoken description that says what happens in an event or process
achieve	MM	verb	1	1. to succeed in doing or having what you planned or intended, usually after a lot of effort
achievement	MM	noun	1	1. a particular thing that you have achieved
acknowledge	MM	verb	1	1. to accept or admit that something exists, is true, or is real
across	LM	preposition	1	1. from one side of something to the other
acting	MM	noun	1	1. the job or skill of performing in plays and films
action	MM	noun	1	1. the process of doing something, especially in order to stop a bad situation from developing or continuing // 2. something you do, especially something that seems wrong or unusual to other people
activity	LM	noun	1	1. things that people do, especially in order to achieve a particular aim
actor	MM	noun	1	1. someone who performs in plays and films, especially as their job
actual	LM	adjective	1	1. used to emphasize that something is real or exact
actuality	LM	noun	1	2. the state of being real or really existing
actually	MM	adverb	1	1. used for emphasizing what is really true or what really happened
adapt	LM	verb	1	2. to change something to make it suitable for a different purpose
add	LM	verb	1	1. PUT WITH SOMETHING ELSE to put something with something else or with a group of other things
addicted	MM	adjective	1	1a. enjoying a particular activity very much and spending as much time as you can doing it

Appendix N: Results of Research Question 1a (RQ1a)

Table 79

Metaphor Use in Learner Oral Discourse (Pre-Test): Experimental Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	142	50	13	9.15
	ST0102	280	88	27	9.64
	ST0103	286	76	31	10.84
	ST0104	193	55	22	11.40
	ST0105	412	100	37	8.98
	ST0106	155	48	18	11.61
	ST0107	286	83	30	10.49
	ST0108	292	105	28	9.59
	ST0109	240	63	26	10.83
	ST0110	166	61	17	10.24
GR0102	ST0201	255	78	32	12.55
	ST0202	159	46	20	12.58
	ST0203	304	104	28	9.21
	ST0204	197	56	14	7.11
	ST0205	177	49	33	18.64
	ST0206	229	76	29	12.66
	ST0207	167	53	24	14.37
	ST0208	351	116	33	9.40
	ST0209	179	64	18	10.06
	ST0210	254	85	21	8.27
Total		4,724	1,456	501	10.88

Table 80

Metaphor Use in Learner Oral Discourse (Pre-Test): Control Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	161	54	16	9.94
	ST0302	145	37	29	20.00
	ST0303	209	57	32	15.31
	ST0304	191	68	19	9.95
	ST0305	165	46	21	12.73
	ST0306	212	56	27	12.74
	ST0307	164	52	18	10.98
	ST0308	219	59	34	15.53
	ST0309	200	63	27	13.50
	ST0310	236	64	25	10.59
GR0204	ST0401	178	48	19	10.67
	ST0402	160	49	14	8.75
	ST0403	224	64	18	8.04
	ST0404	222	62	19	8.56
	ST0405	187	59	27	14.44
	ST0406	242	77	31	12.81
	ST0407	226	59	20	8.85
	ST0408	168	44	20	11.90
	ST0409	263	81	20	7.60
	ST0410	193	58	17	8.81
Total		3,965	1,157	453	11.58

Table 81

Metaphor Use in Learner Written Discourse (Pre-Test): Experimental Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	201	66	33	16.42
	ST0102	181	83	28	15.47
	ST0103	192	67	39	20.31
	ST0104	172	72	24	13.95
	ST0105	173	74	24	13.87
	ST0106	185	72	37	20.00
	ST0107	222	81	42	18.92
	ST0108	153	66	24	15.69
	ST0109	182	93	33	18.13
	ST0110	179	72	32	17.88
GR0102	ST0201	142	42	24	16.90
	ST0202	190	68	27	14.21
	ST0203	185	86	24	12.97
	ST0204	172	72	30	17.44
	ST0205	189	74	27	14.29
	ST0206	174	65	30	17.24
	ST0207	216	93	35	16.20
	ST0208	187	76	27	14.44
	ST0209	178	57	37	20.79
	ST0210	154	70	27	17.53
Total		3,627	1,449	604	16.63

Table 82

Metaphor Use in Learner Written Discourse (Pre-Test): Control Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	147	62	24	16.33
	ST0302	175	67	19	10.86
	ST0303	154	68	24	15.58
	ST0304	127	54	20	15.75
	ST0305	163	63	28	17.18
	ST0306	174	64	30	17.24
	ST0307	185	78	28	15.14
	ST0308	169	75	21	12.43
	ST0309	132	54	18	13.64
	ST0310	185	77	29	15.68
GR0204	ST0401	226	79	44	19.47
	ST0402	136	51	25	18.38
	ST0403	179	63	25	13.97
	ST0404	198	81	40	20.20
	ST0405	164	64	32	19.51
	ST0406	168	63	20	11.90
	ST0407	234	96	41	17.52
	ST0408	193	65	33	17.10
	ST0409	183	75	25	13.66
	ST0410	171	76	21	12.28
Total		3,463	1,375	547	15.69

Table 83

Metaphor Use in Learner Oral Discourse (Post-Test): Experimental Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	216	75	22	10.19
	ST0102	346	86	53	15.32
	ST0103	285	82	25	8.77
	ST0104	188	72	17	9.04
	ST0105	317	100	61	19.24
	ST0106	201	72	25	12.44
	ST0107	229	84	19	8.30
	ST0108	259	85	24	9.27
	ST0109	252	95	24	9.52
	ST0110	162	60	15	9.26
GR0102	ST0201	259	80	21	8.11
	ST0202	122	37	19	15.57
	ST0203	298	95	39	13.09
	ST0204	280	77	27	9.64
	ST0205	213	63	22	10.33
	ST0206	194	76	19	9.79
	ST0207	131	45	16	12.21
	ST0208	216	67	24	11.11
	ST0209	154	57	21	13.64
	ST0210	139	51	9	6.47
Total		4,461	1,459	502	11.07

Table 84

Metaphor Use in Learner Oral Discourse (Post-Test): Control Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	153	45	21	13.73
	ST0302	107	32	15	14.02
	ST0303	157	55	10	6.37
	ST0304	136	36	8	5.88
	ST0305	125	47	9	7.20
	ST0306	178	57	24	13.48
	ST0307	177	58	20	11.30
	ST0308	149	53	13	8.72
	ST0309	117	33	16	13.68
	ST0310	153	50	8	5.23
GR0204	ST0401	215	57	24	11.16
	ST0402	141	47	10	7.09
	ST0403	232	75	25	10.78
	ST0404	156	35	9	5.77
	ST0405	180	57	23	12.78
	ST0406	250	83	22	8.80
	ST0407	179	51	19	10.61
	ST0408	135	46	14	10.37
	ST0409	135	46	10	7.41
	ST0410	122	39	10	8.20
Total		3,197	1,002	310	9.63

Table 85

Metaphor Use in Learner Written Discourse (Post-Test): Experimental Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	204	64	31	15.20
	ST0102	180	52	35	19.44
	ST0103	183	50	39	21.31
	ST0104	203	65	38	18.72
	ST0105	253	81	54	21.34
	ST0106	170	53	36	21.18
	ST0107	170	49	36	21.18
	ST0108	184	68	38	20.65
	ST0109	165	48	40	24.24
	ST0110	191	66	32	16.75
GR0102	ST0201	180	55	31	17.22
	ST0202	173	46	32	18.50
	ST0203	170	61	32	18.82
	ST0204	208	71	31	14.90
	ST0205	204	74	27	13.24
	ST0206	186	62	35	18.82
	ST0207	216	62	37	17.13
	ST0208	257	90	55	21.40
	ST0209	189	49	37	19.58
	ST0210	195	63	52	26.67
Total		3,881	1,229	748	19.31

Table 86

Metaphor Use in Learner Written Discourse (Post-Test): Control Group

B2 groups	Participants	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	160	49	30	18.75
	ST0302	241	79	38	15.77
	ST0303	152	51	22	14.47
	ST0304	128	50	23	17.97
	ST0305	173	70	32	18.50
	ST0306	213	63	38	17.84
	ST0307	200	63	28	14.00
	ST0308	183	67	27	14.75
	ST0309	168	61	29	17.26
	ST0310	197	73	24	12.18
GR0204	ST0401	227	69	45	19.82
	ST0402	230	71	49	21.30
	ST0403	217	90	38	17.51
	ST0404	159	55	26	16.35
	ST0405	177	75	22	12.43
	ST0406	180	71	24	13.33
	ST0407	197	74	27	13.71
	ST0408	178	56	28	15.73
	ST0409	210	69	39	18.57
	ST0410	169	52	30	17.75
Total		3,759	1,308	619	16.40

Table 87

Metaphor Density Variations in Learner Oral Discourse (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				AI	RI	%
GR0101	ST0101	9.15	10.19	1.03	0.11	11
	ST0102	9.64	15.32	5.68	0.59	59
	ST0103	10.84	8.77	-2.07	-0.19	-19
	ST0104	11.40	9.04	-2.36	-0.21	-21
	ST0105	8.98	19.24	10.26	1.14	114
	ST0106	11.61	12.44	0.82	0.07	7
	ST0107	10.49	8.30	-2.19	-0.21	-21
	ST0108	9.59	9.27	-0.32	-0.03	-3
	ST0109	10.83	9.52	-1.31	-0.12	-12
	ST0110	10.24	9.26	-0.98	-0.10	-10
GR0102	ST0201	12.55	8.11	-4.44	-0.35	-35
	ST0202	12.58	15.57	3.00	0.24	24
	ST0203	9.21	13.09	3.88	0.42	42
	ST0204	7.11	9.64	2.54	0.36	36
	ST0205	18.64	10.33	-8.32	-0.45	-45
	ST0206	12.66	9.79	-2.87	-0.23	-23
	ST0207	14.37	12.21	-2.16	-0.15	-15
	ST0208	9.40	11.11	1.71	0.18	18
	ST0209	10.06	13.64	3.58	0.36	36
	ST0210	8.27	6.47	-1.79	-0.22	-22
Total		10.88	11.07	0.18	0.06	6

Table 88

Metaphor Density Variations in Learner Oral Discourse (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				AI	RI	%
GR0203	ST0301	9.94	13.73	3.79	0.38	38
	ST0302	20.00	14.02	-5.98	-0.30	-30
	ST0303	15.31	6.37	-8.94	-0.58	-58
	ST0304	9.95	5.88	-4.07	-0.41	-41
	ST0305	12.73	7.20	-5.53	-0.43	-43
	ST0306	12.74	13.48	0.75	0.06	6
	ST0307	10.98	11.30	0.32	0.03	3
	ST0308	15.53	8.72	-6.80	-0.44	-44
	ST0309	13.50	13.68	0.18	0.01	1
	ST0310	10.59	5.23	-5.36	-0.51	-51
GR0204	ST0401	10.67	11.16	0.49	0.05	5
	ST0402	8.75	7.09	-1.66	-0.19	-19
	ST0403	8.04	10.78	2.74	0.34	34
	ST0404	8.56	5.77	-2.79	-0.33	-33
	ST0405	14.44	12.78	-1.66	-0.12	-12
	ST0406	12.81	8.80	-4.01	-0.31	-31
	ST0407	8.85	10.61	1.76	0.20	20
	ST0408	11.90	10.37	-1.53	-0.13	-13
	ST0409	7.60	7.41	-0.20	-0.03	-3
	ST0410	8.81	8.20	-0.61	-0.07	-7
Total		11.58	9.63	-1.96	-0.14	-14

Table 89

Metaphor Density Variations in Learner Written Discourse (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				AI	RI	%
GR0101	ST0101	16.42	15.20	-1.22	-0.07	-7
	ST0102	15.47	19.44	3.97	0.26	26
	ST0103	20.31	21.31	1.00	0.05	5
	ST0104	13.95	18.72	4.77	0.34	34
	ST0105	13.87	21.34	7.47	0.54	54
	ST0106	20.00	21.18	1.18	0.06	6
	ST0107	18.92	21.18	2.26	0.12	12
	ST0108	15.69	20.65	4.97	0.32	32
	ST0109	18.13	24.24	6.11	0.34	34
	ST0110	17.88	16.75	-1.12	-0.06	-6
GR0102	ST0201	16.90	17.22	0.32	0.02	2
	ST0202	14.21	18.50	4.29	0.30	30
	ST0203	12.97	18.82	5.85	0.45	45
	ST0204	17.44	14.90	-2.54	-0.15	-15
	ST0205	14.29	13.24	-1.05	-0.07	-7
	ST0206	17.24	18.82	1.58	0.09	9
	ST0207	16.20	17.13	0.93	0.06	6
	ST0208	14.44	21.40	6.96	0.48	48
	ST0209	20.79	19.58	-1.21	-0.06	6
	ST0210	17.53	26.67	9.13	0.52	52
Total		16.63	19.31	2.68	0.18	18

Table 90

Metaphor Density Variations in Learner Written Discourse (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
				AI	RI	%
GR0203	ST0301	16.33	18.75	2.42	0.15	15
	ST0302	10.86	15.77	4.91	0.45	45
	ST0303	15.58	14.47	-1.11	-0.07	-7
	ST0304	15.75	17.97	2.22	0.14	14
	ST0305	17.18	18.50	1.32	0.08	8
	ST0306	17.24	17.84	0.60	0.03	3
	ST0307	15.14	14.00	-1.14	-0.08	-8
	ST0308	12.43	14.75	2.33	0.19	19
	ST0309	13.64	17.26	3.63	0.27	27
	ST0310	15.68	12.18	-3.49	-0.22	-22
GR0204	ST0401	19.47	19.82	0.35	0.02	2
	ST0402	18.38	21.30	2.92	0.16	16
	ST0403	13.97	17.51	3.55	0.25	25
	ST0404	20.20	16.35	-3.85	-0.19	-19
	ST0405	19.51	12.43	-7.08	-0.36	-36
	ST0406	11.90	13.33	1.43	0.12	12
	ST0407	17.52	13.71	-3.82	-0.22	-22
	ST0408	17.10	15.73	-1.37	-0.08	-8
	ST0409	13.66	18.57	4.91	0.36	36
	ST0410	12.28	17.75	5.47	0.45	45
Total		15.69	16.40	0.71	0.07	7

Appendix O: Results of Research Question 1b (RQ1b)

Table 91

Use of Open- vs. Closed-Class Metaphors in the Oral Pre-Test: Experimental Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	open	79	50	10	12.66
		closed	9	0	3	33.33
	ST0102	open	124	77	20	16.13
		closed	24	11	7	29.17
	ST0103	open	134	74	25	18.66
		closed	20	2	6	30.00
	ST0104	open	99	55	16	16.16
		closed	9	0	6	66.67
	ST0105	open	184	96	30	16.30
		closed	19	4	7	36.84
	ST0106	open	74	46	14	18.92
		closed	8	2	4	50.00
	ST0107	open	137	79	24	17.52
		closed	18	4	6	33.33
	ST0108	open	157	101	19	12.10
		closed	14	4	9	64.29
	ST0109	open	117	62	19	16.24
		closed	13	1	7	53.85
	ST0110	open	88	57	11	12.50
		closed	12	4	6	50.00
GR0102	ST0201	open	112	72	18	16.07
		closed	22	6	14	63.64
	ST0202	open	83	46	14	16.87
		closed	6	0	6	100.00
	ST0203	open	138	99	19	13.77
		closed	17	5	9	52.94
	ST0204	open	85	53	10	11.76
		closed	7	3	4	57.14
	ST0205	open	84	47	22	26.19
		closed	17	2	11	64.71
	ST0206	open	120	72	19	15.83
		closed	16	4	10	62.50
	ST0207	open	73	51	11	15.07
		closed	18	2	13	72.22
	ST0208	open	165	106	24	14.55
		closed	25	10	9	36.00
	ST0209	open	95	60	12	12.63
		closed	13	4	6	46.15
	ST0210	open	123	80	10	8.13
		closed	16	5	11	68.75
Total		open	2,271	1,383	347	15.40
		closed	303	73	154	53.58

Table 92

Use of Open- vs. Closed-Class Metaphors in the Oral Pre-Test: Control Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	open	84	50	11	13.10
		closed	10	4	5	50.00
	ST0302	open	64	32	23	35.94
		closed	15	5	6	40.00
	ST0303	open	86	54	16	18.60
		closed	22	3	16	72.73
	ST0304	open	95	60	14	14.74
		closed	16	8	5	31.25
	ST0305	open	78	43	14	17.95
		closed	10	3	7	70.00
	ST0306	open	101	52	19	18.81
		closed	16	4	8	50.00
	ST0307	open	74	48	13	17.57
		closed	13	4	5	38.46
	ST0308	open	98	56	24	24.49
		closed	15	3	10	66.67
	ST0309	open	97	60	23	23.71
		closed	8	3	4	50.00
	ST0310	open	104	57	20	19.23
		closed	18	7	5	27.78
GR0204	ST0401	open	81	48	14	17.28
		closed	6	0	5	83.33
	ST0402	open	84	48	9	10.71
		closed	8	1	5	62.50
	ST0403	open	101	63	10	9.90
		closed	14	1	8	57.14
	ST0404	open	98	61	12	12.24
		closed	13	1	7	53.85
	ST0405	open	96	53	22	22.92
		closed	17	6	5	29.41
	ST0406	open	123	75	19	15.45
		closed	17	2	12	70.59
	ST0407	open	94	56	17	18.09
		closed	12	3	3	25.00
	ST0408	open	89	44	16	17.98
		closed	5	0	4	80.00
	ST0409	open	121	80	12	9.92
		closed	18	1	8	44.44
	ST0410	open	86	54	11	12.79
		closed	16	4	6	37.50
Total		open	1,854	1,094	319	17.57
		closed	269	63	134	52.03

Table 93

Use of Open- vs. Closed-Class Metaphors in the Written Pre-Test: Experimental Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	open	105	62	23	21.90
		closed	23	4	10	43.48
	ST0102	open	111	79	20	18.02
		closed	17	4	8	47.06
	ST0103	open	119	63	28	23.53
		closed	18	4	11	61.11
	ST0104	open	110	71	22	20.00
		closed	6	1	2	33.33
	ST0105	open	110	72	17	15.45
		closed	11	2	7	63.64
	ST0106	open	112	67	26	23.21
		closed	19	5	11	57.89
	ST0107	open	138	81	34	24.64
		closed	15	0	8	53.33
	ST0108	open	98	64	18	18.37
		closed	12	2	6	50.00
	ST0109	open	130	88	25	19.23
		closed	19	5	8	42.11
	ST0110	open	101	70	20	19.80
		closed	17	2	12	70.59
GR0102	ST0201	open	69	40	13	18.84
		closed	22	2	11	50.00
	ST0202	open	105	62	15	14.29
		closed	22	6	12	54.55
	ST0203	open	121	85	18	14.88
		closed	11	1	6	54.55
	ST0204	open	102	70	20	19.61
		closed	16	2	10	62.50
	ST0205	open	108	72	17	15.74
		closed	19	2	10	52.63
	ST0206	open	99	63	20	20.20
		closed	15	2	10	66.67
	ST0207	open	118	89	22	18.64
		closed	19	4	13	68.42
	ST0208	open	113	73	19	16.81
		closed	14	3	8	57.14
	ST0209	open	88	55	23	26.14
		closed	17	2	14	82.35
	ST0210	open	94	63	21	22.34
		closed	15	7	6	40.00
Total		open	2,151	1,389	421	19.58
		closed	327	60	183	55.57

Table 94

Use of Open- vs. Closed-Class Metaphors in the Written Pre-Test: Control Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	open	86	56	16	18.60
		closed	20	6	8	40.00
	ST0302	open	100	66	10	10.00
		closed	12	1	9	75.00
	ST0303	open	93	66	16	17.20
		closed	13	2	8	61.54
	ST0304	open	77	52	15	19.48
		closed	12	2	5	41.67
	ST0305	open	98	60	18	18.37
		closed	15	3	10	66.67
	ST0306	open	97	60	19	19.59
		closed	20	4	11	55.00
	ST0307	open	105	77	14	13.33
		closed	20	1	14	70.00
	ST0308	open	99	73	13	13.13
		closed	15	2	8	53.33
	ST0309	open	77	54	10	12.99
		closed	11	0	8	72.73
	ST0310	open	114	76	18	15.79
		closed	17	1	11	64.71
GR0204	ST0401	open	134	74	29	21.64
		closed	26	5	15	57.69
	ST0402	open	80	49	20	25.00
		closed	9	2	5	55.56
	ST0403	open	106	62	19	17.92
		closed	16	1	6	37.50
	ST0404	open	131	78	31	23.66
		closed	12	3	9	75.00
	ST0405	open	100	63	25	25.00
		closed	14	1	7	50.00
	ST0406	open	101	63	16	15.84
		closed	8	0	4	50.00
	ST0407	open	133	93	25	18.80
		closed	25	3	16	64.00
	ST0408	open	115	63	27	23.48
		closed	14	2	6	42.86
	ST0409	open	98	70	14	14.29
		closed	23	5	11	47.83
	ST0410	open	102	73	13	12.75
		closed	20	3	8	40.00
Total		open	2,046	1,328	368	17.84
		closed	322	47	179	56.05

Table 95

Use of Open- vs. Closed-Class Metaphors in the Oral Post-Test: Experimental Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	open	106	71	18	16.98
		closed	14	4	4	28.57
	ST0102	open	147	77	42	28.57
		closed	32	9	11	34.38
	ST0103	open	135	72	18	13.33
		closed	23	10	7	30.43
	ST0104	open	98	66	13	13.27
		closed	10	6	4	40.00
	ST0105	open	173	87	51	29.48
		closed	29	13	10	34.48
	ST0106	open	101	62	20	19.80
		closed	19	10	5	26.32
	ST0107	open	112	75	12	10.71
		closed	20	9	7	35.00
	ST0108	open	142	81	20	14.08
		closed	17	4	4	23.53
	ST0109	open	115	77	16	13.91
		closed	29	18	8	27.59
	ST0110	open	75	51	12	16.00
		closed	14	9	3	21.43
GR0102	ST0201	open	112	71	15	13.39
		closed	19	9	6	31.58
	ST0202	open	56	31	17	30.36
		closed	12	6	2	16.67
	ST0203	open	150	85	31	20.67
		closed	20	10	8	40.00
	ST0204	open	117	68	24	20.51
		closed	15	9	3	20.00
	ST0205	open	99	55	18	18.18
		closed	15	8	4	26.67
	ST0206	open	94	69	14	14.89
		closed	14	7	5	35.71
	ST0207	open	60	37	12	20.00
		closed	14	8	4	28.57
	ST0208	open	105	60	21	20.00
		closed	16	7	3	18.75
	ST0209	open	76	51	13	17.11
		closed	17	6	8	47.06
	ST0210	open	71	45	6	8.45
		closed	9	6	3	33.33
Total		open	2,144	1,291	393	17.99
		closed	358	168	109	30.00

Table 96

Use of Open- vs. Closed-Class Metaphors in the Oral Post-Test: Control Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	open	71	38	17	23.94
		closed	13	7	4	30.77
	ST0302	open	50	30	10	20.00
		closed	10	2	5	50.00
	ST0303	open	71	50	7	9.86
		closed	10	5	3	30.00
	ST0304	open	52	34	4	7.69
		closed	10	2	4	40.00
	ST0305	open	64	42	7	10.94
		closed	10	5	2	20.00
	ST0306	open	92	51	17	18.48
		closed	16	6	7	43.75
	ST0307	open	90	49	18	20.00
		closed	17	9	2	11.76
	ST0308	open	74	44	11	14.86
		closed	15	9	2	13.33
	ST0309	open	65	32	13	20.00
		closed	8	1	3	37.50
	ST0310	open	67	46	2	2.99
		closed	10	4	6	60.00
GR0204	ST0401	open	83	50	15	18.07
		closed	18	7	9	50.00
	ST0402	open	66	41	9	13.64
		closed	8	6	1	12.50
	ST0403	open	122	71	21	17.21
		closed	11	4	4	36.36
	ST0404	open	52	32	7	13.46
		closed	7	3	2	28.57
	ST0405	open	87	51	22	25.29
		closed	8	6	1	12.50
	ST0406	open	123	78	18	14.63
		closed	17	5	4	23.53
	ST0407	open	78	43	15	19.23
		closed	14	8	4	28.57
	ST0408	open	65	41	11	16.92
		closed	11	5	3	27.27
	ST0409	open	71	44	9	12.68
		closed	5	2	1	20.00
	ST0410	open	56	36	5	8.93
		closed	11	3	5	45.45
Total		open	1,499	903	238	15.44
		closed	229	99	72	31.09

Table 97

Use of Open- vs. Closed-Class Metaphors in the Written Post-Test: Experimental Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	open	122	61	24	19.67
		closed	16	3	7	43.75
	ST0102	open	101	50	25	24.75
		closed	15	2	10	66.67
	ST0103	open	105	47	27	25.71
		closed	20	3	12	60.00
	ST0104	open	133	64	27	20.30
		closed	13	1	11	84.62
	ST0105	open	141	77	36	25.53
		closed	29	4	18	62.07
	ST0106	open	108	52	30	27.78
		closed	8	1	6	75.00
	ST0107	open	104	47	31	29.81
		closed	15	2	5	33.33
	ST0108	open	120	68	29	24.17
		closed	15	0	9	60.00
	ST0109	open	104	46	25	24.04
		closed	19	2	15	78.95
	ST0110	open	120	64	25	20.83
		closed	15	2	7	46.67
GR0102	ST0201	open	104	52	24	23.08
		closed	16	3	7	43.75
	ST0202	open	95	44	22	23.16
		closed	17	2	10	58.82
	ST0203	open	112	60	25	22.32
		closed	11	1	7	63.64
	ST0204	open	125	70	23	18.40
		closed	12	1	8	66.67
	ST0205	open	134	73	16	11.94
		closed	13	1	11	84.62
	ST0206	open	120	60	26	21.67
		closed	12	2	9	75.00
	ST0207	open	121	62	24	19.83
		closed	18	0	13	72.22
	ST0208	open	159	84	37	23.27
		closed	33	6	18	54.55
	ST0209	open	108	49	22	20.37
		closed	17	0	15	88.24
	ST0210	open	121	61	39	32.23
		closed	21	2	13	61.90
Total		open	2,357	1,191	537	22.94
		closed	335	38	211	64.02

Table 98

Use of Open- vs. Closed-Class Metaphors in the Written Post-Test: Control Group

B2 groups	Participants	MRW class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	open	97	48	22	22.68
		closed	13	1	8	61.54
	ST0302	open	143	78	25	17.48
		closed	15	1	13	86.67
	ST0303	open	89	48	13	14.61
		closed	13	3	9	69.23
	ST0304	open	82	48	16	19.51
		closed	14	2	7	50.00
	ST0305	open	112	67	23	20.54
		closed	16	3	9	56.25
	ST0306	open	117	61	25	21.37
		closed	20	2	13	65.00
	ST0307	open	113	62	22	19.47
		closed	15	1	6	40.00
	ST0308	open	115	67	18	15.65
		closed	13	0	9	69.23
	ST0309	open	102	61	22	21.57
		closed	13	0	7	53.85
	ST0310	open	122	71	14	11.48
		closed	17	2	10	58.82
GR0204	ST0401	open	147	69	31	21.09
		closed	16	0	14	87.50
	ST0402	open	140	69	32	22.86
		closed	24	2	17	70.83
	ST0403	open	134	87	24	17.91
		closed	19	3	14	73.68
	ST0404	open	99	55	21	21.21
		closed	12	0	5	41.67
	ST0405	open	116	72	15	12.93
		closed	14	3	7	50.00
	ST0406	open	111	66	17	15.32
		closed	18	5	7	38.89
	ST0407	open	120	71	17	14.17
		closed	19	3	10	52.63
	ST0408	open	109	53	22	20.18
		closed	16	3	6	37.50
	ST0409	open	129	66	27	20.93
		closed	16	3	12	75.00
	ST0410	open	104	50	24	23.08
		closed	14	2	6	42.86
Total		open	2,301	1,269	430	18.70
		closed	317	39	189	59.06

Table 99

Use of Metaphors by Word Class in Learner Oral Discourse (Pre-Test): Experimental Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	nouns	19	14	5	26.32
		verbs	29	8	4	13.79
		adjectives	9	8	1	11.11
		adverbs	22	20	0	0.00
	ST0102	nouns	45	34	10	22.22
		verbs	55	25	9	16.36
		adjectives	7	6	1	14.29
		adverbs	17	12	0	0.00
	ST0103	nouns	40	30	10	25.00
		verbs	56	23	11	19.64
		adjectives	5	3	2	40.00
		adverbs	33	18	2	6.06
	ST0104	nouns	30	21	9	30.00
		verbs	45	20	5	11.11
		adjectives	7	7	0	0.00
		adverbs	17	7	2	11.76
	ST0105	nouns	35	20	15	42.86
		verbs	97	43	9	9.28
		adjectives	14	8	6	42.86
		adverbs	38	25	0	0.00
	ST0106	nouns	18	14	4	22.22
		verbs	34	15	8	23.53
		adjectives	7	5	2	28.57
		adverbs	15	12	0	0.00
	ST0107	nouns	31	23	8	25.81
		verbs	64	27	9	14.06
		adjectives	12	8	4	33.33
		adverbs	30	21	3	10.00
	ST0108	nouns	36	33	3	8.33
		verbs	76	37	11	14.47
		adjectives	11	10	1	9.09
		adverbs	34	21	4	11.76
	ST0109	nouns	32	27	5	15.63
		verbs	55	17	9	16.36
		adjectives	9	4	5	55.56
		adverbs	21	14	0	0.00
	ST0110	nouns	25	19	6	24.00
		verbs	34	17	4	11.76
		adjectives	8	7	1	12.50
		adverbs	21	14	0	0.00
GR0102	ST0201	nouns	41	31	10	24.39
		verbs	54	29	6	11.11
		adjectives	2	2	0	0.00
		adverbs	15	10	2	13.33
	ST0202	nouns	18	14	4	22.22
		verbs	45	21	8	17.78
		adjectives	6	5	1	16.67
		adverbs	14	6	1	7.14
	ST0203	nouns	41	31	10	24.39
		verbs	54	30	7	12.96
		adjectives	6	5	1	16.67
		adverbs	37	33	1	2.70
	ST0204	nouns	21	18	3	14.29
		verbs	47	25	6	12.77

		adjectives	4	4	0	0.00
		adverbs	13	6	1	7.69
	ST0205	nouns	25	18	7	28.00
		verbs	40	12	14	35.00
		adjectives	5	4	1	20.00
		adverbs	14	13	0	0.00
	ST0206	nouns	36	25	11	30.56
		verbs	61	30	8	13.11
		adjectives	6	6	0	0.00
		adverbs	17	11	0	0.00
	ST0207	nouns	21	17	4	19.05
		verbs	36	21	5	13.89
		adjectives	5	5	0	0.00
		adverbs	11	8	2	18.18
	ST0208	nouns	41	34	7	17.07
		verbs	87	45	13	14.94
		adjectives	8	8	0	0.00
		adverbs	29	19	4	13.79
	ST0209	nouns	28	22	6	21.43
		verbs	39	14	6	15.38
		adjectives	6	6	0	0.00
		adverbs	22	18	0	0.00
	ST0210	nouns	34	28	6	17.65
		verbs	67	41	4	5.97
		adjectives	2	2	0	0.00
		adverbs	20	9	0	0.00
Total		nouns	617	473	143	23.07
		verbs	1,075	500	156	15.17
		adjectives	139	113	26	15.03
		adverbs	440	297	22	5.12

Table 100

Use of Metaphors by Word Class in Learner Oral Discourse (Pre-Test): Control Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	nouns	22	17	5	22.73
		verbs	39	16	5	12.82
		adjectives	4	3	1	25.00
		adverbs	19	14	0	0.00
	ST0302	nouns	31	20	11	35.48
		verbs	22	7	9	40.91
		adjectives	2	0	2	100.00
		adverbs	9	5	1	11.11
	ST0303	nouns	34	30	4	11.76
		verbs	39	17	11	28.21
		adjectives	2	1	1	50.00
		adverbs	11	6	0	0.00
	ST0304	nouns	31	24	7	22.58
		verbs	38	17	5	13.16
		adjectives	6	4	2	33.33
		adverbs	20	15	0	0.00
	ST0305	nouns	22	15	7	31.82
		verbs	36	18	5	13.89
		adjectives	4	2	2	50.00
		adverbs	16	8	0	0.00
	ST0306	nouns	32	23	8	25.00
		verbs	49	21	10	20.41
		adjectives	4	3	1	25.00
		adverbs	16	5	0	0.00
	ST0307	nouns	20	15	5	25.00
		verbs	38	21	6	15.79
		adjectives	3	2	1	33.33
		adverbs	13	10	1	7.69
	ST0308	nouns	33	22	11	33.33
		verbs	45	20	12	26.67
		adjectives	5	4	1	20.00
		adverbs	15	10	0	0.00
	ST0309	nouns	28	21	7	25.00
		verbs	48	23	14	29.17
		adjectives	3	1	2	66.67
		adverbs	18	15	0	0.00
	ST0310	nouns	27	18	9	33.33
		verbs	62	32	9	14.52
		adjectives	2	1	1	50.00
		adverbs	13	6	1	7.69
GR0204	ST0401	nouns	21	17	4	19.05
		verbs	41	19	7	17.07
		adjectives	4	1	3	75.00
		adverbs	15	11	0	0.00
	ST0402	nouns	11	9	2	18.18
		verbs	48	19	6	12.50
		adjectives	8	8	0	0.00
		adverbs	17	12	1	5.88
	ST0403	nouns	31	27	4	12.90
		verbs	42	18	5	11.90
		adjectives	8	7	1	12.50
		adverbs	20	11	0	0.00
	ST0404	nouns	25	19	6	24.00
		verbs	47	21	5	10.64

		adjectives	6	6	0	0.00
		adverbs	20	15	1	5.00
	ST0405	nouns	31	21	10	32.26
		verbs	46	17	11	23.91
		adjectives	1	1	0	0.00
		adverbs	18	14	1	5.56
	ST0406	nouns	24	16	8	33.33
		verbs	58	27	9	15.52
		adjectives	8	6	2	25.00
		adverbs	33	26	0	0.00
	ST0407	nouns	29	21	8	27.59
		verbs	43	18	7	16.28
		adjectives	9	7	2	22.22
		adverbs	13	10	0	0.00
	ST0408	nouns	21	14	7	33.33
		verbs	40	12	7	17.50
		adjectives	6	4	2	33.33
		adverbs	22	14	0	0.00
	ST0209	nouns	24	22	2	8.33
		verbs	57	27	5	8.77
		adjectives	10	7	3	30.00
		adverbs	30	24	2	6.67
	ST0410	nouns	23	20	3	13.04
		verbs	45	24	5	11.11
		adjectives	3	1	2	66.67
		adverbs	15	9	1	6.67
Total		nouns	520	391	128	24.40
		verbs	883	394	153	18.04
		adjectives	98	69	29	35.90
		adverbs	353	240	9	2.81

Table 101

Use of Metaphors by Word Class in Learner Written Discourse (Pre-Test): Experimental Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	nouns	37	32	5	13.51
		verbs	40	12	12	30.00
		adjectives	12	8	4	33.33
		adverbs	16	10	2	12.50
	ST0102	nouns	43	38	5	11.63
		verbs	33	8	14	42.42
		adjectives	16	15	1	6.25
		adverbs	19	18	0	0.00
	ST0103	nouns	35	27	8	22.86
		verbs	45	9	13	28.89
		adjectives	15	10	5	33.33
		adverbs	24	17	2	8.33
	ST0104	nouns	32	27	5	15.63
		verbs	45	16	14	31.11
		adjectives	18	17	1	5.56
		adverbs	15	11	2	13.33
	ST0105	nouns	35	32	2	5.71
		verbs	42	13	11	26.19
		adjectives	16	15	1	6.25
		adverbs	17	12	3	17.65
	ST0106	nouns	42	31	11	26.19
		verbs	36	9	12	33.33
		adjectives	16	13	3	18.75
		adverbs	18	14	0	0.00
	ST0107	nouns	34	24	10	29.41
		verbs	54	18	15	27.78
		adjectives	22	16	6	27.27
		adverbs	28	23	3	10.71
	ST0108	nouns	30	26	4	13.33
		verbs	37	16	6	16.22
		adjectives	16	9	7	43.75
		adverbs	15	13	1	6.67
	ST0109	nouns	49	40	9	18.37
		verbs	37	15	9	24.32
		adjectives	24	19	5	20.83
		adverbs	20	14	2	10.00
	ST0110	nouns	39	34	5	12.82
		verbs	34	12	12	35.29
		adjectives	6	6	0	0.00
		adverbs	22	18	3	13.64
GR0102	ST0201	nouns	33	24	9	27.27
		verbs	29	12	3	10.34
		adjectives	1	1	0	0.00
		adverbs	6	3	1	16.67
	ST0202	nouns	36	29	7	19.44
		verbs	40	10	6	15.00
		adjectives	12	12	0	0.00
		adverbs	17	11	2	11.76
	ST0203	nouns	39	32	7	17.95
		verbs	40	16	9	22.50
		adjectives	14	13	1	7.14
		adverbs	28	24	1	3.57
	ST0204	nouns	31	28	3	9.68
		Verbs	36	12	13	36.11

		adjectives	15	12	3	20.00
		adverbs	20	18	1	5.00
	ST0205	nouns	38	34	4	10.53
		Verbs	42	14	12	28.57
		adjectives	15	14	1	6.67
		adverbs	13	10	0	0.00
	ST0206	nouns	33	27	6	18.18
		Verbs	34	9	10	29.41
		adjectives	16	14	2	12.50
		adverbs	16	13	2	12.50
	ST0207	nouns	48	40	8	16.67
		Verbs	41	23	13	31.71
		adjectives	17	17	0	0.00
		adverbs	12	9	1	8.33
	ST0208	nouns	40	33	7	17.50
		Verbs	42	15	8	19.05
		adjectives	16	13	3	18.75
		adverbs	15	12	1	6.67
	ST0209	nouns	36	28	8	22.22
		Verbs	34	10	15	44.12
		adjectives	7	7	0	0.00
		adverbs	11	10	0	0.00
	ST0210	nouns	35	28	7	20.00
		Verbs	30	14	9	30.00
		adjectives	13	9	4	30.77
		adverbs	16	12	1	6.25
Total		nouns	745	614	130	17.45
		Verbs	771	263	216	28.12
		adjectives	287	240	47	14.56
		adverbs	348	272	28	8.18

Table 102

Use of Metaphors by Word Class in Learner Written Discourse (Pre-Test): Control Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	nouns	36	29	7	19.44
		Verbs	29	9	7	24.14
		adjectives	11	9	2	18.18
		adverbs	10	9	0	0.00
	ST0302	nouns	34	30	4	11.76
		Verbs	39	15	6	15.38
		adjectives	4	4	0	0.00
		adverbs	23	17	0	0.00
	ST0303	nouns	33	29	4	12.12
		Verbs	30	10	10	33.33
		adjectives	16	14	2	12.50
		adverbs	14	13	0	0.00
	ST0304	nouns	29	25	4	13.79
		Verbs	29	13	7	24.14
		adjectives	10	6	4	40.00
		adverbs	9	8	0	0.00
	ST0305	nouns	37	30	5	13.51
		Verbs	38	14	7	18.42
		adjectives	7	3	4	57.14
		adverbs	16	13	2	12.50
	ST0306	nouns	37	33	4	10.81
		Verbs	37	10	13	35.14
		adjectives	9	7	2	22.22
		adverbs	14	10	0	0.00
	ST0307	nouns	39	35	4	10.26
		Verbs	33	13	7	21.21
		adjectives	15	14	1	6.67
		adverbs	18	15	2	11.11
	ST0308	nouns	41	37	4	9.76
		Verbs	30	13	6	20.00
		adjectives	10	9	1	10.00
		adverbs	18	14	2	11.11
	ST0309	nouns	25	24	1	4.00
		Verbs	27	6	8	29.63
		adjectives	8	8	0	0.00
		adverbs	17	16	1	5.88
	ST0310	nouns	41	35	6	14.63
		Verbs	33	8	11	33.33
		adjectives	18	18	0	0.00
		adverbs	22	15	1	4.55
GR0204	ST0401	nouns	53	42	11	20.75
		Verbs	52	12	15	28.85
		adjectives	11	9	2	18.18
		adverbs	18	11	1	5.56
	ST0402	nouns	29	24	5	17.24
		Verbs	27	5	12	44.44
		adjectives	10	8	2	20.00
		adverbs	14	12	1	7.14
	ST0403	nouns	31	27	4	12.90
		Verbs	41	11	11	26.83
		adjectives	16	14	2	12.50
		adverbs	18	10	2	11.11
	ST0404	nouns	42	32	10	23.81
		verbs	51	16	16	31.37

		adjectives	19	17	2	10.53
		adverbs	19	13	3	15.79
	ST0405	nouns	37	29	8	21.62
		verbs	33	10	12	36.36
		adjectives	10	8	2	20.00
		adverbs	20	16	3	15.00
	ST0406	nouns	26	21	5	19.23
		verbs	43	17	7	16.28
		adjectives	15	12	3	20.00
		adverbs	17	13	1	5.88
	ST0407	nouns	59	47	12	20.34
		verbs	42	18	10	23.81
		adjectives	23	21	2	8.70
		adverbs	9	7	1	11.11
	ST0408	nouns	40	27	13	32.50
		verbs	48	17	10	20.83
		adjectives	16	13	3	18.75
		adverbs	11	6	1	9.09
	ST0209	nouns	47	42	5	10.64
		verbs	29	13	7	24.14
		adjectives	10	10	0	0.00
		adverbs	12	5	2	16.67
	ST0410	nouns	43	37	5	11.63
		verbs	28	11	4	14.29
		adjectives	15	12	3	20.00
		adverbs	16	13	1	6.25
Total		nouns	759	635	121	15.54
		verbs	719	241	186	26.10
		adjectives	253	216	37	15.77
		adverbs	315	236	24	7.44

Table 103

Use of Metaphors by Word Class in Learner Oral Discourse (Post-Test): Experimental Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	nouns	23	18	5	21.74
		verbs	50	24	10	20.00
		adjectives	7	7	0	0.00
		adverbs	26	22	3	11.54
	ST0102	nouns	49	31	18	36.73
		verbs	67	26	16	23.88
		adjectives	13	8	5	38.46
		adverbs	18	12	3	16.67
	ST0103	nouns	27	20	7	25.93
		verbs	69	27	7	10.14
		adjectives	11	8	3	27.27
		adverbs	28	17	1	3.57
	ST0104	nouns	27	21	6	22.22
		verbs	43	21	5	11.63
		adjectives	10	9	1	10.00
		adverbs	18	15	1	5.56
	ST0105	nouns	48	28	20	41.67
		verbs	64	21	15	23.44
		adjectives	16	8	8	50.00
		adverbs	45	30	8	17.78
	ST0106	nouns	28	25	3	10.71
		verbs	45	15	12	26.67
		adjectives	11	9	2	18.18
		adverbs	17	13	3	17.65
	ST0107	nouns	27	23	4	14.81
		verbs	47	20	8	17.02
		adjectives	8	8	0	0.00
		adverbs	30	24	0	0.00
	ST0108	nouns	28	25	3	10.71
		verbs	65	22	10	15.38
		adjectives	9	7	2	22.22
		adverbs	40	27	5	12.50
	ST0109	nouns	38	33	5	13.16
		verbs	39	14	7	17.95
		adjectives	14	11	3	21.43
		adverbs	24	19	1	4.17
	ST0110	nouns	21	17	4	19.05
		verbs	36	18	7	19.44
		adjectives	6	6	0	0.00
		adverbs	12	10	1	8.33
GR0102	ST0201	nouns	34	29	5	14.71
		verbs	54	23	9	16.67
		adjectives	2	2	0	0.00
		adverbs	22	17	1	4.55
	ST0202	nouns	17	10	7	41.18
		verbs	29	14	7	24.14
		adjectives	3	3	0	0.00
		adverbs	7	4	3	42.86
	ST0203	nouns	39	28	11	28.21
		verbs	68	25	14	20.59
		adjectives	10	9	1	10.00
		adverbs	33	23	5	15.15
	ST0204	nouns	33	27	6	18.18
		verbs	56	21	14	25.00

		adjectives	3	3	0	0.00
		adverbs	25	17	4	16.00
	ST0205	nouns	28	19	8	28.57
		verbs	43	15	7	16.28
		adjectives	10	8	2	20.00
		adverbs	18	13	1	5.56
	ST0206	nouns	34	30	4	11.76
		verbs	38	21	8	21.05
		adjectives	6	6	0	0.00
		adverbs	16	12	2	12.50
	ST0207	nouns	25	19	6	24.00
		verbs	26	10	5	19.23
		adjectives	4	3	1	25.00
		adverbs	5	5	0	0.00
	ST0208	nouns	28	22	6	21.43
		verbs	51	18	10	19.61
		adjectives	12	9	3	25.00
		adverbs	14	11	2	14.29
	ST0209	nouns	21	19	2	9.52
		verbs	32	13	7	21.88
		adjectives	8	7	1	12.50
		adverbs	15	12	3	20.00
	ST0210	nouns	16	10	6	37.50
		verbs	37	19	0	0.00
		adjectives	7	7	0	0.00
		adverbs	11	9	0	0.00
Total		nouns	591	454	136	22.59
		verbs	959	387	178	18.50
		adjectives	170	138	32	14.00
		adverbs	424	312	47	11.43

Table 104

Use of Metaphors by Word Class in Learner Oral Discourse (Post-Test): Control Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	nouns	19	14	5	26.32
		verbs	30	9	6	20.00
		adjectives	7	4	3	42.86
		adverbs	15	11	3	20.00
	ST0302	nouns	11	8	3	27.27
		verbs	25	12	5	20.00
		adjectives	2	2	0	0.00
		adverbs	12	8	2	16.67
	ST0303	nouns	16	15	1	6.25
		verbs	36	20	4	11.11
		adjectives	2	2	0	0.00
		adverbs	17	13	2	11.76
	ST0304	nouns	20	16	3	15.00
		verbs	22	9	1	4.55
		adjectives	2	2	0	0.00
		adverbs	8	7	0	0.00
	ST0305	nouns	20	15	5	25.00
		verbs	31	17	2	6.45
		adjectives	0	0	0	0.00
		adverbs	13	10	0	0.00
	ST0306	nouns	27	20	6	22.22
		verbs	40	15	8	20.00
		adjectives	6	5	1	16.67
		adverbs	19	11	2	10.53
	ST0307	nouns	29	24	5	17.24
		verbs	45	14	12	26.67
		adjectives	5	4	1	20.00
		adverbs	11	7	0	0.00
	ST0308	nouns	20	16	4	20.00
		verbs	35	15	6	17.14
		adjectives	8	7	1	12.50
		adverbs	11	6	0	0.00
	ST0309	nouns	14	9	5	35.71
		verbs	31	10	6	19.35
		adjectives	4	3	1	25.00
		adverbs	16	10	1	6.25
	ST0310	nouns	19	19	0	0.00
		verbs	32	16	2	6.25
		adjectives	6	6	0	0.00
		adverbs	10	5	0	0.00
GR0204	ST0401	nouns	25	22	3	12.00
		verbs	41	15	9	21.95
		adjectives	3	2	1	33.33
		adverbs	14	11	2	14.29
	ST0402	nouns	17	13	3	17.65
		verbs	32	14	5	15.63
		adjectives	4	4	0	0.00
		adverbs	13	10	1	7.69
	ST0403	nouns	26	20	6	23.08
		verbs	55	23	11	20.00
		adjectives	7	6	1	14.29
		adverbs	34	22	3	8.82
	ST0404	nouns	11	8	3	27.27
		verbs	30	13	4	13.33

		adjectives	2	2	0	0.00
		adverbs	9	9	0	0.00
	ST0405	nouns	26	15	11	42.31
		verbs	41	19	10	24.39
		adjectives	4	3	1	25.00
		adverbs	16	14	0	0.00
	ST0406	nouns	29	23	5	17.24
		verbs	54	25	6	11.11
		adjectives	13	9	4	30.77
		adverbs	27	21	3	11.11
	ST0407	nouns	20	16	4	20.00
		verbs	36	12	8	22.22
		adjectives	7	7	0	0.00
		adverbs	15	8	3	20.00
	ST0408	nouns	21	16	5	23.81
		verbs	26	8	6	23.08
		adjectives	8	8	0	0.00
		adverbs	10	9	0	0.00
	ST0209	nouns	13	9	4	30.77
		verbs	35	16	3	8.57
		adjectives	7	5	2	28.57
		adverbs	16	14	0	0.00
	ST0410	nouns	14	11	3	21.43
		verbs	29	13	2	6.90
		adjectives	3	3	0	0.00
		adverbs	10	9	0	0.00
Total		nouns	397	309	84	21.53
		verbs	706	295	116	15.94
		adjectives	100	85	16	12.45
		adverbs	296	215	22	6.36

Table 105

Use of Metaphors by Word Class in Learner Written Discourse (Post-Test): Experimental Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0101	ST0101	nouns	35	27	8	22.86
		verbs	51	12	5	9.80
		adjectives	17	8	9	52.94
		adverbs	19	14	2	10.53
	ST0102	nouns	35	24	11	31.43
		verbs	43	9	10	23.26
		adjectives	10	8	2	20.00
		adverbs	13	9	2	15.38
	ST0103	nouns	33	21	12	36.36
		verbs	52	15	11	21.15
		adjectives	6	3	3	50.00
		adverbs	14	8	1	7.14
	ST0104	nouns	42	32	10	23.81
		verbs	63	9	15	23.81
		adjectives	11	11	0	0.00
		adverbs	17	12	2	11.76
	ST0105	nouns	57	36	21	36.84
		verbs	54	17	12	22.22
		adjectives	10	9	1	10.00
		adverbs	20	15	2	10.00
	ST0106	nouns	33	19	14	42.42
		verbs	48	14	11	22.92
		adjectives	15	11	4	26.67
		adverbs	12	8	1	8.33
	ST0107	nouns	31	18	13	41.94
		verbs	49	11	13	26.53
		adjectives	11	8	3	27.27
		adverbs	13	10	2	15.38
	ST0108	nouns	32	23	9	28.13
		verbs	49	13	14	28.57
		adjectives	12	8	4	33.33
		adverbs	27	24	2	7.41
	ST0109	nouns	40	30	10	25.00
		verbs	45	4	9	20.00
		adjectives	10	6	4	40.00
		adverbs	9	6	2	22.22
	ST0110	nouns	38	26	12	31.58
		verbs	49	14	8	16.33
		adjectives	7	5	2	28.57
		adverbs	26	19	3	11.54
GR0102	ST0201	nouns	38	23	15	39.47
		verbs	48	13	8	16.67
		adjectives	5	4	1	20.00
		adverbs	13	12	0	0.00
	ST0202	nouns	31	19	12	38.71
		verbs	45	14	7	15.56
		adjectives	8	6	2	25.00
		adverbs	11	5	1	9.09
	ST0203	nouns	36	24	12	33.33
		verbs	51	18	10	19.61
		adjectives	8	7	1	12.50
		adverbs	17	11	2	11.76
	ST0204	nouns	42	31	11	26.19
		verbs	56	19	9	16.07

		adjectives	13	11	2	15.38
		adverbs	14	9	1	7.14
	ST0205	nouns	38	31	7	18.42
		verbs	61	10	8	13.11
		adjectives	18	17	1	5.56
		adverbs	17	15	0	0.00
	ST0206	nouns	34	26	8	23.53
		verbs	57	10	13	22.81
		adjectives	10	8	2	20.00
		adverbs	19	16	3	15.79
	ST0207	nouns	45	32	13	28.89
		verbs	52	16	7	13.46
		adjectives	13	9	4	30.77
		adverbs	11	5	0	0.00
	ST0208	nouns	60	44	16	26.67
		verbs	64	12	16	25.00
		adjectives	21	17	4	19.05
		adverbs	14	11	1	7.14
	ST0209	nouns	32	21	11	34.38
		verbs	52	10	9	17.31
		adjectives	10	8	2	20.00
		adverbs	14	10	0	0.00
	ST0210	nouns	50	26	24	48.00
		verbs	46	17	10	21.74
		adjectives	16	12	4	25.00
		adverbs	9	6	1	11.11
Total		nouns	782	533	249	31.90
		verbs	1,035	257	205	19.80
		adjectives	231	176	55	24.10
		adverbs	309	225	28	9.09

Table 106

Use of Metaphors by Word Class in Learner Written Discourse (Post-Test): Control Group

B2 groups	Participants	Word class	LUs (tokens)	Non-MRWs (tokens)	MRWs (tokens)	Metaphor density (%)
GR0203	ST0301	nouns	38	26	12	31.58
		verbs	37	6	7	18.92
		adjectives	13	11	2	15.38
		adverbs	9	5	1	11.11
	ST0302	nouns	46	36	10	21.74
		verbs	65	16	13	20.00
		adjectives	11	11	0	0.00
		adverbs	21	15	2	9.52
	ST0303	nouns	27	22	5	18.52
		verbs	42	11	5	11.90
		adjectives	12	9	3	25.00
		adverbs	8	6	0	0.00
	ST0304	nouns	31	24	7	22.58
		verbs	30	8	6	20.00
		adjectives	9	6	3	33.33
		adverbs	12	10	0	0.00
	ST0305	nouns	40	27	13	32.50
		verbs	43	15	7	16.28
		adjectives	12	10	2	16.67
		adverbs	17	15	1	5.88
	ST0306	nouns	41	29	12	29.27
		verbs	48	13	13	27.08
		adjectives	8	8	0	0.00
		adverbs	20	11	0	0.00
	ST0307	nouns	43	34	9	20.93
		verbs	49	13	9	18.37
		adjectives	11	7	4	36.36
		adverbs	10	8	0	0.00
	ST0308	nouns	39	30	9	23.08
		verbs	54	17	8	14.81
		adjectives	10	9	1	10.00
		adverbs	12	11	0	0.00
	ST0309	nouns	42	28	14	33.33
		verbs	30	7	5	16.67
		adjectives	16	15	1	6.25
		adverbs	14	11	2	14.29
	ST0310	nouns	37	30	7	18.92
		verbs	57	21	4	7.02
		adjectives	10	9	1	10.00
		adverbs	18	11	2	11.11
GR0204	ST0401	nouns	47	32	15	31.91
		verbs	73	14	14	19.18
		adjectives	9	8	1	11.11
		adverbs	18	15	1	5.56
	ST0402	nouns	51	38	13	25.49
		verbs	62	15	11	17.74
		adjectives	9	5	4	44.44
		adverbs	18	11	4	22.22
	ST0403	nouns	50	38	12	24.00
		verbs	45	12	11	24.44
		adjectives	13	13	0	0.00
		adverbs	26	24	1	3.85
	ST0404	nouns	31	20	11	35.48
		verbs	40	14	5	12.50

		adjectives	10	6	4	40.00
		adverbs	18	15	1	5.56
	ST0405	nouns	41	31	10	24.39
		verbs	50	19	4	8.00
		adjectives	9	9	0	0.00
		adverbs	16	13	1	6.25
	ST0406	nouns	37	31	6	16.22
		verbs	42	11	7	16.67
		adjectives	15	11	3	20.00
		adverbs	17	13	1	5.88
	ST0407	nouns	41	33	8	19.51
		verbs	51	15	7	13.73
		adjectives	15	14	1	6.67
		adverbs	13	9	1	7.69
	ST0408	nouns	37	23	14	37.84
		verbs	46	13	5	10.87
		adjectives	10	8	2	20.00
		adverbs	16	9	1	6.25
	ST0209	nouns	40	24	16	40.00
		verbs	50	11	7	14.00
		adjectives	18	15	3	16.67
		adverbs	21	16	1	4.76
	ST0410	nouns	33	20	13	39.39
		verbs	49	13	7	14.29
		adjectives	10	7	3	30.00
		adverbs	12	10	1	8.33
Total		nouns	792	576	216	27.33
		verbs	963	264	155	16.12
		adjectives	230	191	38	17.09
		adverbs	316	238	21	6.41

Table 107

Metaphor Types Variations in Learner Oral Discourse (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	MRW class	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
					AI	RI	%
GR0101	ST0101	open	12.66	16.98	4.32	0.34	34%
		closed	33.33	28.57	-4.76	-0.14	-14%
	ST0102	open	16.13	28.57	12.44	0.77	77%
		closed	29.17	34.38	5.21	0.18	18%
	ST0103	open	18.66	13.33	-5.32	-0.29	-29%
		closed	30.00	30.43	0.43	0.01	1%
	ST0104	open	16.16	13.27	-2.90	-0.18	-18%
		closed	66.67	40.00	-26.67	-0.40	-40%
	ST0105	open	16.30	29.48	13.18	0.81	81%
		closed	36.84	34.48	-2.36	-0.06	-6%
	ST0106	open	18.92	19.80	0.88	0.05	5%
		closed	50.00	26.32	-23.68	-0.47	-47%
	ST0107	open	17.52	10.71	-6.80	-0.39	-39%
		closed	33.33	35.00	1.67	0.05	5%
	ST0108	open	12.10	14.08	1.98	0.16	16%
		closed	64.29	23.53	-40.76	-0.63	-63%
	ST0109	open	16.24	13.91	-2.33	-0.14	-14%
		closed	53.85	27.59	-26.26	-0.49	-49%
	ST0110	open	12.50	16.00	3.50	0.28	28%
		closed	50.00	21.43	-28.57	-0.57	-57%
GR0102	ST0201	open	16.07	13.39	-2.68	-0.17	-17%
		closed	63.64	31.58	-32.06	-0.50	-50%
	ST0202	open	16.87	30.36	13.49	0.80	80%
		closed	100.00	16.67	-83.33	-0.83	-83%
	ST0203	open	13.77	20.67	6.90	0.50	50%
		closed	52.94	40.00	-12.94	-0.24	-24%
	ST0204	open	11.76	20.51	8.75	0.74	74%
		closed	57.14	20.00	-37.14	-0.65	-65%
	ST0205	open	26.19	18.18	-8.01	-0.31	-31%
		closed	64.71	26.67	-38.04	-0.59	-59%
	ST0206	open	15.83	14.89	-0.94	-0.06	-6%
		closed	62.50	35.71	-26.79	-0.43	-43%
	ST0207	open	15.07	20.00	4.93	0.33	33%
		closed	72.22	28.57	-43.65	-0.60	-60%
	ST0208	open	14.55	20.00	5.45	0.38	38%
		closed	36.00	18.75	-17.25	-0.48	-48%
	ST0209	open	12.63	17.11	4.47	0.35	35%
		closed	46.15	47.06	0.90	0.02	2%
	ST0210	open	8.13	8.45	0.32	0.04	4%
		closed	68.75	33.33	-35.42	-0.52	-52%
Total		open	15.40	17.99	2.59	0.20	20%
		closed	53.58	30.00	-23.58	-0.37	-37%

Table 108

Metaphor Types Variations in Learner Oral Discourse (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	MRW class	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
					AI	RI	%
GR0203	ST0301	open	13.10	23.94	10.85	0.83	83%
		closed	50.00	30.77	-19.23	-0.38	-38%
	ST0302	open	35.94	20.00	-15.94	-0.44	-44%
		closed	40.00	50.00	10.00	0.25	25%
	ST0303	open	18.60	9.86	-8.75	-0.47	-47%
		closed	72.73	30.00	-42.73	-0.59	-59%
	ST0304	open	14.74	7.69	-7.04	-0.48	-48%
		closed	31.25	40.00	8.75	0.28	28%
	ST0305	open	17.95	10.94	-7.01	-0.39	-39%
		closed	70.00	20.00	-50.00	-0.71	-71%
	ST0306	open	18.81	18.48	-0.33	-0.02	-2%
		closed	50.00	43.75	-6.25	-0.13	-13%
	ST0307	open	17.57	20.00	2.43	0.14	14%
		closed	38.46	11.76	-26.70	-0.69	-69%
	ST0308	open	24.49	14.86	-9.62	-0.39	-39%
		closed	66.67	13.33	-53.33	-0.80	-80%
	ST0309	open	23.71	20.00	-3.71	-0.16	-16%
		closed	50.00	37.50	-12.50	-0.25	-25%
	ST0310	open	19.23	2.99	-16.25	-0.84	-84%
		closed	27.78	60.00	32.22	1.16	116%
GR0204	ST0401	open	17.28	18.07	0.79	0.05	5%
		closed	83.33	50.00	-33.33	-0.40	-40%
	ST0402	open	10.71	13.64	2.92	0.27	27%
		closed	62.50	12.50	-50.00	-0.80	-80%
	ST0403	open	9.90	17.21	7.31	0.74	74%
		closed	57.14	36.36	-20.78	-0.36	-36%
	ST0404	open	12.24	13.46	1.22	0.10	10%
		closed	53.85	28.57	-25.27	-0.47	-47%
	ST0405	open	22.92	25.29	2.37	0.10	10%
		closed	29.41	12.50	-16.91	-0.58	-58%
	ST0406	open	15.45	14.63	-0.81	-0.05	-5%
		closed	70.59	23.53	-47.06	-0.67	-67%
	ST0407	open	18.09	19.23	1.15	0.06	6%
		closed	25.00	28.57	3.57	0.14	14%
	ST0408	open	17.98	16.92	-1.05	-0.06	-6%
		closed	80.00	27.27	-52.73	-0.66	-66%
	ST0409	open	9.92	12.68	2.76	0.28	28%
		closed	44.44	20.00	-24.44	-0.55	-55%
	ST0410	open	12.79	8.93	-3.86	-0.30	-30%
		closed	37.50	45.45	7.95	0.21	21%
Total		open	17.57	15.44	-2.13	-0.05	-5%
		closed	52.03	31.09	-20.94	-0.30	-30%

Table 109

Metaphor Types Variations in Learner Written Discourse (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	MRW class	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
					AI	RI	%
GR0101	ST0101	open	21.90	19.67	-2.23	-0.10	-10%
		closed	43.48	43.75	0.27	0.01	1%
	ST0102	open	18.02	24.75	6.73	0.37	37%
		closed	47.06	66.67	19.61	0.42	42%
	ST0103	open	23.53	25.71	2.18	0.09	9%
		closed	61.11	60.00	-1.11	-0.02	-2%
	ST0104	open	20.00	20.30	0.30	0.02	2%
		closed	33.33	84.62	51.28	1.54	154%
	ST0105	open	15.45	25.53	10.08	0.65	65%
		closed	63.64	62.07	-1.57	-0.02	-2%
	ST0106	open	23.21	27.78	4.56	0.20	20%
		closed	57.89	75.00	17.11	0.30	30%
	ST0107	open	24.64	29.81	5.17	0.21	21%
		closed	53.33	33.33	-20.00	-0.38	-38%
	ST0108	open	18.37	24.17	5.80	0.32	32%
		closed	50.00	60.00	10.00	0.20	20%
	ST0109	open	19.23	24.04	4.81	0.25	25%
		closed	42.11	78.95	36.84	0.88	88%
	ST0110	open	19.80	20.83	1.03	0.05	5%
		closed	70.59	46.67	-23.92	-0.34	-34%
GR0102	ST0201	open	18.84	23.08	4.24	0.22	22%
		closed	50.00	43.75	-6.25	-0.13	-13%
	ST0202	open	14.29	23.16	8.87	0.62	62%
		closed	54.55	58.82	4.28	0.08	8%
	ST0203	open	14.88	22.32	7.45	0.50	50%
		closed	54.55	63.64	9.09	0.17	17%
	ST0204	open	19.61	18.40	-1.21	-0.06	-6%
		closed	62.50	66.67	4.17	0.07	7%
	ST0205	open	15.74	11.94	-3.80	-0.24	-24%
		closed	52.63	84.62	31.98	0.61	61%
	ST0206	open	20.20	21.67	1.46	0.07	7%
		closed	66.67	75.00	8.33	0.13	13%
	ST0207	open	18.64	19.83	1.19	0.06	6%
		closed	68.42	72.22	3.80	0.06	6%
	ST0208	open	16.81	23.27	6.46	0.38	38%
		closed	57.14	54.55	-2.60	-0.05	-5%
	ST0209	open	26.14	20.37	-5.77	-0.22	-22%
		closed	82.35	88.24	5.88	0.07	7%
	ST0210	open	22.34	32.23	9.89	0.44	44%
		closed	40.00	61.90	21.90	0.55	55%
Total		open	19.58	22.94	3.36	0.19	19%
		closed	55.57	64.02	8.45	0.21	21%

Table 110

Metaphor Types Variations in Learner Written Discourse (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	MRW class	Pre-test MRW (%)	Post-test MRW (%)	Evolution rates		
					AI	RI	%
GR0203	ST0301	open	18.60	22.68	4.08	0.22	22%
		closed	40.00	61.54	21.54	0.54	54%
	ST0302	open	10.00	17.48	7.48	0.75	75%
		closed	75.00	86.67	11.67	0.16	16%
	ST0303	open	17.20	14.61	-2.60	-0.15	-15%
		closed	61.54	69.23	7.69	0.13	13%
	ST0304	open	19.48	19.51	0.03	0.00	0%
		closed	41.67	50.00	8.33	0.20	20%
	ST0305	open	18.37	20.54	2.17	0.12	12%
		closed	66.67	56.25	-10.42	-0.16	-16%
	ST0306	open	19.59	21.37	1.78	0.09	9%
		closed	55.00	65.00	10.00	0.18	18%
	ST0307	open	13.33	19.47	6.14	0.46	46%
		closed	70.00	40.00	-30.00	-0.43	-43%
	ST0308	open	13.13	15.65	2.52	0.19	19%
		closed	53.33	69.23	15.90	0.30	30%
	ST0309	open	12.99	21.57	8.58	0.66	66%
		closed	72.73	53.85	-18.88	-0.26	-26%
	ST0310	open	15.79	11.48	-4.31	-0.27	-27%
		closed	64.71	58.82	-5.88	-0.09	-9%
GR0204	ST0401	open	21.64	21.09	-0.55	-0.03	-3%
		closed	57.69	87.50	29.81	0.52	52%
	ST0402	open	25.00	22.86	-2.14	-0.09	-9%
		closed	55.56	70.83	15.28	0.28	28%
	ST0403	open	17.92	17.91	-0.01	0.00	0%
		closed	37.50	73.68	36.18	0.96	96%
	ST0404	open	23.66	21.21	-2.45	-0.10	-10%
		closed	75.00	41.67	-33.33	-0.44	-44%
	ST0405	open	25.00	12.93	-12.07	-0.48	-48%
		closed	50.00	50.00	0.00	0.00	0%
	ST0406	open	15.84	15.32	-0.53	-0.03	-3%
		closed	50.00	38.89	-11.11	-0.22	-22%
	ST0407	open	18.80	14.17	-4.63	-0.25	-25%
		closed	64.00	52.63	-11.37	-0.18	-18%
	ST0408	open	23.48	20.18	-3.29	-0.14	-14%
		closed	42.86	37.50	-5.36	-0.13	-13%
	ST0409	open	14.29	20.93	6.64	0.47	47%
		closed	47.83	75.00	27.17	0.57	57%
	ST0410	open	12.75	23.08	10.33	0.81	81%
		closed	40.00	42.86	2.86	0.07	7%
Total		open	17.84	18.70	0.86	0.11	11%
		closed	56.05	59.06	3.01	0.10	10%

Appendix P: MRW Prepositions in Learner Discourse

Table 111

Use of MRW Prepositions in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

MRW prepositions	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	in	50	37.31	in	52	33.77
	with	30	22.39	with	45	29.22
	on	19	14.18	on	21	13.64
	to	17	12.69	to	11	7.14
	from	8	5.97	about	8	5.19
	about	5	3.73	at	8	5.19
	at	3	2.24	from	4	2.60
	by	2	1.49	without	2	1.30
				after	1	0.65
				apart	1	0.65
				by	1	0.65
Total	8	134	100.00	11	154	100.00

Table 112

Use of MRW Prepositions in Learner Oral Discourse (Control vs. Experimental Groups): Post-Test

MRW prepositions	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	with	26	36.11	with	33	30.28
	about	11	15.28	in	25	22.94
	from	10	13.89	on	15	13.76
	in	9	12.50	about	13	11.93
	on	7	9.72	from	10	9.17
	to	5	6.94	to	9	8.26
	at	4	5.56	under	2	1.83
				between	1	0.92
				out	1	0.92
Total	7	72	100.00	9	109	100.00

Table 113

Use of MRW Prepositions in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

MRW prepositions	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	in	104	58.10	in	110	60.11
	with	20	11.17	with	22	12.02
	on	13	7.26	into	11	6.01
	about	10	5.59	to	10	5.46
	into	8	4.47	on	7	3.83
	to	7	3.91	without	6	3.28
	from	6	3.35	about	5	2.73
	at	2	1.12	at	3	1.64
	between	2	1.12	between	3	1.64
	by	2	1.12	from	3	1.64
	over	2	1.12	by	2	1.09
	after	1	0.56	after	1	0.55
	before	1	0.56			
	throughout	1	0.56			
Total	14	179	100.00	12	183	100.00

Table 114

Use of MRW Prepositions in Learner Written Discourse (Control vs. Experimental Groups): Post-Test

MRW prepositions	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	in	85	44.97	in	90	42.65
	with	34	17.99	with	24	11.37
	into	16	8.47	on	16	7.58
	to	13	6.88	about	15	7.11
	about	12	6.35	to	15	7.11
	from	10	5.29	from	12	5.69
	on	6	3.17	into	12	5.69
	at	5	2.65	without	10	4.74
	without	3	1.59	by	5	2.37
	after	2	1.06	at	3	1.42
	by	2	1.06	through	3	1.42
	through	1	0.53	between	2	0.95
				under	2	0.95
				over	1	0.47
				until	1	0.47
Total	12	189	100.00	15	211	100.00

Appendix Q: MRW Nouns in Learner Discourse

Table 115

Use of MRW Nouns in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

MRW nouns	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	time	39	30.47	thing	28	19.58
	thing	30	23.44	time	21	14.69
	problem	13	10.16	subject	20	13.99
	way	9	7.03	exercise	13	9.09
	exercise	7	5.47	way	8	5.59
	point	4	3.13	point	7	4.90
	subject	4	3.13	problem	7	4.90
	view	4	3.13	view	7	4.90
	hour	2	1.56	hand	5	3.50
	mark	2	1.56	future	4	2.80
	child	1	0.78	mark	4	2.80
	end	1	0.78	end	2	1.40
	extent	1	0.78	part	2	1.40
	fault	1	0.78	content	1	0.70
	hand	1	0.78	disconnection	1	0.70
	lesson	1	0.78	energy	1	0.70
	level	1	0.78	example	1	0.70
	middle	1	0.78	eye	1	0.70
	motivation	1	0.78	fact	1	0.70
	network	1	0.78	game	1	0.70
	part	1	0.78	obligation	1	0.70
	study	1	0.78	occasion	1	0.70
	tablet	1	0.78	resource	1	0.70
	year	1	0.78	schedule	1	0.70
				situation	1	0.70
				solution	1	0.70
				tablet	1	0.70
				week	1	0.70
Total	24	128	100.00	28	143	100.00

Table 116

Use of MRW Nouns in Learner Oral Discourse (Control vs. Experimental Groups): Post-Test

MRW nouns	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	time	21	25.00	thing	38	27.94
	thing	20	23.81	time	23	16.91
	point	9	10.71	point	7	5.15
	way	9	10.71	view	7	5.15
	view	7	8.33	way	7	5.15
	centre	2	2.38	problem	5	3.68
	damage	2	2.38	team	5	3.68
	freedom	2	2.38	life	3	2.21
	lesson	2	2.38	part	3	2.21
	child	1	1.19	air	2	1.47
	floor	1	1.19	country	2	1.47
	life	1	1.19	facility	2	1.47
	medium	1	1.19	key	2	1.47
	network	1	1.19	screen	2	1.47
	part	1	1.19	situation	2	1.47
	rest	1	1.19	weather	2	1.47
	rise	1	1.19	age	1	0.74
	series	1	1.19	centre	1	0.74
	world	1	1.19	down	1	0.74
				experience	1	0.74
				eye	1	0.74
				fact	1	0.74
				feeling	1	0.74
				floor	1	0.74
				group	1	0.74
				independence	1	0.74
				interest	1	0.74
				liberty	1	0.74
				medium	1	0.74
				nature	1	0.74
				relation	1	0.74
				routine	1	0.74
				spirit	1	0.74
				stuff	1	0.74
				subject	1	0.74
				thinking	1	0.74
				traffic	1	0.74
				up	1	0.74
				wall	1	0.74
				world	1	0.74
Total	19	84	100.00	40	136	100.00

Table 117

Use of MRW Nouns in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

MRW nouns	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	thing	16	13.22	thing	17	13.08
	hand	11	9.09	way	11	8.46
	problem	7	5.79	point	8	6.15
	way	7	5.79	life	7	5.38
	aim	6	4.96	contrast	6	4.62
	life	6	4.96	hand	6	4.62
	culture	4	3.31	view	6	4.62
	day	4	3.31	account	5	3.85
	level	4	3.31	culture	5	3.85
	point	4	3.31	trip	5	3.85
	time	4	3.31	club	4	3.08
	year	4	3.31	time	4	3.08
	advance	3	2.48	part	3	2.31
	place	3	2.48	possibility	3	2.31
	situation	3	2.48	access	2	1.54
	subject	3	2.48	communication	2	1.54
	trip	3	2.48	future	2	1.54
	child	2	1.65	idea	2	1.54
	range	2	1.65	mark	2	1.54
	view	2	1.65	place	2	1.54
	account	1	0.83	subject	2	1.54
	attention	1	0.83	team	2	1.54
	base	1	0.83	world	2	1.54
	communication	1	0.83	year	2	1.54
	conception	1	0.83	area	1	0.77
	condition	1	0.83	argument	1	0.77
	country	1	0.83	child	1	0.77
	course	1	0.83	click	1	0.77
	degree	1	0.83	day	1	0.77
	exchange	1	0.83	degree	1	0.77
	facility	1	0.83	dream	1	0.77
	future	1	0.83	fact	1	0.77
	goal	1	0.83	goal	1	0.77
	industry	1	0.83	impact	1	0.77
	lesson	1	0.83	knowledge	1	0.77
	part	1	0.83	level	1	0.77
	partner	1	0.83	network	1	0.77
	pass	1	0.83	period	1	0.77
	role	1	0.83	progress	1	0.77
	sign	1	0.83	resource	1	0.77
	system	1	0.83	series	1	0.77
	tablet	1	0.83	site	1	0.77
	variety	1	0.83	situation	1	0.77
				source	1	0.77
Total	43	121	100.00	44	130	100.00

Table 118

Use of MRW Nouns in Learner Written Discourse (Control vs. Experimental Groups): Post-Test

MRW nouns	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	subject	56	25.93	subject	51	20.48
	way	29	13.43	point	19	7.63
	point	26	12.04	way	19	7.63
	view	18	8.33	thing	14	5.62
	thing	14	6.48	view	12	4.82
	fact	6	2.78	time	11	4.42
	life	6	2.78	future	7	2.81
	time	6	2.78	role	7	2.81
	part	5	2.31	hand	6	2.41
	lesson	4	1.85	key	6	2.41
	tablet	3	1.39	system	6	2.41
	aim	2	0.93	form	5	2.01
	base	2	0.93	problem	5	2.01
	content	2	0.93	account	4	1.61
	device	2	0.93	material	4	1.61
	material	2	0.93	part	4	1.61
	place	2	0.93	air	3	1.20
	problem	2	0.93	fact	3	1.20
	progress	2	0.93	life	3	1.20
	robot	2	0.93	star	3	1.20
	account	1	0.46	tablet	3	1.20
	age	1	0.46	can	2	0.80
	behaviour	1	0.46	contrast	2	0.80
	capacity	1	0.46	eye	2	0.80
	centre	1	0.46	feeling	2	0.80
	course	1	0.46	lesson	2	0.80
	culture	1	0.46	trip	2	0.80
	diary	1	0.46	worm	2	0.80
	form	1	0.46	access	1	0.40
	freedom	1	0.46	aim	1	0.40
	future	1	0.46	animal	1	0.40
	hand	1	0.46	argument	1	0.40
	idea	1	0.46	break	1	0.40
	impact	1	0.46	cloud	1	0.40
	lockdown	1	0.46	concentration	1	0.40
	matter	1	0.46	degree	1	0.40
	period	1	0.46	device	1	0.40
	pillar	1	0.46	digital	1	0.40
	resource	1	0.46	discipline	1	0.40
	role	1	0.46	dream	1	0.40
	solution	1	0.46	dump	1	0.40
	trip	1	0.46	end	1	0.40
	variety	1	0.46	exercise	1	0.40
	world	1	0.46	experience	1	0.40
	year	1	0.46	facility	1	0.40
				field	1	0.40
				hill	1	0.40
				hurdle	1	0.40
				idea	1	0.40
				impression	1	0.40
				lead	1	0.40
				lockdown	1	0.40
				money	1	0.40

				nutshell	1	0.40
				path	1	0.40
				place	1	0.40
				platform	1	0.40
				progress	1	0.40
				relation	1	0.40
				resource	1	0.40
				revolution	1	0.40
				situation	1	0.40
				spirit	1	0.40
				thinking	1	0.40
				top	1	0.40
				water	1	0.40
				weather	1	0.40
				year	1	0.40
Total	45	216	100.00	68	249	100.00

Appendix R: MRW Verbs in Learner Discourse

Table 119

Use of MRW Verbs in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

MRW verbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	have	55	35.95	have	49	31.41
	give	16	10.46	give	22	14.10
	spend	8	5.23	go	7	4.49
	go	7	4.58	make	7	4.49
	solve	5	3.27	see	7	4.49
	think	5	3.27	send	6	3.85
	make	4	2.61	get	5	3.21
	put	4	2.61	spend	5	3.21
	see	4	2.61	pass	4	2.56
	take	4	2.61	prepare	4	2.56
	fail	3	1.96	put	4	2.56
	get	3	1.96	relax	4	2.56
	look	3	1.96	feel	3	1.92
	copy	2	1.31	copy	2	1.28
	enter	2	1.31	distract	2	1.28
	focus	2	1.31	do	2	1.28
	organise	2	1.31	review	2	1.28
	pay	2	1.31	solve	2	1.28
	reach	2	1.31	consider	1	0.64
	send	2	1.31	explode	1	0.64
	call	1	0.65	foment	1	0.64
	concentrate	1	0.65	grow	1	0.64
	connect	1	0.65	help	1	0.64
	feel	1	0.65	know	1	0.64
	follow	1	0.65	look	1	0.64
	grow	1	0.65	motivate	1	0.64
	help	1	0.65	organise	1	0.64
	inspire	1	0.65	realise	1	0.64
	keep	1	0.65	resolve	1	0.64
	know	1	0.65	share	1	0.64
	motivate	1	0.65	show	1	0.64
	relax	1	0.65	speak	1	0.64
	report	1	0.65	stay	1	0.64
	reserve	1	0.65	struggle	1	0.64
	revise	1	0.65	take	1	0.64
	show	1	0.65	understand	1	0.64
	sound	1	0.65	win	1	0.64
	want	1	0.65			
Total	38	153	100.00	37	156	100.00

Table 120

Use of MRW Verbs in Learner Oral Discourse (Control vs. Experimental Groups): Post-Test

MRW verbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	have	20	17.24	have	43	24.16
	spend	15	12.93	go	17	9.55
	take	11	9.48	live	17	9.55
	grow	8	6.90	spend	16	8.99
	live	7	6.03	grow	8	4.49
	go	6	5.17	speak	7	3.93
	love	6	5.17	love	6	3.37
	hang	5	4.31	know	5	2.81
	disconnect	3	2.59	relax	5	2.81
	make	3	2.59	get	4	2.25
	see	3	2.59	help	4	2.25
	care	2	1.72	take	4	2.25
	damage	2	1.72	think	4	2.25
	do	2	1.72	feel	3	1.69
	feel	2	1.72	keep	3	1.69
	get	2	1.72	make	3	1.69
	keep	2	1.72	miss	3	1.69
	realise	2	1.72	disconnect	2	1.12
	arrive	1	0.86	do	2	1.12
	concentrate	1	0.86	see	2	1.12
	connect	1	0.86	walk	2	1.12
	consider	1	0.86	cheer	1	0.56
	experience	1	0.86	come	1	0.56
	focus	1	0.86	concern	1	0.56
	give	1	0.86	connect	1	0.56
	know	1	0.86	consider	1	0.56
	open	1	0.86	disappear	1	0.56
	pay	1	0.86	experience	1	0.56
	relate	1	0.86	face	1	0.56
	surround	1	0.86	find	1	0.56
	think	1	0.86	hang	1	0.56
	treat	1	0.86	hit	1	0.56
	understand	1	0.86	motivate	1	0.56
				produce	1	0.56
				put	1	0.56
				realise	1	0.56
				stay	1	0.56
				value	1	0.56
				watch	1	0.56
Total	33	116	100.00	39	178	100.00

Table 121

Use of MRW Verbs in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

MRW verbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	have	95	51.08	have	91	42.13
	take	12	6.45	take	18	8.33
	live	7	3.76	speak	11	5.09
	mention	7	3.76	get	9	4.17
	consider	5	2.69	go	9	4.17
	go	5	2.69	see	9	4.17
	accept	4	2.15	make	8	3.70
	allow	3	1.61	give	6	2.78
	forget	3	1.61	live	6	2.78
	get	3	1.61	point	4	1.85
	help	3	1.61	say	4	1.85
	spend	3	1.61	understand	3	1.39
	expose	2	1.08	allow	2	0.93
	finish	2	1.08	follow	2	0.93
	know	2	1.08	help	2	0.93
	look	2	1.08	lead	2	0.93
	mean	2	1.08	mention	2	0.93
	open	2	1.08	move	2	0.93
	pass	2	1.08	advance	1	0.46
	solve	2	1.08	apply	1	0.46
	advance	1	0.54	bear	1	0.46
	bring	1	0.54	carry	1	0.46
	create	1	0.54	concentrate	1	0.46
	do	1	0.54	consider	1	0.46
	end	1	0.54	ease	1	0.46
	feel	1	0.54	expand	1	0.46
	fight	1	0.54	find	1	0.46
	form	1	0.54	finish	1	0.46
	grow	1	0.54	forget	1	0.46
	join	1	0.54	hold	1	0.46
	lead	1	0.54	know	1	0.46
	leave	1	0.54	look	1	0.46
	make	1	0.54	mean	1	0.46
	move	1	0.54	open	1	0.46
	pay	1	0.54	prepare	1	0.46
	prepare	1	0.54	provide	1	0.46
	provide	1	0.54	put	1	0.46
	reckon	1	0.54	surf	1	0.46
	remember	1	0.54	talk	1	0.46
	suffer	1	0.54	turn	1	0.46
				update	1	0.46
				upgrade	1	0.46
				value	1	0.46
				win	1	0.46
Total	40	186	100.00	44	216	100.00

Table 122

Use of MRW Verbs in Learner Written Discourse (Control vs. Experimental Groups): Post-Test

MRW verbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	have	26	16.77	have	30	14.63
	mention	19	12.26	go	13	6.34
	take	19	12.26	take	10	4.88
	share	13	8.39	make	9	4.39
	consider	12	7.74	say	9	4.39
	make	5	3.23	speak	9	4.39
	think	5	3.23	hit	8	3.90
	give	4	2.58	see	7	3.41
	go	4	2.58	get	6	2.93
	replace	4	2.58	help	6	2.93
	do	3	1.94	come	5	2.44
	get	3	1.94	follow	5	2.44
	accept	2	1.29	play	5	2.44
	conclude	2	1.29	point	4	1.95
	focus	2	1.29	adapt	3	1.46
	relax	2	1.29	consider	3	1.46
	see	2	1.29	do	3	1.46
	aim	1	0.65	face	3	1.46
	create	1	0.65	give	3	1.46
	deal	1	0.65	replace	3	1.46
	experiment	1	0.65	share	3	1.46
	feel	1	0.65	walk	3	1.46
	finish	1	0.65	deal	2	0.98
	forget	1	0.65	feel	2	0.98
	help	1	0.65	focus	2	0.98
	indicate	1	0.65	look	2	0.98
	live	1	0.65	mean	2	0.98
	look	1	0.65	mention	2	0.98
	lose	1	0.65	open	2	0.98
	pay	1	0.65	progress	2	0.98
	play	1	0.65	stay	2	0.98
	progress	1	0.65	suffer	2	0.98
	reach	1	0.65	understand	2	0.98
	realise	1	0.65	accept	1	0.49
	receive	1	0.65	advance	1	0.49
	reduce	1	0.65	attract	1	0.49
	remember	1	0.65	bear	1	0.49
	send	1	0.65	bring	1	0.49
	settle	1	0.65	create	1	0.49
	show	1	0.65	depend	1	0.49
	spend	1	0.65	digitalise	1	0.49
	suffer	1	0.65	evolve	1	0.49
	treat	1	0.65	expand	1	0.49
	understand	1	0.65	expect	1	0.49
	want	1	0.65	hold	1	0.49
				insert	1	0.49
				judge	1	0.49
				keep	1	0.49
				lift	1	0.49
				live	1	0.49
				lose	1	0.49
				move	1	0.49
				pass	1	0.49

				pay	1	0.49
				plan	1	0.49
				prepare	1	0.49
				realise	1	0.49
				reduce	1	0.49
				result	1	0.49
				save	1	0.49
				send	1	0.49
				substitute	1	0.49
				support	1	0.49
				tend	1	0.49
				throw	1	0.49
				wear	1	0.49
Total	45	155	100.00	66	205	100.00

Appendix S: MRW Adjectives in Learner Discourse

Table 123

Use of MRW Adjectives in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

MRW adjectives	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	free	10	34.48	great	5	19.23
	true	3	10.34	true	4	15.38
	long	2	6.90	free	2	7.69
	prepared	2	6.90	little	2	7.69
	short	2	6.90	responsible	2	7.69
	angry	1	3.45	angry	1	3.85
	crazy	1	3.45	big	1	3.85
	hard	1	3.45	comfortable	1	3.85
	helpful	1	3.45	dependent	1	3.85
	high	1	3.45	difficult	1	3.85
	interactive	1	3.45	fair	1	3.85
	little	1	3.45	independent	1	3.85
	next	1	3.45	insecure	1	3.85
	social	1	3.45	negative	1	3.85
	special	1	3.45	smart	1	3.85
				typical	1	3.85
Total	15	29	100.00	16	26	100.00

Table 124

Use of MRW Adjectives in Learner Oral Discourse (Control vs. Experimental Groups): Post-Test

MRW adjectives	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	great	4	25.00	fit	6	18.75
	social	3	18.75	great	4	12.50
	free	2	12.5	big	3	9.38
	true	2	12.5	independent	3	9.38
	concerned	1	6.25	down	2	6.25
	fit	1	6.25	free	2	6.25
	full	1	6.25	new	2	6.25
	independent	1	6.25	true	2	6.25
	new	1	6.25	able	1	3.13
				isolated	1	3.13
				long	1	3.13
				low	1	3.13
				natural	1	3.13
				real	1	3.13
				social	1	3.13
				ugly	1	3.13
Total	9	16	100.00	16	32	100.00

Table 125

Use of MRW Adjectives in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

MRW adjectives	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	free	5	13.51	free	7	14.89
	huge	3	8.11	great	5	10.64
	new	3	8.11	new	5	10.64
	able	2	5.41	true	4	8.51
	big	2	5.41	able	2	4.26
	hard	2	5.41	clear	2	4.26
	high	2	5.41	high	2	4.26
	public	2	5.41	huge	2	4.26
	small	2	5.41	modern	2	4.26
	true	2	5.41	old	2	4.26
	comfortable	1	2.70	strong	2	4.26
	digital	1	2.70	aware	1	2.13
	economical	1	2.70	difficult	1	2.13
	fabulous	1	2.70	equal	1	2.13
	fantastic	1	2.70	fair	1	2.13
	great	1	2.70	further	1	2.13
	hereditary	1	2.70	hard	1	2.13
	low	1	2.70	inclusive	1	2.13
	perfect	1	2.70	incredible	1	2.13
	popular	1	2.70	mature	1	2.13
	social	1	2.70	outstanding	1	2.13
	weak	1	2.70	perfect	1	2.13
				social	1	2.13
Total	22	37	100.00	23	47	100.00

Table 126

Use of MRW Adjectives in Learner Written Discourse (Control vs. Experimental Groups): Post-Test

MRW adjectives	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	big	6	15.79	big	6	10.91
	clear	6	15.79	digital	5	9.09
	digital	3	7.89	hard	5	9.09
	near	3	7.89	clear	4	7.27
	prepared	3	7.89	safe	3	5.45
	large	2	5.26	true	3	5.45
	comfortable	1	2.63	comfortable	2	3.64
	concerned	1	2.63	easy	2	3.64
	dangerous	1	2.63	full	2	3.64
	distant	1	2.63	great	2	3.64
	essential	1	2.63	long	2	3.64
	far	1	2.63	negative	2	3.64
	general	1	2.63	new	2	3.64
	huge	1	2.63	perfect	2	3.64
	incredible	1	2.63	able	1	1.82
	interactive	1	2.63	concerned	1	1.82
	long	1	2.63	deep	1	1.82
	new	1	2.63	difficult	1	1.82
	traditional	1	2.63	helpful	1	1.82
	true	1	2.63	huge	1	1.82
	untouched	1	2.63	large	1	1.82
				modern	1	1.82
				near	1	1.82
				rushed	1	1.82
				short	1	1.82
				strong	1	1.82
				whole	1	1.82
Total	21	38	100.00	27	55	100.00

Appendix T: MRW Adverbs in Learner Discourse

Table 127

Use of MRW Adverbs in Learner Oral Discourse (Control vs. Experimental Groups): Pre-Test

MRW adverbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	alone	2	22.22	hard	8	36.36
	back	1	11.11	really	6	27.27
	hard	1	11.11	completely	3	13.64
	incredibly	1	11.11	before	2	9.09
	out	1	11.11	free	1	4.55
	there	1	11.11	off	1	4.55
	together	1	11.11	up	1	4.55
	up	1	11.11			
Total	8	9	100.00	7	22	100.00

Table 128

Use of MRW Adverbs in Learner Oral Discourse (Control vs. Experimental Groups): Post-Test

MRW adverbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	out	7	31.82	out	17	36.17
	up	4	18.18	up	8	17.02
	outside	3	13.64	completely	6	12.77
	really	3	13.64	really	5	10.64
	together	2	9.09	together	5	10.64
	before	1	4.55	about	1	2.13
	completely	1	4.55	around	1	2.13
	simply	1	4.55	before	1	2.13
				outside	1	2.13
				simply	1	2.13
				there	1	2.13
Total	8	22	100.00	11	47	100.00

Table 129

Use of MRW Adverbs in Learner Written Discourse (Control vs. Experimental Groups): Pre-Test

MRW adverbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	before	7	29.17	before	7	25.00
	up	6	25.00	generally	6	21.43
	generally	4	16.67	out	4	14.29
	really	3	12.50	really	2	7.14
	alone	1	4.17	away	1	3.57
	away	1	4.17	clearly	1	3.57
	naturally	1	4.17	completely	1	3.57
	quickly	1	4.17	fast	1	3.57
				hard	1	3.57
				overall	1	3.57
				quickly	1	3.57
				there	1	3.57
				up	1	3.57
Total	8	24	100.00	13	28	100.00

Table 130

Use of MRW Adverbs in Learner Written Discourse (Control vs. Experimental Groups): Post-Test

MRW adverbs	Control group			Experimental group		
	Types	AF	RF (%)	Types	AF	RF (%)
	up	8	38.10	generally	5	17.86
	really	3	14.29	out	5	17.86
	before	2	9.52	up	4	14.29
	completely	2	9.52	really	3	10.71
	generally	2	9.52	before	2	7.14
	along	1	4.76	completely	2	7.14
	down	1	4.76	on	2	7.14
	fast	1	4.76	far	1	3.57
	soon	1	4.76	in	1	3.57
				low	1	3.57
				overall	1	3.57
				there	1	3.57
Total	9	21	100.00	12	28	100.00

Appendix U: Results of Research Question 2 (RQ2)

Table 131

Experimental Group's Overall Performance at B2 Level: Pre-Test

B2 groups	Participants	*Reading*	*Use of English*	*Writing*	*Listening*	*Speaking*	Overall Score	Grade
GR0101	ST0101	160	122	153	72	163	143.00	B1 level
	ST0102	147	132	53	130	162	136.00	Not reported
	ST0103	147	137	159	129	170	149.00	B1 level
	ST0104	64	77	141	123	162	132.00	Not reported
	ST0105	185	152	171	163	184	170.00	Grade C
	ST0106	130	77	139	143	160	139.00	Not reported
	ST0107	190	167	180	169	184	181.00	Grade A
	ST0108	190	190	167	190	173	185.00	Grade A
	ST0109	153	48	151	151	163	144.00	B1 level
	ST0110	83	48	145	72	157	129.00	Not reported
GR0102	ST0201	123	37	53	129	177	129.00	Not reported
	ST0202	168	141	159	151	163	157.00	B1 level
	ST0203	156	77	153	72	177	145.00	B1 level
	ST0204	137	56	153	123	169	140.00	B1 level
	ST0205	144	141	141	129	163	144.00	B1 level
	ST0206	140	77	151	80	157	137.00	Not reported
	ST0207	77	48	132	123	163	128.00	Not reported
	ST0208	130	132	152	137	173	147.00	B1 level
	ST0209	163	128	159	147	156	151.00	B1 level
	ST0210	128	56	141	35	151	125.00	Not reported
Total		140.75	102.15	142.65	123.40	166.35	145.55	B1 level

Table 132

Control Group's Overall Performance at B2 Level: Pre-Test

B2 groups	Participants	*Reading*	*Use of English*	*Writing*	*Listening*	*Speaking*	Overall Score	Grade
GR0203	ST0301	172	122	145	151	160	152.00	B1 level
	ST0302	130	48	147	123	162	134.00	Not reported
	ST0303	162	147	144	123	164	149.00	B1 level
	ST0304	140	132	160	160	167	153.00	B1 level
	ST0305	166	48	148	129	160	141.00	B1 level
	ST0306	140	37	152	130	160	137.00	Not reported
	ST0307	156	132	160	123	163	148.00	B1 level
	ST0308	166	77	153	129	164	147.00	B1 level
	ST0309	128	77	147	137	162	139.00	Not reported
	ST0310	153	122	153	129	169	147.00	B1 level
GR0204	ST0401	130	48	152	80	162	130.00	Not reported
	ST0402	179	147	160	72	176	157.00	B1 level
	ST0403	181	137	153	169	176	164.00	Grade C
	ST0404	172	141	147	169	160	160.00	Grade C
	ST0405	168	122	152	166	160	155.00	B1 level
	ST0406	156	122	148	129	169	145.00	B1 level
	ST0407	130	37	126	123	160	128.00	Not reported
	ST0408	172	147	130	129	151	148.00	B1 level
	ST0409	147	48	140	61	157	132.00	Not reported
	ST0410	163	67	144	151	167	149.00	B1 level
Total		155.55	97.90	148.05	129.15	163.45	145.75	B1 level

Table 133

Experimental Group's Speaking Performance at B2 Level: Pre-Test

B2 groups	Participants	Grammar & vocab.	Discourse mgmt.	Pron.	Interactive comm.	Global mark	Overall Score	Grade
GR0101	ST0101	3.00	3.00	3.50	3.50	3.00	163.00	Grade C
	ST0102	3.00	3.50	3.00	3.00	3.00	162.00	Grade C
	ST0103	3.50	4.00	3.50	3.50	3.50	170.00	Grade C
	ST0104	3.00	3.00	3.50	3.00	3.00	162.00	Grade C
	ST0105	4.00	4.50	4.50	4.50	4.00	184.00	Grade A
	ST0106	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0107	5.00	4.00	5.00	3.50	4.00	184.00	Grade A
	ST0108	4.00	4.00	3.50	4.00	3.50	173.00	Grade B
	ST0109	3.00	3.00	3.50	3.50	3.00	163.00	Grade C
	ST0110	2.50	3.00	3.00	3.00	3.00	157.00	B1 level
GR0102	ST0201	3.50	4.00	4.00	4.00	4.00	177.00	Grade B
	ST0202	3.00	3.50	3.00	3.50	3.00	163.00	Grade C
	ST0203	3.50	4.00	4.00	4.00	4.00	177.00	Grade B
	ST0204	3.00	3.50	3.50	4.00	3.50	169.00	Grade C
	ST0205	3.00	3.00	3.50	3.50	3.00	163.00	Grade C
	ST0206	2.00	3.50	3.00	3.00	3.00	157.00	B1 level
	ST0207	3.00	3.50	3.00	3.50	3.00	163.00	Grade C
	ST0208	3.50	4.00	3.50	3.50	4.00	173.00	Grade B
	ST0209	2.00	3.50	2.50	3.00	3.00	156.00	B1 level
	ST0210	2.00	3.50	2.50	3.00	2.50	151.00	B1 level
Total		3.13	3.55	3.43	3.48	3.30	166.35	Grade C

Table 134

Control Group's Speaking Performance at B2 Level: Pre-Test

B2 groups	Participants	Grammar & vocab.	Discourse mgmt.	Pron.	Interactive comm.	Global mark	Overall Score	Grade
GR0203	ST0301	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0302	3.00	3.00	3.50	3.00	3.00	162.00	Grade C
	ST0303	3.00	3.50	3.50	3.50	3.00	164.00	Grade C
	ST0304	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
	ST0305	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0306	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0307	3.50	3.00	3.50	3.00	3.00	163.00	Grade C
	ST0308	3.00	3.50	3.50	3.50	3.00	164.00	Grade C
	ST0309	3.00	3.50	3.00	3.00	3.00	162.00	Grade C
	ST0310	3.50	3.50	3.50	3.50	3.50	169.00	Grade C
GR0204	ST0401	3.00	3.00	3.50	3.00	3.00	162.00	Grade C
	ST0402	3.50	4.00	3.50	4.00	4.00	176.00	Grade B
	ST0403	3.50	4.00	3.50	4.00	4.00	176.00	Grade B
	ST0404	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0405	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0406	3.50	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0407	2.50	3.00	3.50	3.00	3.00	160.00	Grade C
	ST0408	3.00	3.00	2.50	2.50	2.50	151.00	B1 level
	ST0409	2.50	3.00	3.00	3.00	3.00	157.00	B1 level
	ST0410	3.50	3.00	3.50	3.50	3.50	167.00	Grade C
Total		3.10	3.25	3.28	3.23	3.18	163.45	Grade C

Table 135

Experimental Group's Writing Performance at B2 Level: Pre-Test

B2 groups	Participants	Content	Comm. achv.	Organisation	Language	Overall Score	Grade
GR0101	ST0101	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0102	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0103	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0104	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0105	3.50	3.00	3.50	3.00	164.00	Grade C
	ST0106	2.00	2.00	3.00	2.00	140.00	B1 level
	ST0107	4.00	4.00	4.00	4.00	180.00	Grade A
	ST0108	2.50	2.00	3.50	3.00	153.00	B1 level
	ST0109	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0110	2.00	2.50	3.00	2.00	144.00	B1 level
GR0102	ST0201	1.00	1.00	1.00	1.00	53.00	Not reported
	ST0202	2.50	2.50	3.00	2.00	147.00	B1 level
	ST0203	2.50	2.00	3.00	2.00	144.00	B1 level
	ST0204	2.50	2.50	3.00	2.00	147.00	B1 level
	ST0205	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0206	2.00	2.00	3.00	2.50	144.00	B1 level
	ST0207	2.00	2.00	2.50	1.50	130.00	Not reported
	ST0208	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0209	2.50	2.50	2.50	2.50	147.00	B1 level
	ST0210	2.00	2.00	2.00	2.00	130.00	Not reported
Total		2.25	2.20	2.68	2.18	138.40	Not reported

Table 136

Control Group's Writing Performance at B2 Level: Pre-Test

B2 groups	Participants	Content	Comm. achv.	Organisation	Language	Overall Score	Grade
GR0203	ST0301	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0302	2.00	2.00	2.50	1.50	130.00	Not reported
	ST0303	2.00	2.00	2.50	2.50	140.00	B1 level
	ST0304	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0305	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0306	2.00	2.00	2.50	1.50	130.00	Not reported
	ST0307	2.50	2.50	2.50	2.50	147.00	B1 level
	ST0308	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0309	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0310	2.00	2.00	2.50	2.00	137.00	Not reported
GR0204	ST0401	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0402	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0403	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0404	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0405	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0406	2.00	2.00	2.50	2.00	137.00	Not reported
	ST0407	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0408	2.00	2.00	2.00	2.00	130.00	Not reported
	ST0409	2.50	3.00	3.00	2.00	151.00	B1 level
	ST0410	2.50	2.50	2.00	2.50	144.00	B1 level
Total		2.13	2.15	2.33	2.08	136.35	Not reported

Table 137

Experimental Group's Overall Performance at B2 Level: Post-Test

B2 groups	Participants	*Reading*	*Use of English*	*Writing*	*Listening*	*Speaking*	Overall Score	Grade
GR0101	ST0101	147	128	170	151	176	156.00	B1 level
	ST0102	149	77	161	137	167	148.00	B1 level
	ST0103	174	147	164	151	180	164.00	Grade C
	ST0104	137	147	161	147	176	155.00	B1 level
	ST0105	182	167	171	176	188	177.00	Grade B
	ST0106	156	122	153	123	160	144.00	B1 level
	ST0107	190	190	174	185	186	186.00	Grade A
	ST0108	188	179	190	190	188	188.00	Grade A
	ST0109	174	141	160	151	170	162.00	Grade C
	ST0110	130	77	163	80	167	141.00	B1 level
GR0102	ST0201	56	77	151	72	170	130.00	Not reported
	ST0202	162	132	153	147	167	153.00	B1 level
	ST0203	140	67	161	129	170	144.00	B1 level
	ST0204	162	128	153	163	167	156.00	B1 level
	ST0205	140	77	155	176	169	153.00	B1 level
	ST0206	130	67	165	129	167	143.00	B1 level
	ST0207	128	19	148	80	163	129.00	Not reported
	ST0208	149	141	163	137	170	155.00	B1 level
	ST0209	160	137	174	176	167	164.00	Grade C
	ST0210	37	128	157	123	169	136.00	Not reported
Total		144.55	117.40	162.35	141.15	171.85	154.20	B1 level

Table 138

Control Group's Overall Performance at B2 Level: Post-Test

B2 groups	Participants	*Reading*	*Use of English*	*Writing*	*Listening*	*Speaking*	Overall Score	Grade
GR0203	ST0301	181	141	170	166	157	164.00	Grade C
	ST0302	77	137	163	80	169	143.00	B1 level
	ST0303	160	156	167	72	164	153.00	B1 level
	ST0304	149	137	170	147	164	155.00	B1 level
	ST0305	153	152	167	137	160	155.00	B1 level
	ST0306	153	156	167	137	157	155.00	B1 level
	ST0307	162	141	173	147	167	160.00	Grade C
	ST0308	140	137	169	137	170	153.00	B1 level
	ST0309	69	67	163	129	164	137.00	Not reported
	ST0310	144	141	171	151	167	157.00	B1 level
GR0204	ST0401	162	122	183	137	166	157.00	B1 level
	ST0402	185	137	185	156	167	167.00	Grade C
	ST0403	181	141	183	180	177	172.00	Grade C
	ST0404	170	152	180	156	176	167.00	Grade C
	ST0405	163	128	180	163	166	162.00	Grade C
	ST0406	179	147	177	147	171	165.00	Grade C
	ST0407	56	147	182	72	162	144.00	B1 level
	ST0408	130	163	174	160	160	167.00	Grade C
	ST0409	137	56	170	156	160	147.00	B1 level
	ST0410	170	132	164	129	162	153.00	B1 level
Total		146.05	134.05	172.90	137.95	165.30	156.65	B1 level

Table 139

Experimental Group's Speaking Performance at B2 Level: Post-Test

B2 groups	Participants	Grammar & vocab.	Discourse mgmt.	Pron.	Interactive comm.	Global mark	Overall Score	Grade
GR0101	ST0101	3.50	4.00	3.50	4.00	4.00	176.00	Grade B
	ST0102	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
	ST0103	4.00	4.00	4.00	4.00	4.00	180.00	Grade A
	ST0104	3.00	4.00	4.00	4.00	4.00	176.00	Grade B
	ST0105	4.50	4.50	4.50	4.50	4.50	188.00	Grade A
	ST0106	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0107	4.00	4.50	5.00	4.00	4.50	186.00	Grade A
	ST0108	4.50	4.50	4.50	4.50	4.50	188.00	Grade A
	ST0109	3.50	3.50	3.50	4.00	3.50	170.00	Grade C
	ST0110	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
GR0102	ST0201	3.50	3.50	3.50	4.00	3.50	170.00	Grade C
	ST0202	3.50	3.50	3.50	3.00	3.50	167.00	Grade C
	ST0203	3.50	3.50	3.50	4.00	3.50	170.00	Grade C
	ST0204	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
	ST0205	3.50	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0206	3.50	3.50	3.00	3.50	3.50	167.00	Grade C
	ST0207	2.50	3.00	3.00	3.50	3.50	163.00	Grade C
	ST0208	3.50	4.00	3.50	3.50	3.50	170.00	Grade C
	ST0209	3.50	3.50	3.00	3.50	3.50	167.00	Grade C
	ST0210	3.00	4.00	3.50	3.50	3.50	169.00	Grade C
Total		3.45	3.73	3.63	3.73	3.70	171.85	Grade C

Table 140

Control Group's Speaking Performance at B2 Level: Post-Test

B2 groups	Participants	Grammar & vocab.	Discourse mgmt.	Pron.	Interactive comm.	Global mark	Overall Score	Grade
GR0203	ST0301	2.50	3.00	3.00	3.00	3.00	157.00	B1 level
	ST0302	3.50	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0303	3.00	3.50	3.50	3.50	3.00	164.00	Grade C
	ST0304	3.00	3.00	3.00	3.50	3.50	164.00	Grade C
	ST0305	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0306	2.50	3.00	3.00	3.00	3.00	157.00	B1 level
	ST0307	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
	ST0308	3.50	3.50	4.00	3.50	3.50	170.00	Grade C
	ST0309	3.00	3.00	3.00	3.50	3.50	164.00	Grade C
	ST0310	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
GR0204	ST0401	3.00	3.50	3.50	3.00	3.50	166.00	Grade C
	ST0402	3.00	3.50	3.50	3.50	3.50	167.00	Grade C
	ST0403	3.50	4.00	4.00	4.00	4.00	177.00	Grade B
	ST0404	3.50	3.50	4.00	4.00	4.00	176.00	Grade B
	ST0405	3.00	3.00	3.50	3.50	3.50	166.00	Grade C
	ST0406	3.00	4.00	4.00	4.00	3.50	171.00	Grade C
	ST0407	3.00	3.00	3.00	3.50	3.00	162.00	Grade C
	ST0408	3.00	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0409	2.50	3.00	3.00	3.50	3.00	160.00	Grade C
	ST0410	3.00	3.00	3.00	3.50	3.00	162.00	Grade C
Total		3.03	3.30	3.38	3.45	3.35	165.30	Grade C

Table 141

Experimental Group's Writing Performance at B2 Level: Post-Test

B2 groups	Participants	Content	Comm. achv.	Organisation	Language	Overall Score	Grade
GR0101	ST0101	3.50	3.00	3.50	3.00	164.00	Grade C
	ST0102	3.50	3.00	3.50	3.00	165.00	Grade C
	ST0103	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0104	3.50	3.00	3.50	3.00	164.00	Grade C
	ST0105	3.00	3.00	3.50	3.00	163.00	Grade C
	ST0106	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0107	4.00	4.00	4.50	4.00	182.00	Grade A
	ST0108	5.00	5.00	5.00	5.00	190.00	Grade A
	ST0109	3.00	2.50	3.50	3.00	160.00	Grade C
	ST0110	3.00	2.50	3.00	2.50	153.00	B1 level
GR0102	ST0201	3.00	3.00	3.00	2.50	157.00	B1 level
	ST0202	3.00	3.00	3.00	2.50	157.00	B1 level
	ST0203	3.50	3.00	3.50	3.00	164.00	Grade C
	ST0204	3.00	3.00	3.00	2.50	157.00	B1 level
	ST0205	3.00	3.00	3.00	2.50	157.00	B1 level
	ST0206	3.00	3.00	3.00	2.50	157.00	B1 level
	ST0207	2.50	2.50	2.50	2.50	147.00	B1 level
	ST0208	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0209	4.00	4.00	3.50	3.50	173.00	Grade B
	ST0210	3.00	3.00	3.00	3.00	160.00	Grade C
Total		3.33	3.18	3.38	3.05	163.40	Grade C

Table 142

Control Group's Writing Performance at B2 Level: Post-Test

B2 groups	Participants	Content	Comm. achv.	Organisation	Language	Overall Score	Grade
GR0203	ST0301	3.50	3.50	3.50	3.00	167.00	Grade C
	ST0302	3.50	3.00	3.50	3.00	164.00	Grade C
	ST0303	3.50	3.00	3.00	3.00	163.00	Grade C
	ST0304	3.00	3.00	3.50	3.00	163.00	Grade C
	ST0305	3.50	3.50	3.50	3.00	167.00	Grade C
	ST0306	3.50	3.50	3.50	3.00	167.00	Grade C
	ST0307	3.50	3.50	3.50	3.00	167.00	Grade C
	ST0308	3.00	3.00	3.50	2.50	160.00	Grade C
	ST0309	3.00	3.00	3.00	3.00	160.00	Grade C
	ST0310	3.00	3.00	3.00	2.50	157.00	B1 level
GR0204	ST0401	3.50	3.50	3.50	3.50	169.00	Grade C
	ST0402	4.00	4.00	4.00	3.50	177.00	Grade B
	ST0403	4.00	4.00	4.00	4.00	180.00	Grade A
	ST0404	4.00	3.50	4.00	4.00	177.00	Grade B
	ST0405	3.50	3.50	4.00	3.00	169.00	Grade C
	ST0406	4.00	3.50	3.50	3.00	169.00	Grade C
	ST0407	3.50	3.50	4.00	3.50	171.00	Grade C
	ST0408	3.50	3.50	3.50	3.00	167.00	Grade C
	ST0409	4.00	4.00	4.00	3.00	173.00	Grade B
	ST0410	3.00	3.00	3.50	3.00	163.00	Grade C
Total		3.50	3.40	3.58	3.13	167.50	Grade C

Table 143

Overall Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	143.00	156.00	13.00	0.09	9
	ST0102	136.00	148.00	12.00	0.09	9
	ST0103	149.00	164.00	15.00	0.10	10
	ST0104	132.00	155.00	23.00	0.17	17
	ST0105	170.00	177.00	7.00	0.04	4
	ST0106	139.00	144.00	5.00	0.04	4
	ST0107	181.00	186.00	5.00	0.03	3
	ST0108	185.00	188.00	3.00	0.02	2
	ST0109	144.00	162.00	18.00	0.13	13
	ST0110	129.00	141.00	12.00	0.09	9
GR0102	ST0201	129.00	130.00	1.00	0.01	1
	ST0202	157.00	153.00	-4.00	-0.03	-3
	ST0203	145.00	144.00	-1.00	-0.01	-1
	ST0204	140.00	156.00	16.00	0.11	11
	ST0205	144.00	153.00	9.00	0.06	6
	ST0206	137.00	143.00	6.00	0.04	4
	ST0207	128.00	129.00	1.00	0.01	1
	ST0208	147.00	155.00	8.00	0.05	5
	ST0209	151.00	164.00	13.00	0.09	9
	ST0210	125.00	136.00	11.00	0.09	9
Total		145.55	154.20	8.65	0.06	6

Table 144

Overall Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	152.00	164.00	12.00	0.08	8
	ST0302	134.00	143.00	9.00	0.07	7
	ST0303	149.00	153.00	4.00	0.03	3
	ST0304	153.00	155.00	2.00	0.01	1
	ST0305	141.00	155.00	14.00	0.10	10
	ST0306	137.00	155.00	18.00	0.13	13
	ST0307	148.00	160.00	12.00	0.08	8
	ST0308	147.00	153.00	6.00	0.04	4
	ST0309	139.00	137.00	-2.00	-0.01	-1
	ST0310	147.00	157.00	10.00	0.07	7
GR0204	ST0401	130.00	157.00	27.00	0.21	21
	ST0402	157.00	167.00	10.00	0.06	6
	ST0403	164.00	172.00	8.00	0.05	5
	ST0404	160.00	167.00	7.00	0.04	4
	ST0405	155.00	162.00	7.00	0.05	5
	ST0406	145.00	165.00	20.00	0.14	14
	ST0407	128.00	144.00	16.00	0.13	13
	ST0408	148.00	167.00	19.00	0.13	13
	ST0409	132.00	147.00	15.00	0.11	11
	ST0410	149.00	153.00	4.00	0.03	3
Total		145.75	156.65	10.90	0.08	8

Table 145

Speaking Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	163.00	176.00	13.00	0.08	8
	ST0102	162.00	167.00	5.00	0.03	3
	ST0103	170.00	180.00	10.00	0.06	6
	ST0104	162.00	176.00	14.00	0.09	9
	ST0105	184.00	188.00	4.00	0.02	2
	ST0106	160.00	160.00	0.00	0.00	0
	ST0107	184.00	186.00	2.00	0.01	1
	ST0108	173.00	188.00	15.00	0.09	9
	ST0109	163.00	170.00	7.00	0.04	4
	ST0110	157.00	167.00	10.00	0.06	6
GR0102	ST0201	177.00	170.00	-7.00	-0.04	-4
	ST0202	163.00	167.00	4.00	0.02	2
	ST0203	177.00	170.00	-7.00	-0.04	-4
	ST0204	169.00	167.00	-2.00	-0.01	-1
	ST0205	163.00	169.00	6.00	0.04	4
	ST0206	157.00	167.00	10.00	0.06	6
	ST0207	163.00	163.00	0.00	0.00	0
	ST0208	173.00	170.00	-3.00	-0.02	-2
	ST0209	156.00	167.00	11.00	0.07	7
	ST0210	151.00	169.00	18.00	0.12	12
Total		166.35	171.85	5.50	0.03	3

Table 146

Speaking Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	160.00	157.00	-3.00	-0.02	-2
	ST0302	162.00	169.00	7.00	0.04	4
	ST0303	164.00	164.00	0.00	0.00	0
	ST0304	167.00	164.00	-3.00	-0.02	-2
	ST0305	160.00	160.00	0.00	0.00	0
	ST0306	160.00	157.00	-3.00	-0.02	-2
	ST0307	163.00	167.00	4.00	0.02	2
	ST0308	164.00	170.00	6.00	0.04	4
	ST0309	162.00	164.00	2.00	0.01	1
	ST0310	169.00	167.00	-2.00	-0.01	-1
GR0204	ST0401	162.00	166.00	4.00	0.02	2
	ST0402	176.00	167.00	-9.00	-0.05	-5
	ST0403	176.00	177.00	1.00	0.01	1
	ST0404	160.00	176.00	16.00	0.10	10
	ST0405	160.00	166.00	6.00	0.04	4
	ST0406	169.00	171.00	2.00	0.01	1
	ST0407	160.00	162.00	2.00	0.01	1
	ST0408	151.00	160.00	9.00	0.06	6
	ST0409	157.00	160.00	3.00	0.02	2
	ST0410	167.00	162.00	-5.00	-0.03	-3
Total		163.45	165.30	1.85	0.01	1

Table 147

"Grammar and Vocabulary" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	3.00	3.50	0.50	0.17	17
	ST0102	3.00	3.00	0.00	0.00	0
	ST0103	3.50	4.00	0.50	0.14	14
	ST0104	3.00	3.00	0.00	0.00	0
	ST0105	4.00	4.50	0.50	0.13	13
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	5.00	4.00	-1.00	-0.20	-20
	ST0108	4.00	4.50	0.50	0.13	13
	ST0109	3.00	3.50	0.50	0.17	17
	ST0110	2.50	3.00	0.50	0.20	20
GR0102	ST0201	3.50	3.50	0.00	0.00	0
	ST0202	3.00	3.50	0.50	0.17	17
	ST0203	3.50	3.50	0.00	0.00	0
	ST0204	3.00	3.00	0.00	0.00	0
	ST0205	3.00	3.50	0.50	0.17	17
	ST0206	2.00	3.50	1.50	0.75	75
	ST0207	3.00	2.50	-0.50	-0.17	-17
	ST0208	3.50	3.50	0.00	0.00	0%
	ST0209	2.00	3.50	1.50	0.75	7
	ST0210	2.00	3.00	1.00	0.50	50
Total		3.13	3.45	0.33	0.14	14

Table 148

"Grammar and Vocabulary" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	3.00	2.50	-0.50	-0.17	-17
	ST0302	3.00	3.50	0.50	0.17	17
	ST0303	3.00	3.00	0.00	0.00	0
	ST0304	3.00	3.00	0.00	0.00	0
	ST0305	3.00	3.00	0.00	0.00	0
	ST0306	3.00	2.50	-0.50	-0.17	-17
	ST0307	3.50	3.00	-0.50	-0.14	-14
	ST0308	3.00	3.50	0.50	0.17	17
	ST0309	3.00	3.00	0.00	0.00	0
	ST0310	3.50	3.00	-0.50	-0.14	-14
GR0204	ST0401	3.00	3.00	0.00	0.00	0
	ST0402	3.50	3.00	-0.50	-0.14	-14
	ST0403	3.50	3.50	0.00	0.00	0
	ST0404	3.00	3.50	0.50	0.17	17
	ST0405	3.00	3.00	0.00	0.00	0
	ST0406	3.50	3.00	-0.50	-0.14	-14
	ST0407	2.50	3.00	0.50	0.20	20
	ST0408	3.00	3.00	0.00	0.00	0
	ST0409	2.50	2.50	0.00	0.00	0
	ST0410	3.50	3.00	-0.50	-0.14	-14
Total		3.10	3.03	-0.08	-0.02	-2

Table 149

"Discourse Management" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	3.00	4.00	1.00	0.33	33
	ST0102	3.50	3.50	0.00	0.00	0
	ST0103	4.00	4.00	0.00	0.00	0
	ST0104	3.00	4.00	1.00	0.33	33
	ST0105	4.50	4.50	0.00	0.00	0
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	4.00	4.50	0.50	0.13	13
	ST0108	4.00	4.50	0.50	0.13	13
	ST0109	3.00	3.50	0.50	0.17	17
	ST0110	3.00	3.50	0.50	0.17	17
GR0102	ST0201	4.00	3.50	-0.50	-0.13	-13
	ST0202	3.50	3.50	0.00	0.00	0
	ST0203	4.00	3.50	-0.50	-0.13	-13
	ST0204	3.50	3.50	0.00	0.00	0
	ST0205	3.00	3.50	0.50	0.17	17
	ST0206	3.50	3.50	0.00	0.00	0
	ST0207	3.50	3.00	-0.50	-0.14	-14
	ST0208	4.00	4.00	0.00	0.00	0
	ST0209	3.50	3.50	0.00	0.00	0
	ST0210	3.50	4.00	0.50	0.14	14
Total		3.55	3.73	0.18	0.06	6

Table 150

"Discourse Management" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	3.00	3.00	0.00	0.00	0
	ST0302	3.00	3.50	0.50	0.17	17
	ST0303	3.50	3.50	0.00	0.00	0
	ST0304	3.50	3.00	-0.50	-0.14	-14
	ST0305	3.00	3.00	0.00	0.00	0
	ST0306	3.00	3.00	0.00	0.00	0
	ST0307	3.00	3.50	0.50	0.17	17
	ST0308	3.50	3.50	0.00	0.00	0
	ST0309	3.50	3.00	-0.50	-0.14	-14
	ST0310	3.50	3.50	0.00	0.00	0
GR0204	ST0401	3.00	3.50	0.50	0.17	17
	ST0402	4.00	3.50	-0.50	-0.13	-13
	ST0403	4.00	4.00	0.00	0.00	0
	ST0404	3.00	3.50	0.50	0.17	17
	ST0405	3.00	3.00	0.00	0.00	0
	ST0406	3.50	4.00	0.50	0.14	14
	ST0407	3.00	3.00	0.00	0.00	0
	ST0408	3.00	3.00	0.00	0.00	0
	ST0409	3.00	3.00	0.00	0.00	0
	ST0410	3.00	3.00	0.00	0.00	0
Total		3.25	3.30	0.05	0.02	2

Table 151

"Pronunciation" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	3.50	3.50	0.00	0.00	0
	ST0102	3.00	3.50	0.50	0.17	17
	ST0103	3.50	4.00	0.50	0.14	14
	ST0104	3.50	4.00	0.50	0.14	14
	ST0105	4.50	4.50	0.00	0.00	0
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	5.00	5.00	0.00	0.00	0
	ST0108	3.50	4.50	1.00	0.29	29
	ST0109	3.50	3.50	0.00	0.00	0
	ST0110	3.00	3.50	0.50	0.17	17
GR0102	ST0201	4.00	3.50	-0.50	-0.13	-13
	ST0202	3.00	3.50	0.50	0.17	17
	ST0203	4.00	3.50	-0.50	-0.13	-13
	ST0204	3.50	3.50	0.00	0.00	0
	ST0205	3.50	3.50	0.00	0.00	0
	ST0206	3.00	3.00	0.00	0.00	0
	ST0207	3.00	3.00	0.00	0.00	0
	ST0208	3.50	3.50	0.00	0.00	0
	ST0209	2.50	3.00	0.50	0.20	20
	ST0210	2.50	3.50	1.00	0.40	40
Total		3.43	3.63	0.20	0.07	7

Table 152

"Pronunciation" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	3.00	3.00	0.00	0.00	0
	ST0302	3.50	3.50	0.00	0.00	0
	ST0303	3.50	3.50	0.00	0.00	0
	ST0304	3.50	3.00	-0.50	-0.50	-50
	ST0305	3.00	3.00	0.00	0.00	0
	ST0306	3.00	3.00	0.00	0.00	0
	ST0307	3.50	3.50	0.00	0.00	0
	ST0308	3.50	4.00	0.50	0.50	50
	ST0309	3.00	3.00	0.00	0.00	0
	ST0310	3.50	3.50	0.00	0.00	0
GR0204	ST0401	3.50	3.50	0.00	0.00	0
	ST0402	3.50	3.50	0.00	0.00	0
	ST0403	3.50	4.00	0.50	0.50	50
	ST0404	3.00	4.00	1.00	1.00	100
	ST0405	3.00	3.50	0.50	0.50	50
	ST0406	3.50	4.00	0.50	0.50	50
	ST0407	3.50	3.00	-0.50	-0.50	-50
	ST0408	2.50	3.00	0.50	0.50	50
	ST0409	3.00	3.00	0.00	0.00	0
	ST0410	3.50	3.00	-0.50	-0.50	-50
Total		3.28	3.38	0.10	0.10	10

Table 153

"Interactive Communication" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	3.50	4.00	0.50	0.14	14
	ST0102	3.00	3.50	0.50	0.17	17
	ST0103	3.50	4.00	0.50	0.14	14
	ST0104	3.00	4.00	1.00	0.33	33
	ST0105	4.50	4.50	0.00	0.00	0
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	3.50	4.00	0.50	0.14	14
	ST0108	4.00	4.50	0.50	0.13	13
	ST0109	3.50	4.00	0.50	0.14	14
	ST0110	3.00	3.50	0.50	0.17	17
GR0102	ST0201	4.00	4.00	0.00	0.00	0
	ST0202	3.50	3.00	-0.50	-0.14	-14
	ST0203	4.00	4.00	0.00	0.00	0
	ST0204	4.00	3.50	-0.50	-0.13	-13
	ST0205	3.50	3.50	0.00	0.00	0
	ST0206	3.00	3.50	0.50	0.17	17
	ST0207	3.50	3.50	0.00	0.00	0
	ST0208	3.50	3.50	0.00	0.00	0
	ST0209	3.00	3.50	0.50	0.17	17
	ST0210	3.00	3.50	0.50	0.17	17
Total		3.48	3.73	0.25	0.08	8

Table 154

"Interactive Communication" Performance Variations at B2 (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	3.00	3.00	0.00	0.00	0
	ST0302	3.00	3.50	0.50	0.17	17
	ST0303	3.50	3.50	0.00	0.00	0
	ST0304	3.50	3.50	0.00	0.00	0
	ST0305	3.00	3.00	0.00	0.00	0
	ST0306	3.00	3.00	0.00	0.00	0
	ST0307	3.00	3.50	0.50	0.17	17
	ST0308	3.50	3.50	0.00	0.00	0
	ST0309	3.00	3.50	0.50	0.17	17
	ST0310	3.50	3.50	0.00	0.00	0
GR0204	ST0401	3.00	3.00	0.00	0.00	0
	ST0402	4.00	3.50	-0.50	-0.13	-13
	ST0403	4.00	4.00	0.00	0.00	0
	ST0404	3.00	4.00	1.00	0.33	33
	ST0405	3.00	3.50	0.50	0.17	17
	ST0406	3.50	4.00	0.50	0.14	14
	ST0407	3.00	3.50	0.50	0.17	17
	ST0408	2.50	3.00	0.50	0.20	20
	ST0409	3.00	3.50	0.50	0.17	17
	ST0410	3.50	3.50	0.00	0.00	0
Total		3.23	3.45	0.23	0.08	8

Table 155

"Global Mark" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	3.00	4.00	1.00	0.33	33
	ST0102	3.00	3.50	0.50	0.17	17
	ST0103	3.50	4.00	0.50	0.14	14
	ST0104	3.00	4.00	1.00	0.33	33
	ST0105	4.00	4.50	0.50	0.13	13
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	4.00	4.50	0.50	0.13	13
	ST0108	3.50	4.50	1.00	0.29	29
	ST0109	3.00	3.50	0.50	0.17	17
	ST0110	3.00	3.50	0.50	0.17	17
GR0102	ST0201	4.00	3.50	-0.50	-0.13	-13
	ST0202	3.00	3.50	0.50	0.17	17
	ST0203	4.00	3.50	-0.50	-0.13	-13
	ST0204	3.50	3.50	0.00	0.00	0
	ST0205	3.00	3.50	0.50	0.17	17
	ST0206	3.00	3.50	0.50	0.17	17
	ST0207	3.00	3.50	0.50	0.17	17
	ST0208	4.00	3.50	-0.50	-0.13	-13
	ST0209	3.00	3.50	0.50	0.17	17
	ST0210	2.50	3.50	1.00	0.40	40
Total		3.30	3.70	0.40	0.14	14

Table 156

"Global Mark" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	3.00	3.00	0.00	0.00	0
	ST0302	3.00	3.50	0.50	0.17	17
	ST0303	3.00	3.00	0.00	0.00	0
	ST0304	3.50	3.50	0.00	0.00	0
	ST0305	3.00	3.00	0.00	0.00	0
	ST0306	3.00	3.00	0.00	0.00	0
	ST0307	3.00	3.50	0.50	0.17	17
	ST0308	3.00	3.50	0.50	0.17	17
	ST0309	3.00	3.50	0.50	0.17	17
	ST0310	3.50	3.50	0.00	0.00	0
GR0204	ST0401	3.00	3.50	0.50	0.17	17
	ST0402	4.00	3.50	-0.50	-0.13	-13
	ST0403	4.00	4.00	0.00	0.00	0
	ST0404	3.00	4.00	1.00	0.33	33
	ST0405	3.00	3.50	0.50	0.17	17
	ST0406	3.50	3.50	0.00	0.00	0
	ST0407	3.00	3.00	0.00	0.00	0
	ST0408	2.50	3.00	0.50	0.20	20
	ST0409	3.00	3.00	0.00	0.00	0
	ST0410	3.50	3.00	-0.50	-0.14	-14
Total		3.18	3.35	0.18	0.06	6

Table 157

Writing Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	137.00	164.00	27.00	0.20	20
	ST0102	130.00	165.00	35.00	0.27	27
	ST0103	137.00	169.00	32.00	0.23	23
	ST0104	130.00	164.00	34.00	0.26	26
	ST0105	164.00	163.00	-1.00	-0.01	-1
	ST0106	140.00	160.00	20.00	0.14	14
	ST0107	180.00	182.00	2.00	0.01	1
	ST0108	153.00	190.00	37.00	0.24	24
	ST0109	137.00	160.00	23.00	0.17	17
	ST0110	144.00	153.00	9.00	0.06	6
GR0102	ST0201	53.00	157.00	104.00	1.96	196
	ST0202	147.00	157.00	10.00	0.07	7
	ST0203	144.00	164.00	20.00	0.14	14
	ST0204	147.00	157.00	10.00	0.07	7
	ST0205	137.00	157.00	20.00	0.15	15
	ST0206	144.00	157.00	13.00	0.09	9
	ST0207	130.00	147.00	17.00	0.13	13
	ST0208	137.00	169.00	32.00	0.23	23
	ST0209	147.00	173.00	26.00	0.18	18
	ST0210	130.00	160.00	30.00	0.23	23
Total		138.40	163.40	25.00	0.18	18

Table 158

Writing Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	137.00	167.00	30.00	0.22	22
	ST0302	130.00	164.00	34.00	0.26	26
	ST0303	140.00	163.00	23.00	0.16	16
	ST0304	160.00	163.00	3.00	0.02	2
	ST0305	137.00	167.00	30.00	0.22	22
	ST0306	130.00	167.00	37.00	0.28	28
	ST0307	147.00	167.00	20.00	0.14	14
	ST0308	130.00	160.00	30.00	0.23	23
	ST0309	130.00	160.00	30.00	0.23	23
	ST0310	137.00	157.00	20.00	0.15	15
GR0204	ST0401	130.00	169.00	39.00	0.30	30
	ST0402	137.00	177.00	40.00	0.29	29
	ST0403	130.00	180.00	50.00	0.38	38
	ST0404	130.00	177.00	47.00	0.36	36
	ST0405	130.00	169.00	39.00	0.30	30
	ST0406	137.00	169.00	32.00	0.23	23
	ST0407	130.00	171.00	41.00	0.32	32
	ST0408	130.00	167.00	37.00	0.28	28
	ST0409	151.00	173.00	22.00	0.15	15
	ST0410	144.00	163.00	19.00	0.13	13
Total		136.35	167.50	31.15	0.23	23

Table 159

"Content" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	2.00	3.50	1.50	0.75	75
	ST0102	2.00	3.50	1.50	0.75	75
	ST0103	2.00	3.50	1.50	0.75	75
	ST0104	2.00	3.50	1.50	0.75	75
	ST0105	3.50	3.00	-0.50	-0.14	-14
	ST0106	2.00	3.00	1.00	0.50	50
	ST0107	4.00	4.00	0.00	0.00	0
	ST0108	2.50	5.00	2.50	1.00	100
	ST0109	2.00	3.00	1.00	0.50	50
	ST0110	2.00	3.00	1.00	0.50	50
GR0102	ST0201	1.00	3.00	2.00	2.00	200
	ST0202	2.50	3.00	0.50	0.20	20
	ST0203	2.50	3.50	1.00	0.40	40
	ST0204	2.50	3.00	0.50	0.20	20
	ST0205	2.00	3.00	1.00	0.50	50
	ST0206	2.00	3.00	1.00	0.50	50
	ST0207	2.00	2.50	0.50	0.25	25
	ST0208	2.00	3.50	1.50	0.75	75
	ST0209	2.50	4.00	1.50	0.60	60
	ST0210	2.00	3.00	1.00	0.50	50
Total		2.25	3.33	1.08	0.56	56

Table 160

"Content" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	2.00	3.50	1.50	0.75	75
	ST0302	2.00	3.50	1.50	0.75	75
	ST0303	2.00	3.50	1.50	0.75	75
	ST0304	3.00	3.00	0.00	0.00	0
	ST0305	2.00	3.50	1.50	0.75	75
	ST0306	2.00	3.50	1.50	0.75	75
	ST0307	2.50	3.50	1.00	0.40	40
	ST0308	2.00	3.00	1.00	0.50	50
	ST0309	2.00	3.00	1.00	0.50	50
	ST0310	2.00	3.00	1.00	0.50	50
GR0204	ST0401	2.00	3.50	1.50	0.75	75
	ST0402	2.00	4.00	2.00	1.00	100
	ST0403	2.00	4.00	2.00	1.00	100
	ST0404	2.00	4.00	2.00	1.00	100
	ST0405	2.00	3.50	1.50	0.75	75
	ST0406	2.00	4.00	2.00	1.00	100
	ST0407	2.00	3.50	1.50	0.75	75
	ST0408	2.00	3.50	1.50	0.75	75
	ST0409	2.50	4.00	1.50	0.60	60
	ST0410	2.50	3.00	0.50	0.20	20
Total		2.13	3.50	1.38	0.67	67

Table 161

"Communicative Achievement" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	2.00	3.00	1.00	0.50	50
	ST0102	2.00	3.00	1.00	0.50	50
	ST0103	2.00	3.50	1.50	0.75	75
	ST0104	2.00	3.00	1.00	0.50	50
	ST0105	3.00	3.00	0.00	0.00	0
	ST0106	2.00	3.00	1.00	0.50	50
	ST0107	4.00	4.00	0.00	0.00	0
	ST0108	2.00	5.00	3.00	1.50	150
	ST0109	2.00	2.50	0.50	0.25	25
	ST0110	2.50	2.50	0.00	0.00	0
GR0102	ST0201	1.00	3.00	2.00	2.00	200
	ST0202	2.50	3.00	0.50	0.20	20
	ST0203	2.00	3.00	1.00	0.50	50
	ST0204	2.50	3.00	0.50	0.20	20
	ST0205	2.00	3.00	1.00	0.50	50
	ST0206	2.00	3.00	1.00	0.50	50
	ST0207	2.00	2.50	0.50	0.25	25
	ST0208	2.00	3.50	1.50	0.75	75
	ST0209	2.50	4.00	1.50	0.60	60
	ST0210	2.00	3.00	1.00	0.50	50
Total		2.20	3.18	0.98	0.44	44

Table 162

"Communicative Achievement" Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	2.00	3.50	1.50	0.75	75
	ST0302	2.00	3.00	1.00	0.50	50
	ST0303	2.00	3.00	1.00	0.50	50
	ST0304	3.00	3.00	0.00	0.00	0
	ST0305	2.00	3.50	1.50	0.75	75
	ST0306	2.00	3.50	1.50	0.75	75
	ST0307	2.50	3.50	1.00	0.40	40
	ST0308	2.00	3.00	1.00	0.50	50
	ST0309	2.00	3.00	1.00	0.50	50
	ST0310	2.00	3.00	1.00	0.50	50
GR0204	ST0401	2.00	3.50	1.50	0.75	75
	ST0402	2.00	4.00	2.00	1.00	100
	ST0403	2.00	4.00	2.00	1.00	100
	ST0404	2.00	3.50	1.50	0.75	75
	ST0405	2.00	3.50	1.50	0.75	75
	ST0406	2.00	3.50	1.50	0.75	75
	ST0407	2.00	3.50	1.50	0.75	75
	ST0408	2.00	3.50	1.50	0.75	75
	ST0409	3.00	4.00	1.00	0.33	33
	ST0410	2.50	3.00	0.50	0.20	20
Total		2.15	3.40	1.25	0.58	58

Table 163

"Organisation" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	2.50	3.50	1.00	0.40	40
	ST0102	2.00	3.50	1.50	0.75	75
	ST0103	2.50	3.50	1.00	0.40	40
	ST0104	2.00	3.50	1.50	0.75	75
	ST0105	3.50	3.50	0.00	0.00	0
	ST0106	3.00	3.00	0.00	0.00	0
	ST0107	4.00	4.50	0.50	0.13	13
	ST0108	3.50	5.00	1.50	0.43	43
	ST0109	2.50	3.50	1.00	0.40	40
	ST0110	3.00	3.00	0.00	0.00	0
GR0102	ST0201	1.00	3.00	2.00	2.00	200
	ST0202	3.00	3.00	0.00	0.00	0
	ST0203	3.00	3.50	0.50	0.17	17
	ST0204	3.00	3.00	0.00	0.00	0
	ST0205	2.50	3.00	0.50	0.20	20
	ST0206	3.00	3.00	0.00	0.00	0
	ST0207	2.50	2.50	0.00	0.00	0
	ST0208	2.50	3.50	1.00	0.40	40
	ST0209	2.50	3.50	1.00	0.40	40
	ST0210	2.00	3.00	1.00	0.50	50
Total		2.68	3.38	0.70	0.35	35

Table 164

"Organisation" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	2.50	3.50	1.00	0.40	40
	ST0302	2.50	3.50	1.00	0.40	40
	ST0303	2.50	3.00	0.50	0.20	20
	ST0304	3.00	3.50	0.50	0.17	17
	ST0305	2.50	3.50	1.00	0.40	40
	ST0306	2.50	3.50	1.00	0.40	40
	ST0307	2.50	3.50	1.00	0.40	40
	ST0308	2.00	3.50	1.50	0.75	75
	ST0309	2.00	3.00	1.00	0.50	50
	ST0310	2.50	3.00	0.50	0.20	20
GR0204	ST0401	2.00	3.50	1.50	0.75	75
	ST0402	2.50	4.00	1.50	0.60	60
	ST0403	2.00	4.00	2.00	1.00	100
	ST0404	2.00	4.00	2.00	1.00	100
	ST0405	2.00	4.00	2.00	1.00	100
	ST0406	2.50	3.50	1.00	0.40	40
	ST0407	2.00	4.00	2.00	1.00	100
	ST0408	2.00	3.50	1.50	0.75	75
	ST0409	3.00	4.00	1.00	0.33	33
	ST0410	2.00	3.50	1.50	0.75	75
Total		2.33	3.58	1.25	0.57	57

Table 165

"Language" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Experimental Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0101	ST0101	2.00	3.00	1.00	0.50	50
	ST0102	2.00	3.00	1.00	0.50	50
	ST0103	2.00	3.50	1.50	0.75	75
	ST0104	2.00	3.00	1.00	0.50	50
	ST0105	3.00	3.00	0.00	0.00	0
	ST0106	2.00	3.00	1.00	0.50	50
	ST0107	4.00	4.00	0.00	0.00	0
	ST0108	3.00	5.00	2.00	0.67	67
	ST0109	2.00	3.00	1.00	0.50	50
	ST0110	2.00	2.50	0.50	0.25	25
GR0102	ST0201	1.00	2.50	1.50	1.50	150
	ST0202	2.00	2.50	0.50	0.25	25
	ST0203	2.00	3.00	1.00	0.50	50
	ST0204	2.00	2.50	0.50	0.25	25
	ST0205	2.00	2.50	0.50	0.25	25
	ST0206	2.50	2.50	0.00	0.00	0
	ST0207	1.50	2.50	1.00	0.67	67
	ST0208	2.00	3.50	1.50	0.75	75
	ST0209	2.50	3.50	1.00	0.40	40
	ST0210	2.00	3.00	1.00	0.50	50
Total		2.18	3.05	0.88	0.46	46

Table 166

"Language" Performance Variations at B2 Level (Pre-Test vs. Post-Test): Control Group

B2 groups	Participants	Pre-test score	Post-test score	Evolution rates		
				AI	RI	%
GR0203	ST0301	2.00	3.00	1.00	0.50	50
	ST0302	1.50	3.00	1.50	1.00	100
	ST0303	2.50	3.00	0.50	0.20	20
	ST0304	3.00	3.00	0.00	0.00	0
	ST0305	2.00	3.00	1.00	0.50	50
	ST0306	1.50	3.00	1.50	1.00	100
	ST0307	2.50	3.00	0.50	0.20	20
	ST0308	2.00	2.50	0.50	0.25	25
	ST0309	2.00	3.00	1.00	0.50	50
	ST0310	2.00	2.50	0.50	0.25	25
GR0204	ST0401	2.00	3.50	1.50	0.75	75
	ST0402	2.00	3.50	1.50	0.75	75
	ST0403	2.00	4.00	2.00	1.00	100
	ST0404	2.00	4.00	2.00	1.00	100
	ST0405	2.00	3.00	1.00	0.50	50
	ST0406	2.00	3.00	1.00	0.50	50
	ST0407	2.00	3.50	1.50	0.75	75
	ST0408	2.00	3.00	1.00	0.50	50
	ST0409	2.00	3.00	1.00	0.50	50
	ST0410	2.50	3.00	0.50	0.20	20
Total		2.08	3.13	1.05	0.54	54