

**Investigating CLIL effects on
learners' academic performance
in a Chinese tertiary education programme**

Cibo ZHOU, EdD

**Thesis submitted to the University of Nottingham
for the degree of Doctor of Education**

February, 2022

DECLARATION

I declare that this thesis represents my own work and has not been submitted in any form for another degree at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

A handwritten signature in black ink that reads "David". The signature is written in a cursive style with a large 'D' and a small 'd'.

February, 22nd, 2022

ABSTRACT

Given the advantage of packaging foreign language enrichment measures into content teaching, the Content and Language Integrated Learning (CLIL) approach is regarded by some Chinese researchers as a potential alternative to the current mainstream EFL approach at the tertiary level widely seen as ineffective. However, a majority of the CLIL research in China remains centred on introducing the CLIL concept and evaluating feasibility from theoretical perspectives, while little empirical research has been documented. Thus, the present study investigated the effects of CLIL on learners' academic performance and language proficiency in a Chinese tertiary educational context, attempting to provide "hard evidence" on CLIL achievements in Chinese education.

In this study, based on the elucidation of the umbrella term of CLIL, a critical review of previous empirical studies and methodological instruments, and the theoretical reinterpretation of the 4Cs framework, a two-phase mixed-methods quasi-experiment was designed and conducted with both quantitative and qualitative methods. The quantitative methods, including pre- and post-tests of linguistic and content performance, were administered to scrutinize the statistical relationship between CLIL exposure and learners' attainment in language and content. The key findings were verified through a qualitative method of interviews in the second phase. The triangulated data were expected to provide more reliable and dependable evidence of the effectiveness of CLIL in Chinese educational contexts.

By investigating the temporal changes of learners' academic performance for one semester, the study led to findings that present a compelling, theory-driven explanation for how and why CLIL education impacted Chinese university learners' academic performance, accompanied by the theoretical, empirical and pedagogical implications for policymakers and practitioners to enhance target learners' academic learning.

ACKNOWLEDGEMENTS

Thanks!

First and foremost, I would like to express my gratitude to my supervisor, Prof. Anwei Feng. Many thanks to him for his precious guidance, invaluable support, and inspiring comments. In all stages of the present research, he has been available to encourage me to go ahead and develop the thesis creatively and critically.

Moreover, I'd like to express my thank to Dr. Irina Hawker, for her helpful and valuable advice on the research design and thesis development, which was of great importance and significance in my research.

Finally, I would also like to thank all the students who participated in this study for their enthusiasm and willingness to volunteer and answer all my questions. I also thank all my colleagues for their contributions to this thesis.

Thanks!

TABLE OF CONTENTS

CHAPTER ONE- Introduction	1
1.1 Background	1
1.2 Rationale	3
1.2.1 Positionality	4
1.2.2 Practical perspective	5
1.2.3 Theoretical perspective	9
1.3 Aims	11
1.4 Overview of the thesis structure	11
CHAPTER TWO- Literature Review	13
2.1 Delineating the umbrella term of CLIL	13
2.1.1 Conceptualising CLIL from policymakers' perspective	17
2.1.2 Conceptualising CLIL from practitioners' perspective	20
2.1.3 Conceptualising CLIL from an integrated perspective	23
2.2 Review of CLIL studies	30
2.2.1 Overview of CLIL effect in quantitative studies	31
2.2.2 Deep insights of CLIL attainments in qualitative studies	36
2.2.3 Investigating CLIL effects from the integrated perspective	40
CHAPTER THREE- The 4Cs Framework of CLIL research	46
3.1 The theoretical framework of CLIL	46
3.1.1 Content	47
3.1.2 Communication	50
3.1.3 Cognition	53
3.1.4 Culture	55
3.2 The 4Cs framework of CLIL in the present study	58

3.3 Aims and research questions	60
CHAPTER FOUR- Research Methodology	65
4.1 The research paradigm	65
4.1.1 The ontological and epistemological stance: Pragmatism	66
4.1.2 The research paradigm in the present study	73
4.1.2.1 The ontological and epistemological stance	73
4.1.2.2 Methodology	74
4.2 The research design	77
4.2.1 Samples	78
4.2.2 The selection of mixed-method strategies	79
4.3 Research procedures and methods	84
4.3.1 Phase one: The pre-test-post-test control and experimental group design	85
4.3.2 Phase two: The semi-structured open-ended interviews	88
4.4 Ethical issues	91
4.5 Data collection and analysis	92
4.5.1 Quantitative data from the tests	94
4.5.2 Qualitative data from the interviews	94
CHAPTER FIVE- Data Analysis	97
5.1 Quantitative findings from Phase One	97
5.1.1 Statistic descriptive of quantitative tests	98
5.1.2 Relationships between CLIL learners' language competence and content performance in pre- and post-tests	101
5.2 Qualitative findings from Phase Two	108
5.2.1 Analysis of learners' interviews	109
5.2.2 Teacher's feedback	124

CHAPTER SIX- Discussions, Conclusions and Implications	131
6.1 Discussions	132
6.1.1 Improvements in writing skills	132
6.1.2 The role of language learning in CLIL contexts	133
6.1.3 Translanguaging	135
6.1.4 Cognitive thinking processes in CLIL learning	137
6.1.5 Changes in learners' intercultural understanding	139
6.2 Implications	142
6.3 Limitations	144
REFERENCE	146
APPENDICES	159

LIST OF TABLES

Table 5.1	Descriptive statistics of pre-and post-tests for the experimental group	99
Table 5.2	Descriptive statistics of pre-and post-tests for the control group	100
Table 5.3	Descriptive statistics between the experimental and control group	101
Table 5.4	The results of Pearson Correlation Coefficient of Control Group	102
Table 5.5	The results of Pearson Correlation Coefficient of Experimental group	103
Table 5.6	Background information of interview participants	109

LIST OF FIGURES

Figure 2.1	The continuum of content and language integration	24
Figure 2.2	The integrated grid of CLIL	26
Figure 2.3	Language continua of CLIL	28
Figure 3.1	The 4Cs framework	47
Figure 4.1	Sequential Explanatory Design	82

CHAPTER ONE - Introduction

1.1 Background

The 21st century has witnessed a historical paradigm shift where “the grand rhetoric of global perspective sits alongside postmodernist interpretations of fragmented societies” (Coyle, 2007, p. 543). The worldwide interchange of commodity, knowledge and human resource in the era of globalised society bolsters a growing need in the job market for a huge bilingual/multilingual workforce (Banegas, 2014; Hüttner & Smit, 2014; A. Llinares & McCabe, 2020). This shift in job orientation leads to the change of language patterns on a global scale, a situation depicted by Maurais (2003, p. 13) as a “new linguistic world order”. These changes arouse a great demand on mainstream education to “improve language-learning opportunities and linguistic educational outcomes” (Juan-Garau & Salazar-Noguera, 2015, p. 1). In response, university graduates’ competence in second/foreign language (L2) communication has been accentuated in diverse countries, driving educational policymakers and practitioners to alter outmoded L2 educational approaches and seek appropriate alternatives which allow students to “stand their ground in international contexts” (Dalton-Puffer, 2007).

Among the variegated L2 educational approaches practised in European contexts over the last two decades, the dual-focused Content and Language Integrated Learning (CLIL) approach stands out to attract genuine attention in its birthplace. As an innovative educational approach, CLIL has been considered from its inception by some European policymakers and grass-root actors as “a pragmatic European solution” to globalisation, and evaluated, mainly from an economic viewpoint, as “a highly successful and efficient way of channelling resources towards language acquisition without putting more pressure on an already hefty curriculum” (European Commission, 1995; Eurydice, 2006; Marsh, 2002, p. 11). This ascendancy is termed in a simplified way as “delivering ‘two’ (foreign language and subject content) for the price of ‘one’ (teaching unit)” (Dalton-Puffer, Nikula, & Smit, 2010, p. 284). Given these advantages,

CLIL has gained strong support from key stakeholders like the European Union, L2 teachers, parents, and learners. The approach has had an exponential and massive uptake across diverse European nations and has been gradually embedded in mainstream education from preschool to vocational education as an established teaching approach (Järvinen, 2010; Pérez-Cañado, 2012). A large number of studies has been conducted to verify the effectiveness of CLIL in enhancing learner attainment in European settings. As commented by Deller and Price (2005, p. 29), CLIL is “spreading fast and here to stay”, in both the academic world and classroom praxis (Goris, Denessen, & Verhoeven, 2019; Hemmi & Banegas, 2021; Infante, Benvenuto, & Lastrucci, 2009).

In a similar vein, China is seeking pragmatic solutions to globalisation as well. Since adopting the Reform and Open-up policy in 1978, China goes global at an unprecedented pace. The increasing participation in the international community and global economy entrenches the use of English and other foreign languages in Chinese society. As a result, a consensus seems to be reached throughout the country that the role and status of foreign language education need further and faster development (Bruton, 2015; Feng, 2009; MOE, 2015). Various nationwide attempts have been made to meet the requirements of fast growth in English education. However, despite the national drive for foreign language education, the result of its achievement in Chinese universities does not appear to be satisfactory (Sheng, 2012; Zhou & Zhang, 2014). The mainstream EFL (English as a Foreign Language) approach in Chinese tertiary education has been extensively criticized for its “high time consumption and low effectiveness”, arousing widespread disillusion and dissatisfaction among key stakeholders and practitioners (Cai, 2010, p. 306).

Given the rapid and massive practices in European contexts and the advantage of packaging foreign language enrichment measure into content teaching, CLIL is viewed by some Chinese researchers as a potential solution to globalisation and an alternative

to the current mainstream EFL approach (Chang, 2006; Hu, Li, & Lei, 2014; H. Liu, 2019; R. N. Wei & Feng, 2015). However, with the increasing focus on CLIL, fears and uncertainties begin to arise among potential stakeholders and practitioners in China. They wonder whether the CLIL approach represents the future of EFL, by playing a beneficial role in improving learners' learning outcomes; or merely a fad by over-generalising positive findings from specific circumstances to various heterogeneous contexts (B. Liu, 2010; X. W. Wang, 2011).

The leading cause of these fears and uncertainties mainly lies in the paucity of pertinent CLIL research in the Chinese context. This country has been so far a largely unproven territory to CLIL, with related research still in its infancy (Hu et al., 2014). A majority of the sporadic CLIL research in China remains centred on introducing CLIL concept and evaluating feasibility from theoretical perspectives (Luo, 2006; H. W. Yang & Xu, 2011), while little empirical research has been reported (Lei & Hu, 2014). Concerning the target context of Chinese higher education, the situation is more elusive. No official documents have been promulgated to endorse CLIL implementation from the policymakers' perspective, and little classroom praxis has been documented from the bottom-up level. The scarcity of "hard evidence" on CLIL achievements in Chinese education has triggered doubt and confusion among potential stakeholders. In comparison with the overwhelming bulk of CLIL practice and research in Europe, little research has shown to what extent these Europe-based findings can be extrapolated to non-European circumstances like China. Given the immense differences in sociocultural, linguistic, and educational aspects between China and Europe, more empirical evidence of CLIL effectiveness in the Chinese context is needed to defuse fears.

1.2 Rationale

In response to the gaps mentioned above, this present study sought to investigate CLIL effects on learners' academic performance in a Chinese higher education

programme. The rationale for the selection of this target context is provided in the following sections.

1.2.1 Positionality

Before articulating my choice of research context in the current study, one prerequisite issue is to allow my positionality to step in. According to Curtis, Murphy, and Shields (2013), positionality describes the identity taken by the researcher in relation to the research topic and research settings. The recognition of a researcher's identities might gain insight into how he/she approaches a research setting and engages with research participants. In this sense, a description and analysis of my multiple identities might better inform how I perceive my intended study and the research questions proposed.

My positionality towards the current research has been largely influenced by the entanglement of my multiple overlapping identities arose from my personal working and educational experiences. For over twenty years, I have been working as an English language teacher in a Chinese public university since I got it as my first job in life. The reason for my adherence to it lies in my commitment to the belief that effective English language education might equip my students with a helpful language tool for their globalised mobility and employability in future. However, as an insider of the Chinese college English education system, I am deeply aware of teachers and learners' dissatisfaction with the mainstream EFL approach and their bottom-up initiatives for an alternative solution. As a result, I have experienced a number of attempts to change the current English language education approach in my working context in the past two decades. I have undergone the ebb and flow of scores of alternatives, yet most of which ended in vain. The previous frustrating experiences in class made me relatively conservative to attempt a new educational approach.

In contrast to the insider status, my EdD study equips me with a researcher stance

to investigate CLIL effectiveness from an outsider perspective. My doctorate study allows me to access a large body of literature canvassing CLIL benefits in various contexts. It gradually changes my conservative attitude and arouses my interest in producing research-based evidence of CLIL effectiveness in my working context. It is anticipated that my research on this innovative approach might add evidence to gradually influence policymakers' understanding of CLIL and result in improvement in English language education from a top-down process. In turn, the shift in the authoritative allocation of values from policymakers, especially in a relatively centralized educational system like my working context, might as well enhance classroom practice, benefit grassroots teachers and students, and fulfil the commitment I have adhered to in the past two decades.

According to John W. Creswell (2013), the work of researchers is inevitably impacted by their positionalities. By identifying and acknowledging my positionality relatively transparently, I expect to demonstrate how my research focus has come into being. As illustrated in this section, the fluidity of my role as a teacher and a researcher, along with the interplay of my position as an insider and an outsider, has deeply influenced my research. It affects not only my research interest and context, but my choice of epistemological stance and research methodology as well.

1.2.2 Practical perspective

Practically, considerations for the selection of the target research context pertain to the following key words of 'Chinese', 'higher education', and 'academic performance'.

Firstly, the Chinese context is selected in the present study. Under both top-down and grassroots demand, China is seeking an alternative to the mainstream EFL approach, which entails opportunities for CLIL research in the Chinese context (Chang, 2006; Feng, 2009; Hu et al., 2014). With deepening participation in globalisation, the Chinese

economy has been greatly reconfigured and revitalized with growing interconnection with other parts of the world. The “increasingly borderless economic global dependency” calls for more workforce with the ability to communicate in a lingua franca, which in turn is becoming “a prerequisite for individual success” (Coyle, Hood, & Marsh, 2010, p. 8). It is highly likely, but not sufficiently documented, that many foreign language teaching approaches have been or are being scrutinized in China to seek the most suitable ones (Hu & Lei, 2014; Tong & Shi, 2012). Considering the number of L2 learning populations in China and the already hefty curricular burden in education, there might be considerable interest and need for the CLIL approach that has already exhibited its effectiveness in Europe. Thus an empirical CLIL study in the Chinese context like the present one might provide additional opportunity for Chinese learners’ better accommodation of globalised mobility and employability and maximize their potential for success. In addition, research on CLIL effectiveness in the Chinese context might generate impacts and insights on the transferability and feasibility of CLIL beyond Europe. It might contribute to the expansion of CLIL into the country with the largest number of English language learners, which in turn will facilitate the sustainability and future development of this dual-focused approach.

Secondly, higher education is selected as the target education level in the present study. This selection originated from my personal learning experiences and working context. In my undergraduate study in a Chinese university, I received English language education in the mainstream EFL approach, now known as the grammar-translation approach, taking time and efforts to recite texts and vocabulary. However, my good performance in the EFL course did not bring me advantages in my postgraduate study in a British university. I suffered great difficulties due to the lack of content related language training. When I started my teaching career, the ineffectiveness of the mainstream EFL approach appeared to be similarly apparent to my students. Thus, my study experiences as an English language learner and my profession as an English language teacher in the Chinese tertiary educational contexts provide me with a dual

view of both learners and teachers' dissatisfaction with the conventional EFL approach, together with their bottom-up initiatives searching for an alternative solution.

Besides personal motives and convenience, other factors like policy support and the availability of potential CLIL teachers have been taken into account for selecting higher education rather than other educational levels. In comparison with most European countries, China is a newcomer to bilingual education. As it happened, the official promotion to use English as the medium of instruction for specific courses in Chinese universities aligns with the basic principles of CLIL. However, the top-down policy support has been in effect in China for only about twenty years and has hitherto been limited merely to higher education, which makes attempts of CLIL in other educational levels more far-fetched (MOE, 2001). Another factor taken into account is the availability of qualified CLIL teachers. Without corresponding official endorsement, in Chinese primary and secondary education, it is not common to find bilingual teachers, not to say qualified CLIL teachers with capability in the CLIL language, proficiency in foreign language and content subject teaching, and more importantly, "an understanding of the CLIL approach and methodology" (Ioannou Georgiou, 2012, p. 500). In contrast, with a relatively higher educational background and more bilingual experience abroad, more qualified CLIL teachers are available to guarantee the successful implementation of CLIL in this study at the university level.

Lastly, among the variegated factors of measuring CLIL effectiveness, learners' academic performance has been chosen as the main research target. This selection is largely made in accordance with the exam-driven feature of Chinese education system (Yan & He, 2015). Both traditionally and contemporarily, learners' academic performance has been foregrounded by key stakeholders in the target research context of Chinese higher education. Good academic results, to some extent, indicate more career choices and job security to learners, more financial allocation and better admission to universities, as well as more public support and central government fiscal

appropriation to policymakers. As mentioned in the prior section, CLIL is still at an initial stage in China. The innovative CLIL approach may not be very familiar to some stakeholders. Purely theoretical descriptions of CLIL as “a fusion where the best of language education joins together with the best of general education”(Ioannou Georgiou, 2012, p. 496), or “the ultimate opportunity to practice and improve a foreign language” (Pérez-Vidal, 2013, p. 59) might not fully convince key stakeholders of CLIL effectiveness in the Chinese context. One practical solution to defuse their fears and uncertainties is to present them learners’ academic achievement after a certain period of CLIL exposure. A kind of “hard evidence” that CLIL can potentially enhance learners’ academic performance in a more efficient way than the mainstream EFL approach. In the present study, to be precise, CLIL learners’ academic performances in both content and language were assessed and compared with that of non-CLIL counterparts to reveal CLIL effectiveness. This practice largely coincides with the standpoint of Coyle et al. (2010, p. 1) that “the educational success of CLIL is in the content- and language - learning outcomes realized in classrooms”.

However, CLIL learners’ academic achievements in the duality of content and language are not the entire story of CLIL effectiveness. It is documented in the bulk of evidence that CLIL brings far more impacts on learners, such as cognitive development and intercultural awareness (Aguilar, 2017; Goris et al., 2019; Ting, 2011). CLIL learners’ academic achievement is not only the manifestation of “learners’ progression in knowledge, skills and understanding of the content”, and “the development of appropriate language knowledge and skills” (Coyle, 2007, p. 550). It also involves learners’ engagement in related cognitive processing and communication in the academic context, and deepening intercultural awareness. Therefore, in this study, CLIL learners’ progress in cognitive and cultural aspects was investigated alongside their academic performance in content and language. The four interrelated aspects of content, communication (language), cognition and culture merge to form the 4Cs framework. This framework is not only the theoretical basis of CLIL, but also the practical tool to

plan CLIL lessons, linking theory and practice (Coyle et al., 2010; Fazzi & Lasagabaster, 2021; Wiesemes, 2009).

1.2.3 Theoretical perspective

Theoretically, the present study attempts to add evidence to a succinct definition of CLIL with clear borderlines and ensure that CLIL practitioners and researchers are engaging in the same approach rather than others.

As a newly emerged educational approach, CLIL has gained a rapid and massive realisation across Europe and beyond. It becomes “the new fashionable approach and nearly everyone either wants to do it or wants to be seen to be doing it” (Ioannou Georgiou, 2012, p. 497). To some extent, this approach is on the brink of becoming a victim of its own success. It has aroused considerable debates and disagreements on its concept among researchers and has generated numerous theoretical frameworks comprising discrepant understandings of the construct of CLIL within different perspectives. In the literature, with the massive and extensive uptake in variegated educational settings over the past two decades, there appears more than 216 types of CLIL models from different perspectives developed from diverse sociolinguistic and political contexts (Aguilar & Muñoz, 2014; Coyle, 2007). Consequently, in an effort to accommodate the myriad of variant models, CLIL becomes an umbrella term. Any efforts to provide an explicit and theoretically ‘fit’ definition or to draw a definite boundary of CLIL from either policymakers’ or practitioners’ perspective prove to be risky and controversial (Cenoz, Genesee, & Gorter, 2014; Dalton-Puffer, Llinares, Lorenzo, & Nikula, 2014).

In response to the complications on “how large or how small the CLIL umbrella is”, one possible solution is to start from the common denominators of CLIL and adopt an integrated perspective, as detailed in Section 2.1 (Ioannou Georgiou, 2012, p. 497). For instance, by incorporating an integrated perspective, the dichotomies of content and

language in CLIL could be replaced by a continuum that allows for diverse content and language integrated programmes (Met, 1998). In this study, an integrated grid has been adopted to divide four quadrants by two intersecting continua (Paran, 2013), together with a set of four typological continua (Cenoz, 2015). With the position in the grid and the four typological continua, it could be further elaborated that content learning in CLIL is placed in the ‘knowledge for learning’ domain, involving ‘content and cognition’; whereas the ‘culture-bound phenomenon’ of language learning is regarded as ‘a medium for learning’, in which ‘communication (language) and intercultural understanding’ are included (Coyle, 2007). This articulation contributes to the literature by developing the two main foci of CLIL into “one integrated learning process” with four interrelated facets of content, communication, cognition, and culture (Dalton-Puffer, 2011, p. 196). The four facets interplay with each other at diverse dimensions within a specific context and consequently converge to establish the 4Cs framework to illustrate a quality CLIL learning process from a holistic view (Coyle et al., 2010, p. 41).

CLIL is a complex phenomenon that has been conceptualized from different perspectives, reflecting various stakeholders’ stance and ultimately informing different methodologies and research methods. Thus, with the change of geo-political and economic landscape, these prototypical characteristics “need to be honed, sharpened, and fine-tuned in line with the demands of the contexts where it is being applied” (Pérez-Cañado, 2016, p. 316), especially in the target context of the current study, where is under-researched. With the above considerations, the research focus in the present study lies in investigating CLIL effects on learners’ academic performance in a Chinese tertiary education programme. It is expected that the research would increase our insights into this knowledge area with an in-depth of the ‘voice’ of students in the under-researched context. Further details would be presented in the following sections.

1.3 Aims

The main aim of the current study is thus to gain new insights into the CLIL approach by investigating CLIL effects on learners' academic performance in a Chinese tertiary educational context. To this end, Coyle's 4Cs framework in CLIL research is adopted. It is expected that based on this theoretical framework, findings of this study could support those in the literature that concern CLIL as an evidenced alternative approach to enhance learners' academic proficiency for globalised mobility and employability, or indeed reveal the limitations. Hence, the main research question is proposed as following:

To what extent does CLIL facilitate Chinese college learners' development of the desired academic performance?

In addition, the main research question could be decoded into the following sub-questions:

- 1) Does CLIL improve learners' academic performance in subject content?
- 2) Does CLIL enhance learners' linguistic and communicative competence?
- 3) Does CLIL increase learners' intellectually cognitive proficiency?
- 4) Does CLIL facilitate learners' intercultural understanding?

As a whole, by examining and understanding potential outcomes of CLIL approach in the target context, the study is expected to present a compelling, theory-driven explanation for how and why CLIL affects learners' academic performance, and provide some theoretical, empirical and pedagogical implications for both researchers and practitioners to facilitate learners' academic proficiency as their knowledge learning process.

1.4 Overview of the thesis structure

The thesis contains six chapters. After introducing the main background and context where the study takes place, Chapter One highlights the significance of the

present study.

Chapter Two presents the existing body of knowledge relating to the research focus. There is an argumentation on some key issues relevant to the research topic, followed by literature covering related studies.

Chapter Three demonstrates the theoretical framework adopted in this study. Four key components, namely content, communication, cognition and culture are drilled from the 4Cs framework to form the main framework to investigate the effectiveness of the CLIL approach in an authentic Chinese college setting.

Chapter Four illustrates how the aforementioned framework relates to the possible methodological approach adopted. Under the philosophical implications of pragmatism, mixed-methods research was conducted to gain a rich and nuanced understanding of CLIL effects on Chinese college learners. In addition, this chapter focuses on an important part of social science research, namely ethical issues, to raise the awareness of moral concerns in designing, conducting and analysing the study.

Chapter Five describes in detail the results and key themes emerging from the two stages of data analysis, including quantitative and qualitative data.

Chapter Six maps a more extensive discussion, based on findings from Chapter Five, highlighting potential links between the literature, theoretical framework, the methodology adopted and the interpretations of findings. In addition, the last chapter seeks to depict knowledge contributions of the study and discuss further the implications and recommendations accordingly.

CHAPTER TWO- Literature Review

This chapter reviews the extant body of knowledge pertinent to my research scope, namely the conceptualisation of the umbrella term of CLIL and different perspectives of synthesising previous research on CLIL outcomes, particularly concerning the target research context of Chinese higher education. The logic behind the organisation of the wide-ranging yet somewhat conflicting literature hinges upon the principle of relevancy concerning the aim of this study. Efforts have been made in searching detailed research and presenting it in the following main sections accordingly. Each section comprises sub-sections for maximum clarity of the literature review scope.

In Section 2.1, the review of relevant literature on definitions helps to conceptualize what CLIL might be in the target research circumstances amid the diverse formats of CLIL for reference and reflection in subsequent research. In Section 2.2, the synthesis and stocktaking of related CLIL research facilitate to inform the research design in this study, aiming to bridge the gap in existing knowledge by providing a more in-depth investigation of CLIL effectiveness from an integrated perspective.

2.1 Delineating the umbrella term of CLIL

Although CLIL has been frequently labelled as a newly emerged educational approach, the practice of teaching content in an L2 is nothing new. It dates back to approximately 3000 BC when Akkadians, conquests of Sumer, learned content subjects of “theology, botany, zoology, mathematics and geography in the medium of Sumerian” (Lyster, 2007, p. 7). In effect, similar practices could be discovered throughout history, from the Roman era when some Romans were educated in Greek to French immersion programmes in Canada in the 1970s and bilingual movements in the USA in the 1980s (Coyle, 2007). Based on those predecessors, the concept of CLIL was coined in 1994 by a cohort of key players from the European Network of Administrators, Researchers and Practitioners (EUROCLIC) to describe a dual-focused instruction form combining

both disciplinary knowledge and foreign language enrichment (Coyle, 2007).

As an innovative educational approach, CLIL has been considered from its inception by both EU and national-level policymakers and local grassroots actors as “a pragmatic European solution” to globalisation (Dalton-Puffer, Hüttner, & Smit, 2021; European Commission, 1995; Eurydice, 2006; Marsh, 2002, p. 11). Fuelled by this educational belief, CLIL has grown in the past two decades from a term unheard of to a commonplace throughout variegated educational settings in Europe. The extent of its coverage and the pace of spread “has surprised even its most ardent advocates” (Maljers, Marsh, & Wolff, 2007, p. 7).

With the massive and extensive uptake in diverse educational contexts in continental Europe in the past two decades, at least 216 types of CLIL models have emerged in terms of “variables”, “operating factors”, or “environmental parameters”, like duration, compulsory status, intensity, starting age and language proficiency (Coyle, 2007, 2008; Grin, 2005). In an effort to accommodate the myriad of variant models, CLIL becomes “a generic umbrella term” (Dalton-Puffer et al., 2010). The forming of this umbrella term is largely caused by “a wide range of initiatives” from different perspectives developed from diverse sociolinguistic and political contexts in Europe (Marsh, 2002, p. 59). As commented by Hüttner and Smit (2014, p. 166), “(it) is an inevitable consequence of CLIL being a set of localized responses to the rise of English as a global lingua franca”. To some CLIL advocates, the forming of the umbrella term is a signpost of CLIL success, while others claim that “the CLIL umbrella might be stretching too much and that CLIL might be on the verge of becoming a victim of its own success” (Ioannou Georgiou, 2012, p. 497). The potential strengths and complications of the umbrella term are illustrated below.

On the positive side, conceptualising CLIL in an umbrella term endows the approach with an inclusive and flexible nature. The inclusiveness allows CLIL to

embrace a plethora of formats “that it currently includes and which many advocates favour” (Cenoz et al., 2014, p. 257). Besides inclusiveness, the umbrella term entails CLIL a relatively high degree of flexibility. The umbrella term has been claimed to allow content and language integrated learning to occur “in varied, dynamic and relevant learning contexts built on ‘bottom-up’ initiatives as well as ‘top-down’ policy” (Coyle, 2007, p. 546), and to accommodate “the linguistic diversity of the European landscape” (Pérez-Cañado, 2017, p. 82). This flexible nature empowers CLIL with a context-specific stance in commensuration with the “post-method pedagogy of peculiarity” proposed by Kumaravadivelu (2001, p. 538). With CLIL being “the overarching concept”, heterogeneous constituent formats of CLIL can be developed in accordance with different contextual and pedagogical needs (Dalton-Puffer et al., 2014, p. 217) and “avoid the one-size-fits-all model”, which has been proven to collapse due to the inability to adapt to variations (Pérez, Lorenzo, & Pavón, 2016, p. 501). The umbrella term, as summed up by Pérez-Cañado (2017, p. 82), makes CLIL like “a blanket on a large bed shared by many children, each pulling in their own direction, and it is precisely its flexible nature and numerous variations which have allowed it to stretch to meet all needs”.

However, such a flexible and inclusive conceptualisation of CLIL is not exempt from potential weakness (Coyle, 2007). Accompanying the establishment of the umbrella term, issues of generalisability and terminological overlap with other bilingual approaches come up, which might eventually make the CLIL blanket “torn to shreds” if not appropriately handled (Dickey, 2004, p. 11).

The first accusation of the umbrella term of CLIL lies in the potential difficulty of generalising CLIL research findings and practical experience across contexts. As aforementioned, ever since its inception, the term CLIL “has acquired some characteristics of a brand name, complete with the symbolic capital of positive description: innovative, modern, effective, efficient and forward-looking” (Dalton-

Puffer et al., 2010, p. 3). CLIL thus becomes a “fashionable” approach (Ioannou Georgiou, 2012, p. 497). In the process of including diverse renditions, CLIL might face the risk of “being watered down or losing the essential characteristics” that have led to its prior success or leading to potential misapplication and misappropriation. More importantly, “the extensive variety of CLIL models ... may lead to communication between researchers, teachers, and policymakers being obstructed” (Ioannou Georgiou, 2012, p. 498). Due to the flexibility and inclusiveness of the umbrella term, one CLIL programme may differ significantly from another. The large discrepancy of different CLIL programmes might “make it a tricky business” to review previous research outcomes, replicate prior studies, compare research findings, and generalize practical experience (Dalton-Puffer et al., 2014, p. 213). Therefore, in carrying out CLIL research like the present study, it is incumbent on researchers and practitioners to provide comprehensive and unequivocal accounts of the CLIL programmes investigated for better communication and sharing in research, practice, and policymaking.

Another potential complication elicited by the umbrella use is the “terminological puzzle concerning the confusion between CLIL and other similar approaches like CBI (Content Based Instruction), Immersion, and Bilingual Education” (Dalton-Puffer et al., 2014, p. 213). It seems little consensus has been hitherto reached amongst seminal CLIL researchers on the similarities and distinctions between CLIL and its akin approaches. One example is from Coyle (2008, p. 97), who differentiates CLIL from bilingual education, immersion, CBI or ESP (English for Specific Purposes) because CLIL “is based on an integrated approach where both language and content are conceptualized on a continuum without an implied preference for either”. In contrast, Jäppinen (2005, p. 149) claims CLIL includes “immersion and some forms of bilingual education”, while Pérez-Vidal and Juan-Garau (2010) seem to find CLIL and bilingual education to be interchangeable terms. These contradictory statements reveal that CLIL “is afflicted with a high lack of terminological clarity” (Paran, 2013, p. 319) and a lack of “a

standardized CLIL blueprint” (Van de Craen, Ceuleers, & Mondt, 2007, p. 197). In the process of including diverse variations of CLIL practice into the umbrella term and taking advantage of its flexibility, it is worth noticing an alarming situation that “there is no widely accepted definition(s) of CLIL and, moreover, no clear understanding of different versions of CLIL” (Cenoz et al., 2014, p. 258). The lack of conceptual clarity, as further noted by Cenoz et al. (2014), might hamper research efforts and educational initiatives.

On reviewing the strengths of the umbrella term of CLIL, it might be inferred that the flexibility and inclusiveness of the umbrella term allow CLIL to cater to diverse educational contexts within and beyond Europe, including the target research context of Chinese higher education. Meanwhile, given the potential complications as aforementioned, there seems “a critical need to refine the definition of CLIL in ways that systematically and coherently recognize this diversity of formats” (Cenoz et al., 2014, p. 257). With regard to the present study, the two conflicting sides of the umbrella term eventually bring about the question of “how large or how small” the CLIL umbrella should be in the target research context (Ioannou Georgiou, 2012, p. 497). Therefore, a succinct definition of CLIL with clear borderlines becomes the prerequisite for subsequent literature review. Two seminal conceptual forms will be subsequently articulated, corresponding to the perspectives of policymakers and practitioners, respectively.

2.1.1 Conceptualising CLIL from policymakers’ perspective

From policymakers’ perspective, without policy instruction of “how the languages are to be taught, learned, and assessed”, CLIL may take any form or shape at the individual level (Sylvén, 2013, p. 303). Among the CLIL definitions from policymakers’ perspective, emphasis has been mainly on “coming to terms with the phenomenon” and providing overview information to potential stakeholders (Dalton-Puffer et al., 2010, p. 9). One seminal EU-endorsed definition is put forward by Marsh (2002, p. 15), the co-

founder and leading advocate of CLIL:

(CLIL is a) dual-focused educational context in which an additional language, thus not usually the first foreign language of the learners involved, is used as a medium in the teaching and learning of non-language content.

With respect to the present study in the Chinese higher education context, this definition appears to be problematic for its political contextualisation and academic overstretch, as detailed below.

From a political perspective, this official definition serves the goals of facilitating multilingualism in education and European integration endorsed by the European Union (EU), the most prominent supporter of CLIL (European Commission, 1995; European Commission Communication, 2003; Ioannou Georgiou, 2012). As featured in a series of EU documents, CLIL has been invested with “a major contribution to make to the Union’s language learning goals”, namely shaping multilingual citizens in a region of at least 23 official languages (European Commission Communication, 2003, p. 8). However, with the prevalence of CLIL beyond European boundaries, this politically contextualized definition and its value-ladenness might impede its acceptance in other nations. For instance, the clarification of “an additional language” as “not usually the first foreign language” corresponds with the EU policy of “Mother Tongue +2 for all citizens” (Eurobarometer, 2012). However, it might not be practical in other contexts such as China.

Regarding the target context of the present study in a mainly habitually monolingual state, there is little political necessity for Chinese policymakers to encourage the learning of a second foreign language among most university students. The acquisition of Chinese and English, probably the two most influential languages

globally, meets most Chinese university graduates' demands. In addition, there is little political drive behind the learning of a foreign language (predominantly English) for national integration in the Chinese context, which is hugely different from the situation that the EU faces.

In addition, this definition is accused of outlining the dual foci in a relatively over-generalized and abstract way, which might not well inform application to local practice (Cenoz et al., 2014). For instance, the educational objectives of CLIL are discounted in this definition, not informing practitioners what competence to improve, the language, the content, or both. Moreover, the expression of “dual-focused” could be understood differently, making it difficult for practitioners to follow. According to Marsh (2002), as long as a dual focus on content and language for instruction is achieved, it can be defined as CLIL, no matter what percentage of the two. Cenoz et al. (2014, p. 245) find this view overstretched by raising a counter-example that “it is difficult to imagine a traditional non-CLIL L2/foreign language class with a less than 10 percent focus on some type of content”.

Another complication of this broad definition is that it delineates the umbrella term in an overlarge way and might not differentiate CLIL from similar bilingual approaches like content based instruction (CBI) or immersion, resulting in more controversy than consensus (Wolff, 2012). Marsh's (2002) flexible definition might be excessively inclusive, at the cost of precision. Such an overstretched way of conceptualising CLIL might make it lose its practical or theoretical utility (Cenoz et al., 2014).

In response to the politically contextualized and academically “large” way of conceptualising CLIL from policymakers' perspective in Marsh's (2002) definition, alternative defining approaches have been proposed from practitioners' perspective to postulate a “smaller or narrowed” definition of CLIL, which will be expounded in the following section.

2.1.2 Conceptualising CLIL from practitioners' perspective

To unravel the political tightness and conceptual overstretch of Marsh's (2002) definition, researchers attempt to conceptualize CLIL in other ways (Lasagabaster & Sierra, 2010; Pérez-Cañado, 2017; Surmont, Van de Craen, Struys, & Somers, 2014). One alternative approach is to provide a distinctive characterisation of CLIL from practitioners' perspective. To distinguish CLIL from immersion, Lasagabaster and Sierra (2010) propose seven differential aspects: the language of instruction, teachers, starting age, teaching materials, language objective, immigrant students, and research, most of which are contextual factors. Despite the accusation of "doing little to resolve the confusion as it suffers from internal contradictions and ungeneralizable data" (Somers & Surmont, 2012, p. 115), the attempt to use contextual factors as the main criteria for characterising CLIL sheds light on subsequent efforts.

Drawing on previous attempts to characterize CLIL and its practice in a range of contexts, Dalton-Puffer (2011, p. 183) distils six "prototypical characteristics" to exemplify a typical CLIL programme "in Europe, South America, and many parts of Asia", which can be summarized as follows:

- (1) **Target language:** Instead of a second language, CLIL tends to use a foreign language or a lingua franca that learners do not regularly use outside classroom.
- (2) **Dominant CLIL language:** English.
- (3) **Teacher selection:** non-native speakers of the CLIL language rather than native speakers, and content teachers instead of language teachers, because "classroom content is not so much taken from everyday life or the general content of the target language culture but rather from content subjects, from academic/scientific disciplines or from the professions" (Wolff, 2012, p. 107).
- (4) **Curricular classification:** timetabled as content courses alongside foreign language lessons.
- (5) **Language equilibrium:** less than 50% of the curriculum is taught in the CLIL language.

- (6) **Starting age:** usually since secondary education when literacy skills in L1 have been acquired.

The provision of the practice-driven prototypical characteristics aims to make CLIL a well-recognized and valuable construct for optimising L2 learning. In turn, these essential properties are expected to be guidelines to ensure CLIL effectiveness in practice in different settings. Nikula (2015) noted that one major advance in the defining characteristics is the clear prioritising of the content subject in Characteristic (4). CLIL is “not language-oriented” but “foremost content subject teaching and learning” (Lasagabaster & Sierra, 2010, p. 108).

However, in an effort to differentiate CLIL from other bilingual approaches, Dalton-Puffer’s (2011) set of prototypical characteristics has been challenged for tightly constraining the umbrella term, at the cost of an over-narrow understanding of CLIL and its potentials (Cenoz, 2015). This weakness has been largely caused by the disregard of the dynamic nature of CLIL in response to the fast-changing socio-political and educational situations in the past years, while the features identified by Dalton-Puffer was at the level of 2011, based on what research was available then.

Characteristic (1) Target Language: the exclusion of a language that is not “a foreign language or a lingua franca” concerning its use outside the classroom might be appropriate around 2011, but not tenable in various parts of the world today, including the research context of Chinese higher education. The boundaries of second/foreign languages have been blurred, more or less, by the increased mobility of the population in current dynamic societies.

Characteristic (2) Dominant CLIL Language: in comparison with the situation around 2011, the dominance of English as the CLIL language seems still unshakable under current circumstances. Given the significant role in foreign trade, knowledge

proliferation, and higher education internationalisation, English will gain its ascendancy in the CLIL language without much perceptible debate in the target research context of Chinese higher education. However, with the rise of emerging economies, changes might occur in the future.

Characteristic (3) Teacher Selection: Cenoz (2015, p. 21) finds the stratification of native and non-native CLIL language speakers not justifiable and draws an analogy that “foreign language lessons are never considered different types of programmes on the basis of being taught by native or non-native teachers”. Globalised employability and mobility of the workforce have already obscured the distinction between native and non-native speakers. The refusal of multilingual teachers with qualified communicative and teaching competence in both native and the CLIL language is not applicable (Lin, 2015).

Characteristic (5) Language equilibrium: it is evidenced that this equilibrium is hard to maintain, especially in tertiary education contexts where more than half of the curriculum taught in the CLIL language has become commonplace in Europe and beyond (Cenoz et al., 2014; Pérez-Vidal & Juan-Garau, 2010).

Characteristic (6) Starting age: it is believed by some CLIL scholars that the introduction of CLIL education should be from secondary education upon learners’ acquisition of literacy skills in L1 (Dalton-Puffer, 2011; Dalton-Puffer & Smit, 2013)). However, the rapid-transforming socio-political and educational contexts of CLIL today make the “one size fits all” criterion impossible, and contrary examples can be found in existing CLIL practices from preschool education throughout Europe and further afield (Hüttner & Smit, 2014, p. 164).

In short, Dalton-Puffer’s (2011) prototypical characteristics might have well facilitated researchers to recognize the complexity of the contexts where CLIL situates

and have well informed potential practitioners of CLIL classroom practice. However, with the change of contexts applied, these prototypical characteristics need to be adjusted, especially in this study where few CLIL precedents could be found.

In view of the articulation in Section 2.1.1 and 2.1.2, CLIL can be regarded as a complex phenomenon that has been conceptualized from different perspectives. The different conceptual forms reflect different stakeholders' stances of viewing CLIL, ultimately informing different methodologies and outcomes in CLIL research, as addressed in the subsequent section.

2.1.3 Conceptualising CLIL from an integrated perspective

As shown in previous sections, it proves risky and controversial to present an evident and theoretically 'fit' definition or draw a coherent borderline of CLIL. One possible solution is to start from the common denominators of CLIL and adopt an integrated perspective in the present study. By incorporating an integrated perspective, content and language dichotomies in CLIL will be replaced by a continuum (Met, 1998). Further on, an integrated grid will be divided by two intersecting continua (Paran, 2013), together with a set of four continua to replace the aforementioned prototypical characteristics (Baker, 2011; Cenoz, 2015).

Met's (1998) continuum of content and language integration

As agreed by both CLIL proponents and opponents, one common denominator of CLIL is its categorisation as a dual-focused content and language integrated programme (Bruton, 2015; Cenoz et al., 2014; Dalton-Puffer et al., 2014; Marsh, 2002). This consensus entails the possibility to differentiate CLIL from other similar programmes via Met's (1998, p. 41) continuum of content and language integration.

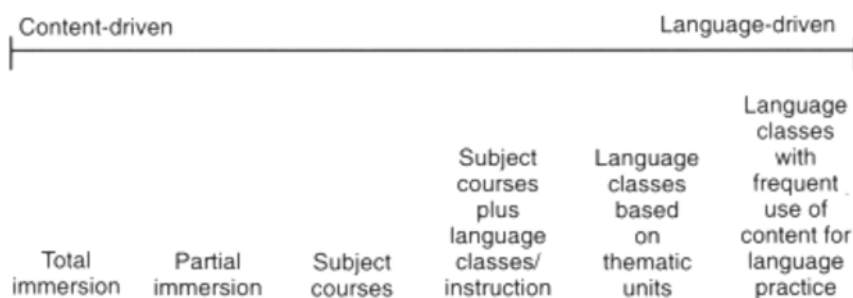


Figure 2.1 The continuum of content and language integration

As illustrated in Figure 2.1, Met’s (1998) continuum allows for diverse content and language integrated programmes according to the extent of integration. The position of each integrated programme is largely determined by “the role that content plays in relation to language development and the amount of explicit language instruction students receive” (Met, 1998, p. 40). At the right end of the continuum are language-driven content programmes like formal foreign language courses in which language is learned as a subject through unsystematic content matter already taught in L1 (Met, 1998). Moving leftwards along the continuum are programmes combining content courses with language courses. In accordance with Dalton-Puffer’s (2011, p. 184) prototypical Characteristic (4), CLIL is primarily categorized in the curriculum as content courses, alongside target language courses which “normally continue as a subject”. This characteristic entails the baseline of CLIL in the continuum, which can be positioned in a range shifting from “subject courses plus language class/instruction” towards the left side of the continuum (Met, 1998, p. 40). At the left end of the continuum are content-driven language programmes like total immersion, in which an additional language acts as a vehicle for content learning. In such programmes, learners’ mastery of content knowledge is considered equally important as language acquisition. This dual-focused nature meets the common denominator of CLIL and can serve as the left end of the scope. In a word, conceptualising CLIL in a continuum rather than dichotomies allows even greater flexibility given the dynamic and diverse realisations of the umbrella term CLIL.

Met's (1998) continuum has proven effective in covering all the language-content integration models in diverse countries and describing the range of settings these models entail (A. Llinares & McCabe, 2020; Ruiz de Zarobe, 2013; Tedick & Cammarata, 2012). However, this encompassing continuum might not reflect "the extent of the actual integration of language and content in teaching" (Paran, 2013, p. 320), which might lead to confusions on the difference between Content Teaching in an Additional Language and CLIL. As aforementioned, CLIL "is based on an integrated approach where both language and content are conceptualized on a continuum without an implied preference for either" (Coyle, 2008, p. 97). If a subject course is taught in an additional language but without any language focus, its position still falls in the CLIL segment in Met's (1998) continuum. However, it might not be categorized as CLIL because its "content and language are not integrated" (Paran, 2013, p. 320). This contradiction is also reflected in CLIL researchers' discrepant views of immersion, as exemplified in Section 2.1.1. Some seminal researchers claim CLIL includes immersion in accordance with its position in Met's (1998) continuum (Jäppinen, 2005; Pérez-Cañado, 2017), while others disagree because there seems to be a lack of integration (Coyle, 2008; Lasagabaster & Ruiz de Zarobe, 2010), though "yet in itself little of a means to discern between CLIL and immersion" (Somers & Surmont, 2012, p. 115).

As we can see, a single line segment along the continuum is not sufficient to conceptualize the actual integration of content and language in CLIL. Further extended forms are needed on the basis of this continuum.

Paran's (2013) integrated grid from practitioners' perspective

In an effort to manifest the actual integration of content and language at curriculum and classroom level, namely from practitioners' perspective, Paran (2013, p. 321) adopts an integrated grid divided by a vertical axis of Content Focus vs. Language Focus, and a horizontal axis of Content Objectives vs. Language Objectives (Figure 2.2).

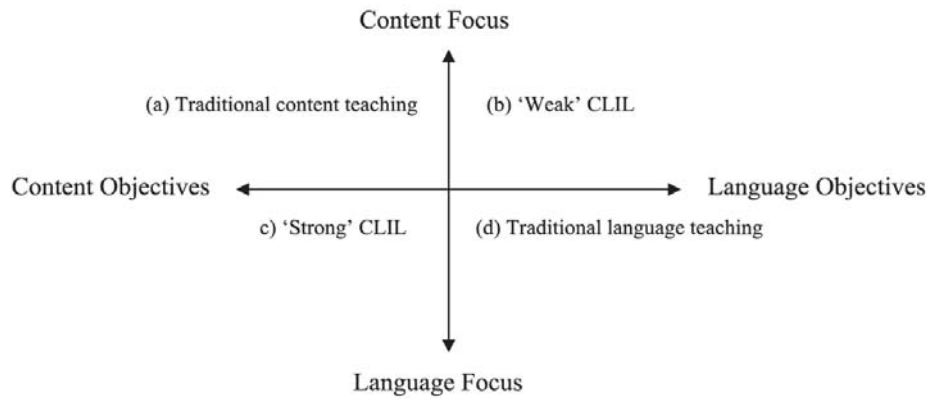


Figure 2.2 *The integrated grid of CLIL* (Paran, 2013, p. 321)

As illustrated in Figure 2.2, the grid is divided into four quadrants by two intersecting continua, which allows for diverse content-language integrated approaches. The position of each approach is mainly decided by their extent of integration in two dimensions, namely the objectives in the curriculum and the focus in the lesson. With both curricular objective and lesson focus in the content subject, Quadrant (a) illustrates conventional content classes in which language is not focused and assessed. In a similar vein, Quadrant (d) presents conventional language classes in which language is learned as a subject through unsystematic content matter; the content of texts is regarded as an incidental vehicle for language learning (Hüttner & Smit, 2014).

The quadrants of interest in CLIL are Quadrant (b) and (c), which are referred to as “weak” and “strong” forms of CLIL (Paran, 2013, p. 321). In Quadrant (b) of weak CLIL, the curricular objectives are language-oriented, but the content focus is not incidental. It is “highlighted, focused on and discussed” (Paran, 2013, p. 321). This form of CLIL has been termed by Lyster and Ballinger (2011, p. 280) as “language classes with thematic units”, which “entails no high-risk assessments of content knowledge”. One example of this weak form of CLIL in China, as quoted by Lyster & Ballinger (2011, p. 280), are some “subjects that are not part of the formal curriculum such as ‘nature and society’ and ‘science and life’ to be taught usually for two lessons per week”.

As to Quadrant (c) of strong CLIL, the curricular objectives are content-oriented, but there is still a focus on language in the class. The basic implication behind the conjunction of the two is that languages are “not learned first and then used”, but “are learned by being used” (Cenoz, 2015, p. 17). Both content knowledge and language development are assessed “in substantive ways” (Lyster & Ballinger, 2011, p. 280). Paran (2013, p. 322) notes that this strong form of CLIL is “not new”. It is a common practice in Africa that learners are schooled in an L2, especially in a majority colonial language. “What is new is the emergence of CLIL in countries where education has traditionally been in the L1” (Paran, 2013, p. 322). Most extant models of CLIL correspond to this quadrant of conceptualisation, and it is where the current model of this present study situates itself.

It can be seen from the illustration that Paran’s (2013) integrated grid is mainly at a curricular and classroom level for practitioners’ consideration. However, CLIL is not just about classroom education. From policymakers’ perspective, sociocultural, political, and economic factors present in the account of CLIL provision (Pérez-Cañado, 2020). As pinpointed by Dalton-Puffer (2011, p. 183), whether a programme is CLIL or not “often depends as much on its cultural and political frame of reference as on the actual characteristics of the programme”, highlighting the essential properties of CLIL from policymakers’ perspective.

Cenoz’s (2015) four typological continua from policymakers’ perspective

Besides the dual integration of content and language, another common denominator of CLIL is its language categorisation as a form of additive bilingualism, which is agreed by both CLIL proponents and opponents (Banegas, 2014; Dallinger, Jonkmann, Hollm, & Fiege, 2016; Dalton-Puffer et al., 2014). This consensus entails the possibility to differentiate CLIL from other language approaches on the basis of Baker’s (2011) established typology of bilingual education. In this typology, different

types of bilingual education are classified according to four dimensions from policymakers' broad social and ideological terms, namely "typical type of child, language of the classroom, societal and educational aim, and aim in language outcome" (Baker, 2011, p. 209). To avoid suggesting different bilingual programmes as static systems and provide scopes to allow for the dynamics and diversity in CLIL practices as aforementioned, Cenoz et al. (2014, p. 17) developed the four typologies into four corresponding continua for CLIL conceptual clarification from the language perspective.

As seen from the four language continua in Figure 2.3, Baker's (2011, p. 209) four dimensions are taken as the essentials to define CLIL programmes. Different language programmes can be placed at various points along the continua. In this figure, the expression of L1 is used as a generic term to indicate the original language of instruction in non-CLIL classrooms, which might be the native language, the majority language, the mother tongue, and the traditional language, etc. In contrast, the term L2 is adopted to indicate the language added to the CLIL classroom besides the native language. It might be the CLIL language, the foreign language, the second language, and the additional language, etc.

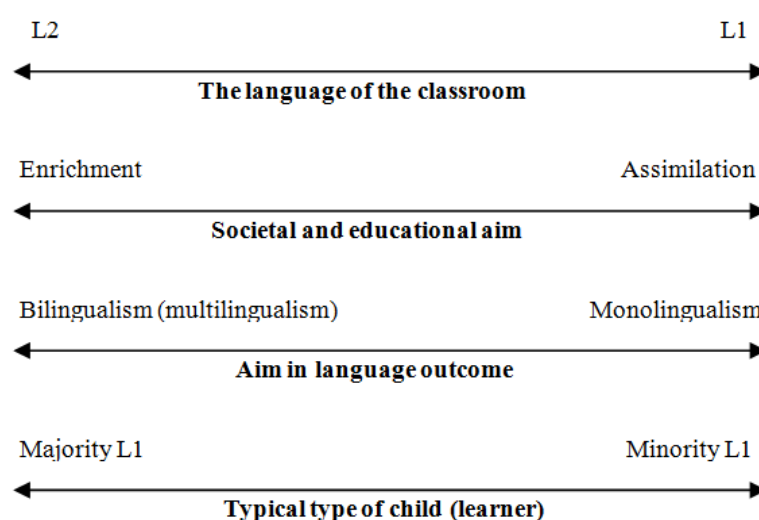


Figure 2.3 Language continua of CLIL, adapted from Cenoz, (2014, p. 18)

In the four continua, most CLIL programmes are featured to shift towards the left end of each continuum. The first continuum indicates the medium of instruction for the

whole curriculum. In CLIL programmes, the number of courses taught in L2 could range from one subject to all, while the percentage of L2 used in each course could vary from zero to full percent. This continuum makes allowance for the diversity of CLIL language use in the whole curriculum. In the second continuum, CLIL programmes aim to enrich language repertoire and culture diversity at no cost of L1 in educational and social contexts, instead of assimilating language minority learners into mainstream society by endangering their native tongue (Baker, 2011). This societal and educational aim is adopted to prepare students with better communicative competence for internationalized mobility and employability. In a similar vein, the aim in language outcome in CLIL programmes is to promote learners' bilingual (sometimes multilingual) and biliterate competence, rather than replacing the L1 with L2, resulting in monolingualism. The specific aim and outcome of different language skills in different CLIL programmes is largely influenced by the use of CLIL language in a sociolinguistic context (Cenoz et al., 2014). As to the type of learners, most CLIL programmes target majority L1 learners instead of minority L1. However, due to the globalised mobility of the population, increasing CLIL programmes might have heterogeneous learners who speak different languages than the majority L1. The four typological continua illustrate the essentials of CLIL from policymakers' perspective and differentiate it from other bilingual language programmes.

However, the four typological essentials are not sufficient to solely sustain the unique features of CLIL. The continua mainly conceptualize CLIL from policymakers' perspective, like overall educational aims, language input, expectant language output, and typical learners. Their intrinsic limitation of addressing the CLIL classroom process is apparent.

Given the preceding flaws in Met's continuum, Paran's integrated grid and Cenoz's four typological continua, CLIL research calls for more holistic and broad frameworks. CLIL interrelates not only with dynamically changing social contexts

external to learners, but also with learners' internal cognitive and affective states. These interrelationships eventually lead CLIL research into a holistic perspective, as illustrated in the following sections.

2.2 Review of CLIL studies

As a newly emergent educational approach, the effectiveness of CLIL needs to be substantiated to bring potential stakeholders credence and confidence (Devos, 2016; Pérez-Cañado, 2012). Thanks to the surge in CLIL realisations throughout Europe and beyond in the past two decades, numerous empirical studies have been documented to investigate its effectiveness (Broca, 2016; Dalton-Puffer et al., 2021; Pérez-Cañado, 2020). Considering the twofold aims of content and language, it is not surprising a majority of extant CLIL studies are in the area of gauging learners' content and language achievements. However, the research outcomes from different studies have portrayed two inclusive or even contrasting tales. On the one hand, researchers in support of the "success story" believe that CLIL gains have been extensively documented in much of the existing research (Pérez-Cañado, 2012; Ruiz de Zarobe, 2015); while on the other hand, sceptics find the "success story" questionable, arguing that few CLIL studies have established conclusive links between positive learning outcomes and CLIL exposure, mostly due to the interference of confounding factors like methodological flaws and social selection (McDougald, 2016; Van de Craen, Mondt, Allain, & Gao, 2007). Therefore, to ensure "the gains observed are truly ascribable to CLIL practice", this section attempts to take stock of extant research findings on CLIL learning outcomes in content and language (Pérez-Cañado, 2012, p. 330).

As illustrated in the preceding sections, CLIL has been conceptualized from policymakers', practitioners' and integrated perspectives respectively. The distinction among these perspectives visualizes the CLIL research space according to researchers' positioning with regard to "how close" they get to the phenomenon in question and

“how fine-grained” the views they take (Smit & Dalton-Puffer, 2007, p. 12), informing diversified research methodologies, namely quantitative, qualitative and mixed-methods approach.

It is true that categorising CLIL studies of this kind of quantitative, qualitative and mixed-methods brings about “inevitable simplifications”. This categorisation of CLIL studies, as argued by Bonnet (2012), does not aim to raise debates on which of these perspectives is more accurate or superior to investigate the CLIL effectiveness, but to resolve the respective shortcomings, and to gain a richer and more nuanced understanding of CLIL effectiveness, which will be detailed in the accompanying sections.

In the following section, the research outcomes of CLIL effectiveness in quantitative studies is reviewed respectively in terms of language and content to inform the research design of this study, together with reflections on their specific methodological strengths and weaknesses.

2.2.1 Overview of CLIL effect in quantitative studies

From the policymakers’ perspective, CLIL has been regarded as a societal phenomenon “with relatively stable characteristics that can be explained in terms of structural determinist ideas”, which implies that the aggregate concept of CLIL follows certain regularities that research needs to identify (Bonnet, 2012, p. 68). With regard to CLIL learners’ achievement, the policymakers’ perspective conceptualizes it as learners’ “performance in problem-solving situations” and reduces it to “an observable surface phenomenon” of measurable content and language input and output (Bonnet, 2012, p. 72). This perspective of CLIL is largely informed by positivist epistemological thinking, which calls for quantitative research methodology and hypothesis testing of CLIL effects on learner attainment through numerical analysis.

Among quantitative CLIL studies, the emphasis has been mainly on entailing policymakers with the evaluation of CLIL effectiveness (Ruiz de Zarobe, 2015). Hence a high proportion of such studies are large-scale quantitative (quasi)experiments contrasting CLIL and non-CLIL learning results for quantitative analysis (N. Hughes, 2013; Pérez-Cañado, 2020). With regard to research findings, most quantitative studies report promising gains in language performance in diverse educational contexts (Pérez-Vidal & Roquet, 2015; Ruiz de Zarobe, 2015). As illustrated in Lasagabaster's (2008) investigation of 198 Spanish secondary school learners, test results showed that CLIL learners notably outperformed non-CLIL controls in general language competence, attempting to establish causality between CLIL exposure and participants' language attainment. In a similar vein, other quantitative studies present CLIL advantages in different language skills, such as vocabulary use (Llach, 2009; Sylvén & Ohlander, 2015; Xanthou, 2011), writing competence (Jexenkicker & Dalton-Puffer, 2010; Llach, 2009), oral narrative competence and communicative strategies (Hüttner & Rieder-Bünemann, 2010), as well as academic reading proficiency (Hellekjær, 2008) and listening skills (Aguilar & Rodríguez, 2012).

In contrast with the large volume of research on linguistic achievement, there appears to be a disjuncture in the literature on content outcome due to the absence of content specialists in CLIL research (Mehisto, 2008). As noted by Nikula, Dalton-Puffer, Llinares, and Lorenzo (2016, p. 1), "even though the duality of content and language lies at the heart of CLIL and has been acknowledged in most studies, CLIL research to date has been mainly characterized by language learning perspectives on learners' general language skills". Besides the imbalance in the research body, findings on content achievement are less conclusive than that of language outcome (Cenoz, 2013; Dallinger et al., 2016).

In quantitative studies of content achievement, comparisons are mainly made between CLIL learners whose content courses are (partially) taught in L2 and non-CLIL

peers in L1. It is a common concern of stakeholders that learning content subjects in an imperfectly known language will affect content understanding and knowledge construction (Dalton-Puffer, 2011; Ruiz de Zarobe, 2015). Therefore, it is expected by CLIL stakeholders that the “gains in content and language learning are balanced and never at the expense of each other” (Ortega, 2015, p. 1). Yet out of stakeholders’ anticipation, quantitative research outcomes appear to be contradictory. On the positive side, earlier findings illustrate CLIL advantages in content attainment even when CLIL learners are tested in L1, as is shown in the study of Belgian primary school learners in mathematics (Van de Craen, Mondt, et al., 2007) and other studies in secondary and tertiary education (Burston & Kyprianou, 2009; Vollmer, Heine, Troschke, Coetzee, & Küttel, 2006). However, in recent years, positive voices have been undercut by studies reporting reduced progress. In a longitudinal study of 1806 German secondary school learners’ performance in history, Dallinger et al. (2016, p. 23) found CLIL-classrooms needed substantially more time to achieve comparable content outcomes than their non-CLIL peers. Similar discouraging results were revealed in the test scores of 294 French university students, with varied CLIL languages of English and German, and the content domains of Law and Computer Science (Roussel & Tricot, 2017). Positioned between positive and negative outcomes were studies reporting little or no significant difference in content attainment between CLIL learners and their peer controls (Admiraal, Westhoff, & de Bot, 2006; Seikkula-Leino, 2007). One well-cited example is a Finnish study of 669 mathematics learners in three age groups (Jäppinen, 2005), in which weak negative effects were found in the younger group (age 7-9), slightly positive for the middle group (10-12), and zero impact for the older (13-15). These neutral results, as noted by Ioannou Georgiou (2012, p. 501), have been considered by some researchers as a winning of CLIL because CLIL programmes can “provide the same level of education and achievement in content (in L2) as would L1 instruction”.

The quantitative studies make contributions to CLIL literature. Firstly, most

similar research attempts to seek causal explanations between CLIL exposure and learner attainment in (quasi)experiments and emphasizes quantitative rigour of analysis. Findings from quantitative research entail outsiders' bird's eye view of the CLIL effect, which are frequently regarded by stakeholders, especially policymakers and educational authorities, with higher credibility and are more generalizable to other contexts (Bonnet, 2012). Secondly, the quantitative research method is relatively quick to conduct and cheap to administer, making large-scale and representative sampling possible. CLIL and non-CLIL subgroups can be extensively compared for outcome evaluation within large samples, providing aggregated snapshots of CLIL effectiveness for quantitative considerations. One seminal example is the large-scale longitudinal CLISS project of Swedish upper-secondary CLIL learners from 2011 to 2014. The overarching aim of the study was to "provide as multifaceted a picture of CLIL in Sweden as possible", including CLIL effectiveness on vocabulary, motivation, and identity (Sylvén & Ohlander, 2015, p. 81). Thirdly, with carefully manipulated conditions in quantitative research, a number of confounding variables might be held constant, allowing researchers more credibly assess CLIL attainments (Dallinger et al., 2016). In a randomized controlled field experiment with three measurement points (pre-, post-, follow-up), thirty regular sixth-grade classes (N=722 German students) were randomly assigned to CLIL/non-CLIL treatment to control for potential pre-existing differences like academic performance and motivation due to "self-selection of applicants and school selection of participants" (Piesche, Jonkmann, Fiege, & Keßler, 2016, p. 108). This result showed that CLIL learners' gaining in physics were smaller than their non-CLIL counterparts immediately after the intervention and at follow-up six weeks later (Piesche et al., 2016, p. 108). In general, quantitative studies tend to provide stakeholders with suggestions of good CLIL practice and analyses of the critical factors that may inform whether/how CLIL programmes should be conducted. As commented by Ana Llinares (2015, p. 60), quantitative studies "have been key in enhancing both the implementation and improvement of CLIL programmes, as well as in stimulating further research in CLIL".

However, criticisms of quantitative studies are augmenting, doubting that some research findings might be biased by methodological flaws (Bruton, 2011; Dallinger et al., 2016). It has been reported in some quantitative studies that confirmation bias might occur owing to the narrow and simplistic focus on testing the unidirectional causal relationship between CLIL exposure and learner attainment. Instead, a number of confounding factors have been identified intervening CLIL learning outcomes, including extra exposure to the CLIL language than non-CLIL learners, higher motivation among CLIL participants, and high drop-out rate from CLIL programmes (Broca, 2016; Bruton, 2015). Rumlich's (2014) longitudinal, quasi-experimental study of prospective CLIL learners is a noteworthy example, in which 968 German sixth-grade CLIL learners have been found to outperform their non-CLIL counterparts even before CLIL learning. It is argued by Rumlich (2014) that this has not been acknowledged in previous studies showing the benefits of CLIL. Nevertheless, even if the confounding factors were ironed out, the claims for or against a CLIL programme might ring hollow by merely hinging on quantitative level studies, without knowing what occurs inside the classroom. According to Bonnet (2012), quantitative studies are generally liable to overlook the context-sensitivity of diverse CLIL models and the question of how/why CLIL influences educational consequences in complex classrooms. The data collected and analysed in quantitative studies are generally "abstracted indices", which might be valuable for descriptive and comparative purposes, yet are distant from classroom activities in which the learning of content and language are organised and enacted (Leung, 2005, p. 243). In other words, most large-scale quantitative CLIL studies are conducted for state considerations, which may not apply to inform classroom practice at local realisations (Moate, 2011). Under many circumstances, the advantage of seeking causality and collecting aggregated, quantitative data in quantitative CLIL research is taken at the cost of the ability to portray qualitative variations.

2.2.2 Deep insights of CLIL attainments in qualitative studies

From practitioners' perspective, CLIL is perceived as an idiosyncratic phenomenon possessing "crucial features that can only be elucidated by taking into account the idea of agency", which means the CLIL community possesses insights into the consequences of classroom practice that "transcend external boundaries and transform structures and systems" (Bonnet, 2012, p. 68). In line with this perspective, CLIL achievement is regarded as "a deep structure" that influences a CLIL learner's actions across situations and is related to his/her "emotional, volitional and reflexive dispositions" (Bonnet, 2012, p. 72). The practitioners' perspective of CLIL research is largely informed by the anti-positivist philosophical thinking, which mainly calls for qualitative approaches to evaluate CLIL effectiveness based on participants' viewpoints and interpretation of how CLIL affects learner outcome through text data (Bonnet, 2012, p. 68). In this way, a relatively small-scale qualitative naturalistic research design is generally preferred to examine diverse variables mediating CLIL education and its consequence in specific contexts (Dalton-Puffer, 2011; Papaja, 2012).

In contrary to quantitative studies with prevalently positive conclusions in language achievement and generally elusive results in content attainment, qualitative research presents more inconsistent outcomes in both language and content due to the research focus on individual learner's idiosyncratic details in the CLIL learning process (Dalton-Puffer, 2011; Ruiz de Zarobe & Cenoz, 2015; Ullmann, 1999).

With deep insights of individual CLIL participant's perceptions and experiences entailed by qualitative studies, an apparent mismatch has been identified between receptive and productive language performances, unveiling more gains in receptive competence among CLIL learners (Aguilar & Rodríguez, 2012). In relation to listening and reading achievements, research findings appear to be more definite (Dalton-Puffer & Smit, 2013; Massler, 2012; Papaja, 2012). Generally positive CLIL effects on reading attainments have been typically demonstrated by opinion questionnaires and interviews

of learner perceptions in a study of science and arts CLIL classes in a Spanish primary school (Pladevall-Ballester, 2015). This finding has been corroborated by teachers' feedback in the same study (Pladevall-Ballester, 2015) and other evidence from self-observation-inspections of 33 British secondary school teachers (Coonan, 2007). Moving on to listening, meeting and interview results of 17 CLIL lecturers at a Spanish university valued listening gains as the most positive among the four language skills, supported by a questionnaire survey of CLIL learners in the same study (Aguilar & Rodríguez, 2012). Similar reports have been documented in other studies as well (Aguilar & Muñoz, 2014; Pladevall-Ballester, 2015).

In contrast with receptive skills, findings on productive competences are generally inconsistent in qualitative studies. In relation to writing, an open-ended questionnaire survey of 87 CLIL learners from nine European countries in a Spanish university valued positive effects on writing, but reported lower achievement than other skills (Aguilar & Rodríguez, 2012). Similar results have been found in other studies in both European and Asian contexts (Guo & Wu, 2013; Järvinen, 2010; Pérez-Vidal & Roquet, 2015). However, voices of unfavourable effects emerge in these contexts as well. In W. Yang & Gosling's (2014) interviews of 78 Taiwanese university participants, no improvement in writing skills was reported. In a similar vein, findings of marginal progress (Dallinger et al., 2016; W. Yang & Gosling, 2014) in speaking proficiency have been reported among positive evidence (Admiraal et al., 2006; Dalton-Puffer, 2008; Gallardo del Puerto, Gomez Lacabex, & García Lecumberri, 2009; Pérez-Vidal & Roquet, 2015). In view of the findings of previous studies, it can be seen that qualitative studies entail a more discrepant picture of CLIL effect on language outcomes than quantitative ones.

Regarding content attainment, findings from the few qualitative studies are not as encouraging as their quantitative counterparts. The number of qualitative studies suggesting positive CLIL effects is even fewer. With the think-aloud method triangulated by written answers and interviews of 13 grade-ten German CLIL learners

and seven peer controls, Heine (2010) found that CLIL participants had a deeper understanding of geographic content knowledge due to additional semantic language learning. Surmont et al. (2014) concur that CLIL learners could better comprehend subject content by developing metalinguistic cognition. Compared with positive reports, relatively more studies find CLIL exposure makes marginal or no detriment to learners' content subject performance (Bonnet, 2012). By comparing ten Swiss primary school CLIL learners and an equal number of Non-CLIL counterparts, neither positive nor negative CLIL impacts were reported on learners' content achievement in classroom-observation and oral subject-knowledge interviews (Badertscher & Bieri, 2009). This finding was corroborated by meeting and interview results of 17 Spanish university lecturers, though their students expressed slightly more negative perceptions on content performance in the study (Aguilar & Rodríguez, 2012). These neutral findings, similar to that of quantitative studies, have been regarded as a CLIL advantage because "learners can produce equally good results even if they studied the content in an imperfectly known language" (Dalton-Puffer, 2011, p. 189). However, these positive and neutral studies have been relativized by negative ones arguing the quality of the content might be sacrificed in the CLIL context. In a series study of Sweden college CLIL learners (Airey, 2009; Airey & Linder, 2006), interview results showed that CLIL learners' understanding of academic concepts might be impaired. In a similar vein, an open-ended question survey in two Spanish universities revealed that both CLIL teachers and learners perceived a slower delivery rate in content (with a slight reduction of content) as a requisite of future CLIL implementation (Dafouz, Núñez, Sancho, & Foran, 2007). Besides, learners were found more cautious of their content knowledge acquisition than teachers. A noteworthy phenomenon among qualitative studies is that teachers' perceptions of content attainment are generally more positive than learners' (Coonan, 2007).

Qualitative CLIL research generally holds its own methodological advantages and attenuates some methodological flaws in quantitative studies. Firstly, the data from

participants' perceptions in qualitative research entail researchers with a deep and rich understanding of CLIL outcomes rather than academic grades (Infante et al., 2009). Researchers can take deep observations of a wide range of variables that might elude in quantitative studies, such as participants' developmental curve in emotion, attitude, and motivation. The provision of insiders' view in the words and categories of participants enables researchers to scrutinise how individuals vary discrepantly in response to CLIL exposure in certain contexts and under certain conditions. For instance, how learners with different L2 proficiency, content knowledge or motivational status gain different educational consequences in CLIL settings (Hou, 2013). In Pladevall-Ballester's (2015) study of Spanish primary CLIL programmes, the stakeholders' perceptions collected in the open-ended questions helped portray learners' diverse developmental trajectory in academic achievement, motivation and attitude toward CLIL. Secondly, small-scale qualitative naturalistic studies commonly employed from practitioners' perspectives enable researchers to explore how CLIL effectiveness occurs/fails in local conditions. With detailed charting of CLIL participants' progressive patterns and perceptions throughout the research, some confounding factors in quantitative-level studies might be attenuated. One example is the high rate of CLIL sample mortality or attrition. As identified in a number of CLIL studies, it might diminish initial representativeness and affect research validity (Admiraal et al., 2006; Apsel, 2012; Ruiz de Zarobe, 2008, 2010). Apsel's (2012, p. 49) episodic/narrative interviews of dropouts from CLIL grammar schools in Germany might facilitate addressing this confounding factor which leads to "silent selection".

However, qualitative studies are not immune to methodological flaws. What is widely criticised in qualitative research is the naturalistic design, in which the equivalence of CLIL learners and non-CLIL peers in prior proficiency cannot be strictly realised, and some confounding variables might not be well controlled. Studies have been documented disclosing CLIL learners with higher initial proficiency than their peer controls in content and language, and other contextual factors like motivation and

family background, which makes CLIL and non-CLIL subgroups not validly comparable (Bruton, 2013, 2015; Dallinger et al., 2016). In addition, most qualitative research is more prone to qualitative studies that are relatively time-consuming and costly to conduct. The data yielded are more at an individual level, which is complex to analyse and more easily influenced by researchers' idiosyncrasies, resulting in relatively lower credibility for quantitative consideration. There are cases where teachers have more incentive to engage in the new educational experience of CLIL, privileging CLIL learners in different ways and showing more positive perceptions from teachers' perspective than learners' (Coyle et al., 2010; Pérez-Cañado, 2020). Moreover, control effects have been reported in some qualitative research, in which repeated interviews, questionnaires, or observations might cause CLIL learners' sensitisation and change in behaviours, eliciting confounding effects in subsequent data collection and biasing internal validity (Bruton, 2011). It can be contended that qualitative CLIL research possibly attenuates certain methodological flaws in some quantitative studies but might cause new research weaknesses in validating CLIL learning consequences, adding potential obstacles to extrapolate research findings from one context to another (Lorenzo, Casal, & Moore, 2010; Navés & Victori, 2010; Pérez-Cañado, 2012).

As aforementioned, both quantitative and qualitative perspectives of CLIL research have methodological strengths and weaknesses respectively, and provide incomplete explanations of CLIL learning outcomes. Therefore, an integrated perspective is needed for "complementary strengths and non-overlapping weaknesses" (Johnson & Onwuegbuzie, 2004, p. 18).

2.2.3 Investigating CLIL effects from the integrated perspective

As discussed in the review of empirical studies in Sections 2.2.1 and 2.2.2, both the policymakers' perspective and their practitioners' counterpart are inclined to provide incomplete explanations of CLIL effects. In view of the divisive dualisms in

investigating CLIL effects, some researchers attempt to draw from the strengths and minimize the weaknesses of both in a third perspective, namely the integrated perspective (Karimi, Lotfi, & Biria, 2017; Kirsch, 2016; Pérez-Cañado, 2017; W. Yang, 2014).

From the integrated perspective, CLIL is concomitantly an individual and societal phenomenon (Li, 2008). It is a multi-faceted complex occurrence whose complexity is strengthened by taking into account the contextual factors in diverse settings (Chaplier & O'Connell, 2015). CLIL learners' achievement is "a commitment to a combination of language fluency and content accuracy", which can be evaluated from both quantitative assessments of learners' learning outcomes and qualitative interpretations of their viewpoints (Muñoz-Luna, 2014, p. 168). In this sense, the integrated perspective is largely informed by a pragmatic philosophical stance, which calls for a mixed research methodology. Different from the dualistic and contradictory policymakers' and practitioners' perspectives, researchers from an integrated perspective hold an eclectic and pluralistic stance and actively invite the two previous perspectives to participate in the dialogue. The metaphysical disputes on whether CLIL is a natural existence or an emergent social world has been compromised. In relation to CLIL research, attempts have been made to fit together the insights provided by both policymakers' and practitioners' perspectives, together with quantitative and qualitative research methods into a workable solution of canvassing CLIL effects (Karimi et al., 2017; Kirsch, 2016; Pérez-Cañado, 2017; W. Yang, 2014).

Different from its two antecedent counterparts, the integrated perspective legitimizes the employment of both quantitative and qualitative methods in CLIL research, making it possible to "follow the research questions in a way that offers the best chance to obtain useful answers" (Johnson & Onwuegbuzie, 2004, p. 17). Mixed methods research in CLIL enables the researcher to obtain both aggregated data and individual information for policymakers' and practitioners' considerations. With the

quantitative data, causal relations between CLIL exposure and educational consequences might be sought to inform policy. Meanwhile, the qualitative information enables CLIL practitioners to reflect on how CLIL affects individual learners in the classroom. When the two types of data are corroborated, greater credence can be gained to the singular conclusion of CLIL attainment.

In contrast with the vast body of research from policymakers' and practitioners' perspectives, the number of studies from the integrated perspective is considerably limited (Gierlinger, 2015; Sandberg, 2015; W. Yang, 2014). However, there is a growing volume of empirical research from the integrated perspective, among which two typical examples would be raised to demonstrate its methodological strengths. In Wiesemes' (2009) study of British primary and secondary school CLIL learners, quantitative test results showed the causal relation between CLIL exposure and positive educational consequence on learners' French and German performance, while qualitative data from classroom observations and semi-structured interviews of participants complemented the causal relation with deep and rich insights. In a different situation from Wiesemes' (2009) study, when the quantitative and qualitative findings conflict, further research will be conducted, providing greater knowledge of CLIL effectiveness. In Yang's (2014) study of Taiwanese university learners, findings from the semi-structured questionnaires revealed negative effects on productive language skills, contrary to their test performance. The conflict between learner perceptions and test scores aroused further research to deepen the understanding of CLIL effects on target learners. In view of the triangulation of different types of data in the two examples, it might be argued that CLIL research is given more possibilities to increase the conclusiveness of findings with the integrated perspective and mixed methods research.

Despite the strengths mentioned above, studies investigating CLIL effects from the integrated perspective are not immune from criticism (Bruton, 2011, 2013; Paran, 2013). Most of the criticism targets controversies in research designs rather than

methodological flaws. One chief concern is the lack of homogeneity of research samples (Pérez-Cañado, 2017). As noted by both CLIL advocates and sceptics (Alonso, 2017; Bruton, 2011, 2013; Paran, 2013; Pérez-Cañado, 2017), the comparability of CLIL and non-CLIL strands has not been strictly guaranteed in a high percentage of CLIL studies. A number of potential intervening variables impinging on CLIL learning outcomes have generally not been factored in and matched in pre-tests and post-tests between experimental and control groups, such as learners' initial FL proficiency and content knowledge, intramural and extramural exposure to the CLIL language, motivation, sociocultural studies, socioeconomic status, class size, types of school, or the linguistic competence of the teacher. These variables have been widely found conducive to bias and discrimination against non-CLIL groups (Dallinger et al., 2016). For instance, Bruton (2015, p. 124) cited the article by Apsel (2012) that in the German context, "CLIL students might receive two extra hours of English for two years before that start, and an hour extra once in CLIL". It is not rare in practice that CLIL classes are reinforced with additional hours in the vehicular language (Coyle, 2013; Dalton-Puffer, 2011). Under such compromised conditions, it is difficult to determine whether the difference between treatment and comparison groups is ascribable to the independent variable of CLIL or other intervening variables like extended exposure to the vehicular language. Without ensuring the homogeneity of these variables between CLIL participants and their controls from the outset, the research on the CLIL effects might be skewed or invalidated. However, in naturalistic contexts of CLIL implementation, as noted by Pérez-Cañado (2017, p. 86), "it is extremely problematic to find homogeneous treatment (CLIL) and comparison (non-CLIL) groups" in already existing classes. In response, the randomisation of samples, namely the use of randomized controlled trials (RCT), has been taken into account in this study to "produce equivalence over a whole range of variables" rather than a few named variables between the experimental and control groups (Cohen, Manion, & Morrison, 2017, p. 277).

Besides the homogeneity of samples, other disputes have been raised in some integrated studies as well, including the neglect of large-scale standard tests on language and content achievement for better generalisability (Moghadam & Fatemipour, 2014; Paran, 2013), the unconcern of measurement of CLIL learners' intangible achievements like culture and cognition (Coyle, 2007; Papadopoulos & Griva, 2014; Sudhoff, 2010), the over-reliance on the teacher's appreciations rather than statistical confirmation (Bruton, 2015; Pérez-Cañado, 2017), the disregard of multiple triangulation between quantitative and qualitative data in the same study (Pérez-Cañado, 2016; W. Yang, 2015), the lack of investigating stakeholder perception on the way CLIL programmes work out (Pladevall-Ballester, 2015; W. Yang, 2015). These healthy controversies and disputes provide significant implications and lay the foundation for a more unbiased, disinterested, and methodologically sound research design adopted in the current study.

Summary

CLIL is “not just a new expression of educational bilingualism” (Lorenzo, 2007, p. 27). It is a well-defined and well-established educational approach (Pérez-Cañado, 2017). In view of the articulation of CLIL conceptualisation and empirical review in this chapter, studies from policymakers' or practitioners' perspectives tend to provide an incomplete portrayal of defining CLIL and evaluating its effectiveness. The integrated perspective has breathed life into the arena of CLIL research and provided a broader view unfettered by its two preceding contradictory perspectives. It has allowed researchers to establish “a taxonomy of different common forms of CLIL so as to circumscribe the diverse contexts in which CLIL is found” (Pérez-Cañado, 2017, p. 86). In other words, as to the research on CLIL effectiveness, the integrated perspective allows researchers to identify the methodological weaknesses in the studies of quantitative/qualitative perspectives and overcome them with the methodological strengths from each other.

However, an appropriate concept and research perspective are not enough for CLIL to earn legitimacy in its future action and research agenda. A more holistic approach is needed, which might be more accurate or superior to understand the nature of CLIL and fit together the insights provided by both into a workable solution and produce a satisfactory product. In the following chapter, the rigour and soundness of its theoretical framework have to be sufficiently justified for stakeholders' investment in time, effort and resources.

CHAPTER THREE- 4Cs Framework of CLIL research

As discussed at the end of the last chapter, based on the integrated perspective of CLIL research, a holistic approach is called for to allow researchers to establish “a taxonomy of different common forms of CLIL so as to circumscribe the diverse contexts in which CLIL is found” (Pérez-Cañado, 2017, p. 86). Thus, Coyle’s 4Cs framework is introduced and subsequently elaborated, with some reinterpretations where necessary. This chapter aims to provide a holistic theoretical framework in the CLIL context that informs the present study.

3.1 The theoretical framework of CLIL

As illustrated in the preceding section, an integrated perspective might not be sufficient to theorize the integration of content and language learning in CLIL. It is foregrounded by Coyle (2007, p. 549) that “quality CLIL is dependent on understanding and operationalising approaches which will not be found solely in the traditional repertoires of either language teaching or subject teaching”. From a holistic perspective, Coyle (2007, p. 550) further elaborates that content learning in CLIL is placed in the “knowledge for learning” domain, involving “content and cognition”; whereas the “culture-bound phenomenon” of language learning is regarded as “a medium for learning”, in which “communication (language) and intercultural understanding” are included. This articulation contributes to the literature by developing the two main foci of CLIL into “one integrated learning process” with four interrelated facets of content, communication, cognition, and culture (Dalton-Puffer, 2011, p. 196). In the same vein, the four facets interplay at diverse dimensions within a specific context and converge to establish the 4Cs framework to illustrate a quality CLIL learning process from a holistic view (Coyle et al., 2010, p. 41). As shown in Figure 4, culture permeates throughout the three facets of content, communication, and cognition. As to Coyle et al. (2010), effective CLIL takes place through the progression in understanding of content, engagement of cognition, interaction in the communicative context, improvement of

language learning and using, development of intercultural awareness, as well as the consideration of contextual variables in the wider educational context in which CLIL is embedded.

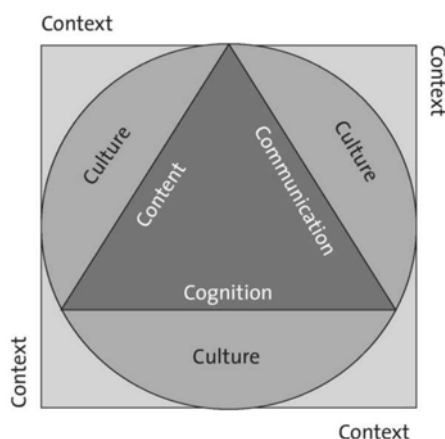


Figure 3.1 The 4Cs framework (Coyle et al., 2010: 41)

The 4Cs framework has been perceived not only as the theoretical basis of CLIL, but also as a practical tool to plan CLIL lessons in diverse contexts, linking theory and practice (Wiesemes, 2009). As Coyle et al. (2010, p. 115) note, “CLIL units will all contain clear objectives, possibly fashioned around the 4Cs”. Therefore, the 4Cs will be elaborated as the objectives or competences this CLIL programme work towards, with some reinterpretations where necessary. Further articulation will be made in the subsequent sections.

3.1.1 Content:

In both CLIL research and practice, there is a trend toward conceptualising content as merely a vehicle for language learning (Cenoz, 2013; Ruiz de Zarobe & Cenoz, 2015). It is claimed that non-linguistic subject matter in CLIL is dictated by learners’ linguistic demands, resulting in simplified or sacrificed content learning for fear of learners’ low L2 proficiency (Bruton, 2011). This misunderstanding of CLIL content has aroused contestations and malpractice in diverse contexts (McDougald, 2016). Coyle (2007) notes that teaching subject content in a target foreign language does not automatically lead to CLIL. Unlike typical language classrooms, “language is neither

the designated subject nor the content of interaction” in CLIL classrooms (Dalton-Puffer, 2007, p. 3).

Contrary to these misunderstandings, placed in the centrality of CLIL is the social-constructivist belief that learners construct their new knowledge upon their current knowledge state (Bruner, 2009; Coyle, 2007; Dalton-Puffer et al., 2021). In accordance with the defining characteristics in Section 2.1.2, CLIL is fundamentally a subject content course in the curriculum. Content learning lies at the heart of the CLIL learning process (Coyle et al., 2010; Dalton-Puffer, 2011). It shares contemporary educational theories with its mainstream counterparts, a foundational one of which is the social-constructivist approach (Vygotsky, 1980). According to Coyle et al. (2010, p. 28), the traditionally dominant learning model in many societies is a teacher-controlled and teacher-led “banking model”, in which new knowledge and skills are believed to be deposited by the expert (the instructor) into the memory bank of the novice (the learner) (Bruner, 1966, 2009). Alternatively, the social-constructivist approach perceives learning as an active process in which learners form new knowledge based on their current knowledge state (Cummins, 2005). According to Dalton-Puffer (2007, p. 7), learners’ learning process is largely devoted to ordering processes in which learners organize and develop new experiences and mental models in terms of their previous ones. In the learning process, learners rely largely on their extant cognitive structures to select and transform new information obtained, invest immediate experience with meaning, and integrate the new information into their knowledge system in an organized way. In this way, the social-constructivist approach brings into CLIL the centrality of learners’ prior experience, active cognitive engagement, and their own interpretation of content knowledge, rather than passive reception of information transmission from teachers, as frequently criticized in some teacher-centred approaches (Meyer, 2010). In other words, content matter means more than learning knowledge and skills. It means “the learner constructing his/her own knowledge and developing skills which are relevant and appropriate” on their current state of knowledge and skills

(Coyle, 2008, p. 550).

The social-constructivist approach offers CLIL a student-centred and interactive learning scenario for effective content learning, which is demanded both cognitively and linguistically. This scenario foregrounds social interaction between learners and their scaffolded learning context, including teachers, peers and resources that are more “expert”. According to Dalton-Puffer (2007, p. 9), learners’ higher cognitive functions in content learning are not “ontogenetically conceivable”, but are realized by “individual’s interaction with the social environment”. When learners deal with new knowledge, they are cognitively challenged. They are likely to be engaged in interacting with the scaffolded learning context to develop their individual thinking. This type of learning is termed by Vygotsky (1980) as ZPD (Zone of Proximal Development), which is “always challenging yet potentially within reach of individual learners on condition that appropriate support, scaffolding and guidance are provided” (Coyle et al., 2010, p. 29). In the CLIL learning context, “meaningful, challenging and authentic input” from the subject content is necessary to make learners’ affective filters open for new knowledge and establish links to prior experiences (Meyer, 2010, p. 13). Alongside cognitive demands, the linguistic requirement is indispensable for learners to construct an understanding of content knowledge (Coyle, 2018). Learners’ interaction with the scaffolded learning context needs to be acted out via a powerful semiotic means, namely linguistic language. In CLIL settings, an imperfectly known L2 is used by learners for social interaction and communication with experts, peers, and other resources. The “linguistic demands of the content as the conduit for learning” needs to be taken into consideration and made accessible to learners (Coyle, 2008, p. 550). Thus, both cognition and language become the prerequisite for learners’ internalisation of knowledge, skills and understanding.

Besides cognitive and linguistic demands, it is theoretically illustrated in the 4Cs framework and empirically evidenced that contextual variables, including teacher

availability, social demands, and so on, might influence the choice of CLIL content (Moate, 2011). In this way, the broad expression of “content” in CLIL might be elaborated as cognitively and linguistically demanded subject learning depends on contextual variables.

3.1.2 Communication:

Along with content, the second facet that needs elucidation is communication. In the 4Cs Framework, Coyle et al. (2010, p. 42) interchangeably use the terms “language” and “communication”, which presumably arouses an unexpected consequence of reducing CLIL language learning to grammatical progression in some research and practice (Aguilar & Rodríguez, 2012). However, the adoption of “communication” in place of “language” is not only “a syntactical device for promoting the C concepts”, but also a strategy to foreground CLIL advantage over mainstream foreign language courses, namely the integration of learners’ linguistic competence and communicative competence (Coyle et al., 2010, p. 42).

A number of extant research outcomes have proven that merely learning content matter in an L2 does not naturally lead to language progress (Coyle & Meyer, 2021; Devos, 2016; S. P. Hughes & Madrid, 2020). According to the social-constructivist approach, CLIL learners’ content learning is realized by their interaction with the social environment via an imperfectly known CLIL language (Dalton-Puffer, 2007). That brings to the fore the preconditional role of learners’ linguistic competence in the vehicular language, without which CLIL cannot succeed. Recent advancements in second language acquisition research have underpinned language learning in CLIL contexts. In line with Krashen’s (1985) input hypothesis, learners’ language is acquired subconsciously with exposure to a sufficient dose of comprehensible input, especially when the situation is featured by positive affective emotions (low affective filter). In CLIL settings, the focus of learning shifts from linguistic knowledge to content knowledge. When learners receive content-related input from teachers, peers, and

materials in the CLIL language, the amount of authentic comprehensible input is largely increased, offering possibilities of enhancing receptive language skills. In this input process, learners are required to acquire the content-obligatory language that is essential for learning the content, such as content-specific vocabulary and content-related grammar elements like tenses and comparatives. For instance, in a science class on measuring insects, learners have to learn simple present forms of “to be” and comparatives of “longer” or “heavier” for better understanding and communication of content knowledge. This kind of content-related input is more related to learners’ linguistic competence. The learning of content in CLIL classes endows learners an authentic use of the vehicular language as it is used for real social, informative, expressive, and other purposes. This authenticity increases CLIL learners’ motivation than their counterparts who simulate real-life situations in mainstream foreign language classes. In addition, the primacy of meaning over form in CLIL classrooms is supposed to make learners’ language mistakes less penalized or corrected, reducing their language learning anxiety and affective filter (Smit & Dalton-Puffer, 2007). Besides input-based theories, SLA research into CLIL has made use of Swain’s (1995) output hypothesis to address the improvement of CLIL learners’ productive language competence. In CLIL lessons, content-related classroom activities and collaborations provide learners with sufficient authentic opportunities to produce comprehensible output voluntarily and naturally in the CLIL language, which “expands their active linguistic repertoire and achieves deeper entrenchment of what they already know” (Dalton-Puffer, 2011, p. 194).

In CLIL learners’ input and output processes, language achievement has not only been made in linguistic competence but communicative competence as well. As per the social-constructivist approach, the nature of language “is not a purely cognitive phenomenon” (Smit & Dalton-Puffer, 2007, p. 10). It is socially distributed and its acquisition is perceived to be “a process which is socially construed” (Smit & Dalton-Puffer, 2007, p. 10). In CLIL contexts, learners are supported in developing content-

related communicative strategies to operate in social settings like classrooms (Devos, 2016). This kind of communicative competence includes a repertoire of content-related speech acts like debating, inquiring, describing, clarifying, and collaborating, which “represent salient aspects of communication reaching beyond mere vocabulary and grammatical forms” and indicates a more sophisticated level of understanding about the CLIL language (Devos, 2016, p. 13). With these tools and strategies, CLIL learners are able to support or be supported by teachers and peers and accomplish classroom activities. This form of language concentrates on the language needed to operate in a foreign language setting like a CLIL class on a specific content subject. It is addressed by Dalton-Puffer (2007, p. 65) as “language for knowledge construction”, without which quality learning in CLIL will not occur. Besides the language for operational use in CLIL classes, Coyle et al. (2010, p. 68) propose a new form of language “to integrate cognitively demanding content with language learning and using” in CLIL, namely “language through learning”. This form of language is predicated on the social-constructivist tenet that effective learning may not occur without the active engagement of language and thinking (Vygotsky, 1980). In CLIL learning, “new meanings are likely to require new language” (Coyle et al., 2010, p. 37). When learners are encouraged to reflect and articulate their understanding of new content matter and language progress in CLIL class in social interactions with teachers and peers, a form of emerging language is discovered, or an already known language is rediscovered by learners themselves “to support and advance their thinking process”. In this way, “a deeper level of learning takes place” (Coyle et al., 2010, p. 38).

In view of the authentic input and output processes in CLIL language learning, it might be safely concluded that the role of language in CLIL settings has been reconceptualized, from learning the language per se, to using the language to learn, and to learning to use the language (Coyle et al., 2010). Different from the mainstream foreign language counterpart, in which “language is learned first and then used” (Cenoz, 2015, p. 17), the CLIL approach combines cognitively challenging content with

language learning and authentic language using in academic contexts. As a result, the CLIL approach enhances learners' linguistic and communicative competence. The concept of communication, as commented by Devos (2016, p. 14), "perhaps best underlines CLIL's distinctiveness in the 4Cs model".

Concerning the gauging of CLIL learners' progress in this "C concept", different language proficiency tests have been adopted in various settings to measure learners' linguistic development. In this study, a Chinese national standardized English proficiency test of CET-6 (College English Test Band 6) is adopted with the related rationale given in Section 4.3.1. As per CLIL learners' progress in communicative competence, it will be further articulated in the succeeding section with learners' attainment in cultural awareness in the cultural facet, which might further develop into learners' intercultural understanding.

3.1.3 Cognition:

In terms of the 4Cs framework, CLIL is intellectually demanding. Learners need to invest additional efforts to comprehend the concepts taught through the L2 when they construct content and language knowledge. This implies that learners need to learn how to apply the attained knowledge through higher-order thinking skills to achieve interaction with teachers and peers in the CLIL classroom (Coyle, 2018; Coyle et al., 2010), and accomplish problem-solving through different media and cultural spheres outside class (C. E. Anderson, 2011).

In the literature of CLIL studies, a number of researchers make efforts to illustrate CLIL learners' cognitive progress and attempt to provide explanations from different perspectives (Coyle et al., 2010; Lorenzo, 2016). The supporting research mainly focuses on CLIL advantage in cognitive control and selective attention (Heine, 2010; Poarch, 2013). With cognitive challenges from both content and language in CLIL learning, Poarch (2013) revealed in her study that CLIL learners had stronger capability

in cognitive control, which prevented their working memory from being overloaded and resulted in a more effective cognitive process. In addition, Heine (2010) found in her research that CLIL learners processed information in a deeper way because the achievement of understanding new content information in the CLIL language takes more cognitive endeavours. As a result, CLIL learners tended to show better long-term retention of the content (Heine, 2010). Despite the positive arguments, Kirschner, Sweller, and Clark (2006) suggest potential negative effects of CLIL from the perspective of cognitive load theory. It is claimed that with simultaneous processing in new content and the foreign language, the challenge might go beyond CLIL learners' working memory capacity, especially among those with relatively lower proficiency in the CLIL language.

These conflicting arguments on CLIL learners' cognitive development reveal that the conceptualisation of cognition in CLIL remains elusive, leading to confusion throughout CLIL practice (Lorenzo, 2016). Such elusiveness might be caused by the scarcity of related taxonomies and frameworks. In response, researchers attempt to integrate knowledge growth and the fostering of thinking and problem-solving skills into an integral learning process. For instance, Bloom (1956) proposed a 'Cognitive Process Dimension' outlining six different categories of thinking processes in the 'cognitive process' domain, namely Remembering, Understanding, Applying, Analysing, Evaluating, and Creating. These categories are ordered from simple to complex, and concrete to abstract (L. W. Anderson & Krathwohl, 2001). In this framework, the lower-order thinking covers remembering, understanding, and applying. The higher-order thinking includes analysing, evaluating and creating. Both the two types of thinking are integral to effective learning. Later, L. W. Anderson and Krathwohl (2001) updated Bloom's taxonomy by adding the 'Knowledge' Dimension to the 'Cognitive Process Dimension'. The knowledge dimension upholds a framework to explore the knowledge demands from different types and levels, covering factual, conceptual, procedural and metacognitive knowledge. Anderson and Krathwohl's

conceptual framework is more complex but logical as it categorises discrepant sorts of thinking in relation to varying kinds of knowledge construction. Thus, this framework is proposed by Coyle et al. (2010, p. 30) because “the transparent connecting of thinking processes to knowledge construction resonates with conceptualising content learning in the CLIL setting”. Within this framework, effective CLIL learning not only involves “the defined knowledge and skills in the curriculum or thematic plan”, but also “how to apply these through creative thinking, problem-solving and cognitive challenges” (Lorenzo, 2016, p. 34).

In short, aligning with Anderson and Krathwohl’s conceptual framework, the broad term of “cognition” in the 4Cs framework can be more related to investigating whether learners need to be cognitively demanded for the sake of achieving understanding, problem solving, information transforming, and discovering new meaning in this study.

3.1.4 Culture:

In the 4Cs framework of CLIL, culture remains to be the least explored but perhaps the most contentious “C” (Devos, 2016). Some CLIL researchers claim to replace the term “culture” with other “C”s, like “community” (Mehisto, Frigols, & Marsh, 2008). Coyle et al. (2010, p. 40) argue that this “forgotten C” is “not an option”, but “a necessity”, because culture permeates throughout the other three “C”s, namely content, cognition and communication.

Given the contestations above, it is essential to clarify what culture means and its role in CLIL. In view of the CLIL literature, the definition of culture is contestable and debateable. These definitions vary in different research domains in social science, eliciting inconsistency and confusion among researchers and practitioners. Thus, considerations of the interconnectedness among culture and other elements are inevitable in clarifying the concept of culture in CLIL. Brown (1980) provides a helpful

explanation linking culture with thinking and language. Brown (1980, p. 138) expounds that “cultural patterns, customs, and ways of life are expressed in language”, and “culture-specific world views (thinking) are reflected in language”. Meanwhile, world views differ among different cultures because language and culture interact with each other. The language employed to express the world views may be “relative and specific to that view” (Brown, 1980, p. 138). With regard to the learning contexts in which a CLIL language is employed, language, cultural awareness, cognitive involvement and thinking are all interrelated to the subject content within a specific setting. In CLIL settings where content and language are integrated, more opportunities are potentially provided to develop learners’ intercultural understanding than that in mainstream foreign language or L1 content classes. With the provision of a multilingual repertoire in CLIL settings, learners are assumed to foster greater awareness of their native language and its influence on thinking, widen their worldviews and bring about specific cognitive changes (Méndez García, 2013). In this sense, the learning of a CLIL language facilitates the realization of cross-cultural communication and the recognition of cultural differences. In CLIL, it is not only the linguistic level that raises intercultural learning processes. With the access to content topics from foreign perspectives and by making comparisons with native perspectives, CLIL learners are prompted to discover the diversities that people from different cultures see and understand each other. In this way, both content cultures and language cultures are fused in CLIL lessons, offering learners more intercultural experiences to deepen their understanding of native and foreign cultures. More intercultural experiences enrich learners’ cultural knowledge, skills and attitudes to “critically apply and analyse social processes and outcomes” (Coyle et al., 2010, p. 40). In this way, cultural awareness in the 4Cs framework of CLIL expands from knowledge about different cultures to intercultural understanding involving different experiences. In other words, CLIL learners’ attainment might transcend cultural awareness to intercultural understanding. With the integration of other Cs, CLIL might add learning value to learners’ intercultural understanding. As Meyer (2010) notes, intercultural understanding prepares learners for increased

communication and cooperation in globalised mobility and employability beyond geographic and cultural borders and is hence regarded as the ‘ultimate educational goal’ in contemporary globalised education.

In reality, the development of learners’ intercultural understanding in a globalised world has been promulgated and endorsed in a number of official documents in Europe, including the *Common European Framework of Reference for Languages* (European Council, 2011), the *Guide for the Development of Language Education Policies in Europe* (European Council, 2003), and the *European Agenda for Culture in a Globalising World* (European Commission, 2007). It is stated by Barroso (2007), the president of the European Commission, that intercultural understanding is an essential driver for personal development, social cohesion and economic growth. Culture is at the heart of EU policies to recognise and respect diversity. These documents reveal that intercultural understanding in CLIL may not be simply learned in a number of specific lessons, for example, about festivals, pop music, customs of a particular region and its native language. Intercultural understanding is developed in a social system where “language, thinking and culture are constructed through interaction” (Coyle et al., 2010, p. 72). From CLIL practitioners’ perspective, learners have to make meaningful interactivity with the teacher, classmates and other resources in and through the CLIL language in the class to foster their intercultural understanding. Whereas from policymakers’ perspective, intercultural learning needs to be extended outside the classroom and realized via social interaction and collaborative meaning-making (Coyle et al., 2010). Overall, CLIL facilitates learners’ intercultural understanding by developing “an ability to see and manage the relationship between themselves and their own cultural beliefs, behaviours and meanings, as expressed in a foreign language, and those of their interlocutors, expressed in the same language - or even a combination of languages” (Byram, 1997, p. 12).

In this study, culture is further expected to extend its potential of developing

learners' intercultural understanding in relation to CLIL learners' intercultural interactions and cognitive thinking, as articulated in the prior section. With regard to the facet of culture, it can be suggested that it is not expected to be a "forgotten C" in CLIL contexts, but rather a "potential C" to develop learners' intercultural understanding in this CLIL programme.

In sum, the four facets of CLIL, content, communication, cognition and culture, have been elaborated individually. However, the four facets do not exist as separate elements. They interrelate with each other and form a "symbiotic relationship" from a holistic view (Coyle et al., 2010, p. 41). It is the symbiosis that brings about quality CLIL learning.

3.2 The 4Cs framework of CLIL in the present study

As Coyle et al. (2010) note, it is complex to integrate the four facets of the 4Cs framework in CLIL research because the means and extent of integration is largely decided by individual learning contexts and projected learning outcomes. Thus, in the following section, the planning of a CLIL Advanced Financial Management (AFM) programme in a Chinese university is introduced to further illustrate the symbiosis of the 4Cs framework.

The first step to plan the target programme is establishing "a shared vision for CLIL" among teachers and programme managers (Coyle et al., 2010, p. 49). Since there is no tradition or few precedents in the target research context, involved professionals need to construct and own shared visions of this CLIL programme, such as develop spontaneous talk or improve finance-related correspondence writing. These shared visions are regarded as "a dynamic and iterative process" varying over time with reflection and review (Coyle et al., 2010, p. 51). However, the shared vision is not the only determinant of CLIL operating in the target research context. The contextual variables, including university types, teacher availability, learners' expectations,

regional and national policies, all contribute to determining the suitable model of CLIL for the target research context.

With regard to a CLIL class in detail, the four facets need to be connected into an integrated whole. In line with the elaborated 4Cs framework, CLIL learners are engaged in a learning process integrating knowledge of financial management as the target content and English as the target language. According to the conceptualisation of CLIL as mentioned above, “it is the content that initially guides the overall planning along the learning route” (Coyle et al., 2010, p. 55). For instance, in a lesson on writing a report on sales figures, both content and cognition are involved in the process of content learning. CLIL learners are required to develop skills of seeking and synthesising useful sales figures and construct knowledge of making related predictions on given data. It is likely that the teacher designs some warm-up activities like a briefing note or a video clip on similar topics, making learners access some key terminologies or other linguistic terms in advance.

Concerning language learning, “the emphasis is always on accessibility of language in order to learn” (Coyle et al., 2010, p. 55). The CLIL language itself is learned by reconstructing content themes with related cognitive processes. For instance, collaborative classroom activities are recommended to increase learners’ access to the content language and improve their communicative competence to fulfil tasks in the CLIL language. Such activities include group discussion on charting the development curves of sales and making predictions, which demands learners to shift properly between past tense for description and future tense for prediction and follow the discourse norms in reporting sales figures.

With regard to culture, the content of financial management provides new learning scenarios different from mainstream L1 content and foreign language classes, in which content cultures from native and foreign perspectives, native language culture, and

target language culture are fused to help learners foster intercultural understanding of “self” and “otherness”. CLIL learners might be demanded to participate in a role-play, exploring the different business norms of presenting sales figures and making predictions to a British manager and a Chinese counterpart respectively, for cultural contrast. It is expected that in combination with the development in other facets, learners might establish a new cultural space of “thirdness”, and intercultural communicative competence will be further developed to relativise and shift between different cultural perspectives effectively and appropriately.

Unlike mainstream content or foreign language courses, the CLIL financial management course needs considerate arrangement in all the four facets. The four facets progress at discrepant rates hinging on the circumstances. In line with the 4Cs framework as illustrated in Figure 3.1, the CLIL learning process is situated and developed in the broader educational context. Contextual factors like teacher availability, learner expectation, national standardised tests, and national policies need to be taken into consideration for improved CLIL effectiveness, which will be detailed in the articulation of research design in the following chapter (Coyle et al., 2010).

As reflected by the explanation of the 4Cs framework from a holistic view, the four elaborated facets provide more clarified objectives or competences for CLIL programmes to achieve. In this study, the effectiveness of the CLIL approach has been scrutinised in a Chinese tertiary context. The evaluation has been mainly focused on the four facets as well, namely, subject content knowledge, cognitive thinking skills, linguistic competence, and intercultural understanding. The explanation of the 4Cs framework caters to be the driving force behind the methodological choices in this study.

3.3 Aims and research questions

As reflected by the conceptualisation, the empirical review, and the elaborated theoretical framework in previous chapters, CLIL is a complex multi-faceted

phenomenon that integrates policymakers and practitioners' perspectives (Cenoz, 2015; Dalton-Puffer et al., 2010). Different perspectives bring CLIL different research foci and methodologies, analysing various components of CLIL. The complexity is further intensified by taking into account a variety of contextual variables in diversified settings.

Owing to the complex nature of CLIL, research on its effectiveness has been left far behind practice. An overwhelming proportion of extant studies are in the area of measuring learners' content and linguistic gains from policymakers' or practitioners' perspectives, whereas other relatively intangible achievements have been gravely underrepresented (Dalton-Puffer, 2011; Ruiz de Zarobe, 2015). Even in Europe, the birthplace of CLIL, the situation is not entirely satisfactory. As commented by Pérez-Cañado (2012, p. 315), "while at first blush it might seem that outcome-oriented investigations into CLIL effects abound throughout Europe, there is still a well-documented paucity of research in this area". The massive and rapid uptake of CLIL "has not been supported by a comparable level of research" from a more holistic perspective, which has given rise to fears among policymakers and practitioners, as well as the disjuncture between the academic world and classroom practice (Mehisto, 2008; W. Yang, 2014). Therefore, to add to the literature of CLIL effects from a more holistic perspective, this study attempts to investigate CLIL effectiveness in facilitating learners' academic performance from an integrated perspective based on the aforementioned conceptualisation, empirical review, and the elaborated theoretical framework.

CLIL remains a relatively uncharted territory to the Chinese context, compared to the significant surge of practice and research worldwide (R. N. Wei & Feng, 2015; W. Yang, 2014). CLIL has so far received little attention from mainstream researchers in the country. Published research on CLIL is limited in China, let alone in influential academic journals and professional conferences. Most of the sporadic research remains centred on introducing the CLIL concept and evaluating feasibility from theoretical

perspectives (Luo, 2006; H. W. Yang & Xu, 2011), while merely a small number of empirical research has been documented (Lei & Hu, 2014). Even among the documented studies, a high percentage of them have been found reporting CLIL research contexts and procedures in a simplified way, adding complication to extrapolate related research findings and replicate effective experience (A. Llinares & McCabe, 2020). Concerning the target context of Chinese higher education, the situation is more elusive. From policymakers' perspective, no official documents have been promulgated to endorse CLIL implementation, while little classroom praxis has been documented from practitioners' initiatives. The scarcity of hard evidence on CLIL effectiveness among Chinese college learners has triggered doubt and confusion among potential stakeholders, hindering its expansion in the nation with the largest population of L2 English learners (Feng, 2009; Zhao & Wang, 2014).

To address the aforementioned gaps, this study aims to focus on the following research questions (RQs):

Main RQ: To what extent does CLIL facilitate Chinese college learners' development of the desired academic performance?

Given the status quo of CLIL in the target research context of Chinese higher education as depicted, the main RQ first aims to verify whether CLIL improves target learners' academic performance in a holistic way to convince key stakeholders. The second aim of the study is to reveal how the desired CLIL effectiveness has been realized among target learners. According to the multi-faceted nature of CLIL and the elaborated 4Cs framework as articulated in previous chapters, the main RQ is divided into the following sub RQs to address the four 'C's, including academic performance in subject content, linguistic and communicative competence, intellectually cognitive proficiency, and intercultural understanding, as listed below:

Sub RQs:

- 1) To what extent does CLIL improve learners' academic performance in subject content?
- 2) To what extent does CLIL enhance learners' linguistic and communicative competence?
- 3) To what extent does CLIL increase learners' intellectually cognitive proficiency?
- 4) To what extent does CLIL facilitate learners' intercultural understanding?

As can be seen from the main RQ and the four sub RQs, there are two parts to each research question. The first part of each RQ is more explanatory, mainly resorts to quantitative assessments of learners' learning outcomes, while the second part is more exploratory, mainly turns to qualitative interpretations of participants' perceptions. The RQs serve as guidelines for the research design and research instruments in the current study, namely a mixed research methodology with a pre-test-post-test control and experimental group design. The detailed research procedure and rationale is addressed in the following chapter.

Summary

In sum, responding to the need for more holistic frameworks in CLIL research, the 4Cs theoretical framework was introduced and analysed in the present study (Coyle, Hood, & Marsh, 2010). By fusing content and language learning in a given context and recognising the symbiotic relation among the four contextualised building blocks of content, communication, cognition and culture, the 4Cs framework entails researchers to capture the multi-faceted complexity of CLIL learning and highlight the significance of CLIL in the overall knowledge learning process. Thus, a study informed by the 4Cs framework is designed and conducted to investigate CLIL effects on learners' academic performance in a Chinese tertiary education programme.

The next chapter is an overview of the methodology selected for this research study. Based on the theoretical framework adopted in this study, the chapter also reveals the underpinning philosophical stance, the research design, and the rationale for selecting a mixed-method research approach, together with details of the research site, participants, ethical issues, data collection process, data management and data validation for the study.

CHAPTER FOUR --Research Methodology

This chapter aims to describe the methodological issues concerning the current research topic and questions. As a starter, a general discussion of the research paradigm in the present study is provided in the first section. By evaluating and comparing the strengths and weaknesses of two traditional paradigms, namely positivism and interpretivism, a new paradigm named pragmatism is taken in the study to entail an immediate and useful middle place philosophically in the CLIL research, informing the mixed-methods research approach. The rationale of adopting a mixed-methods approach is mainly based on the premise that a broad range of sources of data could present a comprehensive and rich picture of CLIL learning among Chinese college learners, because the collection and analysis of both quantitative and qualitative data complement and benefit each other. The second section of this chapter covers the description of the detailed research design, which serves as a research tool for understanding the “human affairs” in the study (Stake, 1995, p. 10). Regarding the research procedure, a mixed research methodology with a pre-test-post-test control and experimental group design is introduced. The following section is an overview of the research “bound system” for this study (John W. Creswell & Miller, 2000), including the research context, participants, together with considerations of related ethical issues related. The last section outlines the process of data collection, management, and analysis, providing potential evidence for the research questions in the present study.

4.1 The research paradigm

In the light of the literature review of CLIL studies in Chapter Two, together with considerations of the elaborated 4Cs theoretical framework as discussed in Chapter Three, a pragmatic stance is taken in the present study to investigate CLIL effects on learners’ academic performance in a Chinese tertiary education programme. Furthermore, Grix (2018) asserts that researchers should proceed in a systematic pattern where the development of an appropriate methodology accords with its ontological and epistemological assumptions. Consequently, inspired by the pragmatic philosophical

underpinnings behind the current research topic, a mixed-method approach is conducted in the study to understand the multi-faceted nature of CLIL holistically and longitudinally. The rationale is provided in the following sections in terms of the philosophical stances, methodological design and methods considerations.

4.1.1 The ontological and epistemological stance: Pragmatism

As noted by Creswell (2018), although a researcher's ontological and epistemological underpinning remains largely hidden in research, it still needs to be made explicit because it ultimately influences the practice of research in an implicit way. As is shown in prior sections, discrepancies exist in the definition and empirical review of CLIL research. Such difference derives from the contradiction between different research perspectives, like policymakers', practitioners' and integrated perspectives as discussed in Chapter Two. Different research perspectives reflect different ways of thinking at the ontological and epistemological levels held by different communities of researchers (Johnson & Christensen, 2008). This section articulates how my design of research has been influenced by the different ontological and epistemological underpinnings in the light of the different research positionalities I take in the current study.

Positivist assumptions

In the social science field, the choice of research paradigm has been in a “war” for more than a century (Johnson & Christensen, 2008, p. 31). Based on conflicting assumptions, concepts, values, and practices, two communities of researchers hold contrasting perspectives of thinking and doing research, namely quantitative and qualitative research paradigms. Quantitative purists contend social phenomena to be objective and independent (John W. Creswell, 2012). It is considered that “time- and context-free generalisations are desirable and possible”, while “causes of social scientific outcomes can be determined reliably and validly” (Johnson & Onwuegbuzie, 2004, p. 14). In this sense, research findings can be generalised to diverse times and

populations, and be replicable in various settings. In line with these positivist assumptions, researchers in the field of education are expected to empirically test and justify presumed hypotheses in a rather formal writing style using the impersonal passive voice and technical terminology involving confirmation and falsification.

With the positivist assumptions underpinning the policymakers' perspective, CLIL is regarded as a societal phenomenon "with relatively stable characteristics that can be explained in terms of structural determinist ideas", which implies that the aggregate concept of CLIL follows regularities that researchers need to identify (Bonnet, 2012, p. 68). CLIL learners' achievement is conceptualised as "performance in problem-solving situations" and is reduced to "an observable surface phenomenon" of measurable content and language input and output (Bonnet, 2012, p. 72). According to this perspective, CLIL effects have been scrutinized in a number of studies in a linear and uni-directional way to test the pre-existing hypothesis of causality between CLIL exposure and learner achievement (Llach, 2009; Ruiz de Zarobe, 2013; Sylvén & Ohlander, 2015; Xanthou, 2011). This perspective of CLIL is primarily informed by positivist epistemological thinking, which calls for quantitative research methodology and hypothesis testing of CLIL effects on learner attainment through numerical analysis.

In the present study, my original view of CLIL has been driven by a positivist attitude from policymakers' perspective, which entails an outsider stance and informs the employment of an experimental design and the collection of numerical test scores. Subscribing to the positivist assumptions, I assume CLIL to be an already constructed theory following certain regularities. Its perceived effectiveness has been replicated in different populations and settings (McDougald, 2016; Ruiz de Zarobe, 2013). Some research findings might be generalised to my research context and benefit my target learners as well. Aligning with the positivist assumptions, I resorted to an experimental design in this study to compromise the influence of potential confounding variables in a contrived setting. Based on learners' precise quantitative test scores of language and

content, this experiment was designed to investigate the causal relation between CLIL exposure and its effectiveness on a group of Chinese university learners in my working context. In addition, the perceived effect of this study might be generalised to a more diverse student profile by policymakers.

However, with the in-depth review of previous CLIL studies, I also realised that when the perceived effectiveness is confirmed in some quantitative studies, little attention has been given to the specifics of a selected group of individuals, like the Chinese university learners in my working circumstances. The causal relation sought in quantitative studies may not well reflect local constituencies' understandings and directly inform CLIL practice in my target context. Hence an interpretivist stance of thinking came to influence my perception of CLIL as well.

Interpretivist assumptions

In contrast, qualitative purists maintain social realities to be subjective and multiple-constructed (Bruton, 2013, 2015; Dallinger et al., 2016). It is believed that “time- and context-free generalisations are neither desirable nor possible” and “research is value-bound, that it is impossible to differentiate fully causes and effects, that logic flows from specific to general” (Johnson & Onwuegbuzie, 2004, p. 14). Put another way, these qualitative purists (interpretivists) hold a subjective view that knowledge in social science is constructed through interactions within the social, cultural and historical context in which communities are situated (Schwandt, 2000). In the educational domain, these qualitative purists tend to describe educational issues in a detailed and rich style, being attached and sometimes involved in the research.

Considering CLIL studies, from this interpretivistic view, CLIL is perceived as an idiosyncratic phenomenon possessing “crucial features that can only be elucidated by taking into account the idea of agency”, which means the CLIL community possesses insights into the consequences of classroom practice which “transcend external

boundaries and transform structures and systems” (Bonnet, 2012, p. 68). CLIL achievement is regarded as “a deep structure” that influences CLIL learners’ actions across situations and is related to their “emotional, volitional and reflexive dispositions” (Bonnet, 2012, p. 72). Hence contextual approach has been more preferred in a number of qualitative CLIL studies to scrutinise diverse variables mediating CLIL effects and educational consequences in specific contexts, such as learner motivation, classroom engagement and teacher-student interaction. The practitioner’s perspective of CLIL research is informed by interpretivist philosophical thinking, which calls for a qualitative approach to evaluate CLIL effectiveness based on participants’ viewpoint and interpretation of how CLIL affects learner outcome through text data (Bonnet, 2012).

The interpretivist stance from the practitioners’ perspective provides me with an insider’s view of CLIL, informing the inclusion of qualitative research techniques and the aggregation of textual data in my study. Subscribing to the interpretivist philosophical thinking, I assume CLIL to be a complex phenomenon, and its effectiveness needs to be perceived from participants’ perceptions in local contexts. The qualitative data in the words and categories of participants provide rich and deep description of CLIL effectiveness and lend themselves to exploring how and why CLIL takes effect in my research context. With the interview feedback from CLIL participants in the present study, findings from my research might be responsive to local situations and stakeholders’ needs and inform practice in my working context.

Despite the aforementioned strengths of investigating CLIL effectiveness from an interpretivist stance, its weaknesses need to be taken into account. One potential complication is to investigate CLIL effectiveness in naturalistic settings in the target research context. As described in Section 1.2, CLIL is still at an initial stage in China. Without explicit policy endorsement from the government, there is a lack of coordinated efforts among key stakeholders to establish CLIL programmes in China’s

public higher education system. Together with the frequently cited lacuna of sample disparity in some previous naturalistic studies as articulated in Section 2.2.2, it is highly difficult to find an existing or ongoing CLIL programme and scrutinise its effect in a non-contrived context in a Chinese university. In addition, findings from one qualitative study with a small sample are difficult to generalise in a fashion to which policymakers are generally accustomed. In contrast with its quantitative counterpart, qualitative research is more prone to make a deep and rich understanding of participant perception at the individual level with textual data. However, it is more labour intensive and time-consuming, resulting in fewer participants being included in the research (Johnson & Onwuegbuzie, 2004). The comparatively small size of sampling might have lower credibility with potential stakeholders, especially among policymakers. Besides, the qualitative data, like the interview feedback, might be regarded by policymakers as more liable to be impacted by researchers' personal biases and participants' idiosyncrasies. Therefore, the qualitative research findings might be overlooked at the high level of policymaking and impeded from generalising to a broader context of the Chinese higher education system.

A pragmatic worldview

As can be seen from the dichotomous views, the primary contrast between quantitative and qualitative paradigms is not only “the absence or presence of quantification”, but the essential issue at the ontological and epistemological level of “how humans and their society should be studied” (Bryman, 2004, p. 14). Due to the divisive difference in underpinning worldviews and philosophies, both quantitative and qualitative purists insist that the two dichotomous paradigms, along with their corresponding research methods, cannot be used together. This “incommensurability thesis” obliges researchers to hold an either/or position and pay allegiance to one research school or thought even though its weaknesses are unequivocal (Johnson & Onwuegbuzie, 2004, p. 14). As noted by Johnson and Christensen (2008, p. 31), the major deficiency of this paradigm polarity is the unawareness that “creative and

thoughtful mixing of assumptions, ideas, and methods can be very helpful and offers a third paradigm”.

However, since the 1990s, the either-or position of positivism and interpretivism has been challenged by pragmatists (James, 2002; J. P. Murphy, 1990; Newman & Benz, 1998). Different from the dualistic epistemology of positivists and interpretivists, pragmatists are prepared to compromise between the two prior polarised views of scientific research. It is held by pragmatists that “a continuum exists between objective and subjective viewpoints and the choice of which depends on the nature of the research question being asked and the particular point in the research process” (Giacobbi Jr, Poczwardowski, & Hager, 2005, p. 22). In contrast with the positivist and interpretivist worldviews which “arise out of antecedent conditions”, the pragmatic worldview “arises out of actions, situations and consequences” (John W. Creswell, 2009, p. 7). It imposes particular attention to applications, like what works and the solutions to research problems (Patton, 2002). Rather than the concern on research methods, pragmatism focuses more on the research problem and resorts to all research strategies available to derive knowledge about the problem (John W. Creswell, 2013). Pragmatism is not solely committed to either the positivist or the interpretivist way of thinking. It “sees positive value in both the quantitative and the qualitative views of human behaviour” (Johnson & Christensen, 2008, p. 35). This entails researchers the freedom to choose both quantitative and qualitative assumptions in their research (David L Morgan, 2007). To pragmatists, “truth is what works at the time”, and the best method is the one that “provides the best understanding of a research problem” (John W. Creswell, 2009, p. 7). In other words, pragmatism could not end the philosophical war between positivism and interpretivism but provides an immediate and useful middle position philosophically in social science research.

Concerning research on CLIL, pragmatic researchers tend to discard the metaphysical disputes on whether CLIL is a natural existence or an emergent social

world. They do not see CLIL as “an absolute unity” but “an external world independent of the mind as well as lodged in the mind” (John W. Creswell, 2009, p. 7). Thus, pragmatic CLIL researchers have the freedom to choose the research perspective and research methods that best meet their needs and purposes, namely the integrated perspective and mixed research methods. The rationale for adopting the integrated perspective in CLIL research lies in its allowance for the researcher to address CLIL effectiveness from practitioners and policymakers’ perspectives. As argued by John W. Creswell and Creswell (2018), a researcher’s ontological and epistemological stance not only influences the choice of research design and approach, but may generate certain allegiances, inclinations or biases as well. A pragmatic position might facilitate CLIL researchers to improve communication between contradictory research stances. According to the integrated perspective underpinned by the pragmatic way of thinking, CLIL is a multi-faceted complex occurrence. Its complexity is further intensified by taking into account its contextual factors in diverse settings, as illustrated in the elaborated 4Cs theoretical framework. The adoption of the integrated perspective allows the researcher to investigate the complex nature of CLIL in the convergence of both policymakers and practitioners’ perspectives.

On the one hand, the CLIL approach might be perceived as an already constructed societal phenomenon with relatively stable characteristics from the positivist perspective. Its effectiveness has been tested and replicated in a large body of empirical studies on different populations, as articulated in previous chapters (Adamson & Coulson, 2014; Alonso, 2017; Pérez-Cañado, 2016). On the other hand, CLIL is regarded from the interpretivist perspective as an idiosyncratic phenomenon with crucial features that can only be expounded by considering stakeholder agency. CLIL participants and researchers are embedded in specific social communities, and subjective judgments could be made throughout the research process (Bonnet, 2012). Therefore, CLIL research may not be entirely objective or value-free.

In view of the multi-facets of CLIL as illustrated in the definition, the 4Cs framework, and the empirical review of previous studies from policymakers', practitioners' and integrated perspectives, a pragmatic stance is adopted in this study. It is expected that the present study's findings could arrive at some extent of generalisation and have credibility with key stakeholders in power, like policymakers. It is also anticipated that the complex phenomenon of why CLIL effects occur in the target research context could be explored with CLIL participants' personal experiences in words.

4.1.2 The research paradigm in the present study

In line with the preceding articulations, both the positivist thinking behind the quantitative perspective and the interpretivist thinking behind the qualitative perspective have respective strengths and weaknesses in conceptualising CLIL and evaluating its effectiveness. Hence, a more holistic and longitudinal research design underpinned by the pragmatic worldview has been adopted in the current research based on the integrated perspective, together with the adoption of a mixed methods approach.

4.1.2.1 The ontological and epistemological stance

The primary aim of the current research is to test and assess whether CLIL enhances target learners' academic performance in a Chinese higher education programme. Given the inceptive status of CLIL in the target research context, the researcher needs to "consider agreement within a research community as a way to approach objectivity" (Giacobbi Jr et al., 2005, p. 22). In other words, the present study's findings need to arrive at some extent of generalisation and have credibility with key stakeholders in power, like policymakers. Therefore, the study is preferred to be conducted in a situation that controls confounding variables and produce quantitative data that is reasonably available for public inspection and policy consideration. In addition, as per the current research, the investigation of whether CLIL affects target learners might not be sufficiently responsive to grassroots CLIL participants' needs.

The quantitative data from the study seems to be too abstract and general to direct application in the target context. In other words, the current study needs CLIL participants' personal experiences in words and categories to explore the complex phenomenon of why CLIL effects occur in the target research context. With deep and rich description and understanding of CLIL participants' perceptions, the current study might be able to demonstrate CLIL effects in a deep and detailed way to potential readers and practitioners. In this way, the integrated perspective allows the researcher to take an explicitly value-oriented qualitative approach to CLIL research.

As noted by pragmatic researchers (Giacobbi Jr et al., 2005; Rorty, 1995), any method to a social phenomenon is inescapably partial. Given its complexity, a better understanding of the multi-facets of CLIL effects can be achieved via multiple uses of ways of knowing. With the integrated perspective of CLIL research, the researcher is allowed to respond to a broader and more complete range of research questions in the current study. It can be addressed not only the 'whether' question to testify CLIL effectiveness, but also the 'how/why' question to explore its occurrence in the target context. Given the respective strengths and weaknesses of policymakers' and practitioners' perspectives in conceptualising CLIL and evaluating its effectiveness, a holistic CLIL study based on an integrated perspective might be given more possibilities to increase the conclusiveness of findings. Therefore, a mixed methods research approach has been employed in the present study, as articulated in detail in the following section.

4.1.2.2 Methodology

In line with the integrated perspective underpinned by the pragmatic worldview, a mixed methods research approach is adopted in the current study. As defined by Johnson & Onwuegbuzie (2004, p. 17), mixed methods research is "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study". Unlike its

two predecessors, the mixed methods approach is not a restrictive form of research. It allows the researcher to take an inclusive and eclectic stance on choosing method and conducting research. It legitimises multiple methods in response to the research questions instead of limiting the researcher's choices. The multiple research methods "offers the best chance to obtain useful answers" to the research question (Johnson & Onwuegbuzie, 2004, p. 18). With regard to a combination of the RQ and the Sub RQs like those in the present study, a mixed methods approach might provide better and more complete answers.

Besides the multiple research methods included, a mixed methods research approach potentially offers the chance to mix these methods in an effective manner. This is regarded as "a major source of justification for the mixed methods approach, because it might produce results superior to monomethod studies" (Johnson & Onwuegbuzie, 2004, p. 18). With the awareness of the strengths and weaknesses of both quantitative and qualitative research, the researcher is allowed to follow the fundamental principle of mixed research (Johnson & Turner, 2003). Aligning with this principle, the researcher might converge multiple types of data from different strategies in a way that "the resulting mixture or combination is likely to result in complementary strengths and nonoverlapping weaknesses" (Johnson & Onwuegbuzie, 2004, p. 18). With the employment of both quantitative and qualitative methods in the same study, the researcher is allowed to use the strengths of one method to overcome the weaknesses of another and provide more robust evidence for a conclusion by converging and corroborating findings from both methods. The textual data obtained from the qualitative method can add meaning to numbers and deepen insights of findings; whereas the numbers from the quantitative method can add precision to texts and increase the generalisability of the results. The two types of data merge to produce complete knowledge to inform theory and practice.

The current research started with a broad quantitative investigation of CLIL

participants' achievements within an experimental design to generalise findings to the population of interest. In the subsequent phase, qualitative interviews were administered to collect deep and detailed participant perceptions to explain initial quantitative results. To be more specific, adding qualitative interviews to the experimental design was used as a manipulation check and a means to review directly with participants the effects of CLIL. By tapping into CLIL participants' and the teacher's perceptions, some potential drawbacks of the experimental methods might be circumvented. For instance, prior CLIL studies have shown that some possible intervening variables, like motivation and sociocultural status, might be concealed in numerical data and affect the causal relation between CLIL exposure and learner achievement (Pérez-Cañado, 2016). The interview of CLIL participants and the teacher was potentially an effective way to reveal these hidden variables. Correspondingly, CLIL participants' test scores were considered as a systematic instrument to gauge CLIL learners' achievements and supplement interview results as measured evidence. In addition, both the quantitative and qualitative findings in the study were improved by adding a component of randomised sample selection from the population of interest to enhance generalisability. As per findings from different methods, when the quantitative statistics and qualitative results were corroborated, greater confidence was held in the singular conclusion. When they conflicted, a greater and deeper understanding of CLIL effects was gained, and research conclusions were accordingly revised.

In view of the above, it can be seen that a mixed methods research approach provided the current study with a full range of methods of data collection and a flexible way of organising these methods. It allowed the researcher to understand CLIL effectiveness in the target research context by making inferences to both quantitative and qualitative databases. These methods and related considerations will be detailed in the following section of the research design.

4.2 The research design

“Research design is a plan for collecting and analysing evidence that will make it possible for the investigator to answer whatever questions he or she has posted” (Ragin, 1994, p. 191). A well-planned and presented research design guides the whole process of the research based on evidence. Generally speaking, the nature of the research topic, research questions, research targets, and the broader practical considerations of time, access, circumstances, feasibility and resources largely determine the research design. Different from designing the quantitative or qualitative research approach mainly focusing on a single way of data collection and analysis, a mixed-method research design is regarded as a procedure for mixing, collecting and analysing both methods in a single study to understand a research question (J. W. Creswell & Plano Clark, 2011). In this sense, designing a mixed-method approach could be more complex by considering various issues, including how, when and what to mix when conducting the research. The mixed-methods research does not simply perform two distinct quantitative and qualitative forms but integrates, links and embeds these two forms into a whole approach.

The present study is designed to employ the mixed-method approach to investigate CLIL effects on learners’ academic performance in a Chinese tertiary education programme. The purpose of this section is to illustrate how the present research design is planned to address the aforementioned research questions and triangulate the research findings by adopting a broad range of data collection methods in the different phases. In detail, a two-phase mixed-method design is arranged, with main attention to the experimental design results. In Phase One, a quantitative study of performance tests is conducted to scrutinise the statistical relationship between CLIL exposure and learners’ attainment in language and content, including pre- and post-tests of linguistic and content attainments. Building on policymakers’ overview of CLIL effectiveness in Phase One, several interviews are conducted in Phase Two among learners and the teacher to achieve a general understanding of their perceptions of their attainment in

CLIL. The interview methods in the second phase are expected to investigate how CLIL learners perceive their progress in the four competences, namely content, cognition, communication and culture. The principles and detailed arrangement of this two-phase mixed-method design would be further justified in the following sections, together with a brief introduction of sampling strategies and the research context in the present study.

4.2.1 Samples

The quality of a piece of research stands or falls not only by the appropriateness of mixed methods and experimental design but also by the suitability of sampling strategies adopted (John W. Creswell, 2012). In this study, more than sixty second-year undergraduates from the university I work for were selected as research participants. The rationale is presented aligning with the four key factors in sampling, namely representativeness, access to the sample, sample size, and the sampling strategy (Cohen et al., 2017).

With regard to sample representativeness, related considerations are made on the selection of sample participants' university and initial English proficiency to guarantee their parameter characteristics are similar or identical to those of the wider population of Chinese university EFL learners. The university where the participants study is a typical public Chinese university. Hence all national policies, curriculum designs, syllabi, and evaluation documents are implemented in this institution, ensuring research participants receive similar policy and institutional influence as their counterparts in other Chinese universities. Concerning the selection of year two students, the reason lies in the consideration of participants' initial language proficiency. In year two, all students in the programme have finished one year's study under the mainstream EFL approach and have passed the national proficiency test of CET-4 (College English Test Band 4), providing necessary language preparation to engage in the CLIL approach.

On the consideration of better access to the sample, research participants were

selected from the university I work for. In my own working context, I would be given more opportunities to “get to the sample”, and “get the information out to the wider public” (Cohen et al., 2017, p. 110). Besides easy access to potential participants and permission from gatekeepers, I am entailed greater freedom to utilise and release research data to the public outside campus.

With respect to sample size, it is undeniable the larger the sample the better, because it brings greater reliability and allows more sophisticated statistics (Cohen et al., 2017). However, to hold constant the confounding effect of class size on learning outcomes, a normal class size in Chinese universities, namely 63, was adopted in this study. All participants were allocated to two sub-classes of CLIL (experimental) and non-CLIL (control) subgroups during the experiment. The size of each group is right beyond the minimum sample size of thirty, as demanded for statistical analysis (Cohen et al., 2017).

Concerning the sampling strategy, all participants were randomly allocated to a CLIL and a non-CLIL subgroup of equal size. An equal chance of being included or excluded from the sample might hold constant some expected and un expected confounding variables in research and “permits two-tailed tests to be administered in statistical analysis of quantitative data” (Cohen et al., 2017, p. 110). Participants’ informed consent was gained on their allocation to two subgroups. According to the pre-arranged plan, those who would disagree with random allocation would be allocated at their will, but would not be regarded as research participants in this study. This would imply a potential dilemma between the pursuit of research validity and the protection of participants’ rights and values. However, all participants agreed with random allocation.

4.2.2 The selection of mixed-method strategies

As Cohen et al. (2017, p. 78) point out, “there is no single blueprint for planning

research”, and the design is largely governed by “the principle of fitness for purpose”. With regard to the planning of a mixed methods research, the major problem is “there are a plethora of designs in existence” for the researcher to decide on (Leech & Onwuegbuzie, 2009, p. 266). In response, different researchers have proposed different typologies to categorise mixed methods strategies. For instance, Johnson and Turner (2003) propose 35 types of designs, while J. W. Creswell and Plano Clark (2011) identify 12 classification systems. In actual research, as Maxwell and Loomis (2003, p. 244) note, the diversity is “far greater than any typology can encompass”. This leaves researchers difficulties in choosing optimal mixed methods research designs for their research. To better design the procedures involving collecting, analysing, and interpreting quantitative and qualitative data in the current study, three strategies that impact the design of a mixed-methods study have been taken into account, namely “timing, weighting, and mixing” (John W. Creswell, 2009, p. 190).

Timing:

The first aspect that needs consideration is ‘timing’, namely the sequence of quantitative and qualitative data collection in the research. If the two methods are conducted in phases, it is called ‘sequential’; while those gathered at the same time is ‘concurrent’ (John W. Creswell, 2009, p. 191). The choice of timing draws upon the researcher’s initial intent of the research.

With regard to the current study, the ‘timing’ is first sequential with an explanatory phase, followed by a qualitative phase exploring the RQs with participants at site. As is shown in the main RQ, the central hypothesis of the present study is an idea that has some support in previous literature but needs further empirical validation. Therefore, an explanatory Phase 1 was designed to scrutinise the causal relationship between CLIL exposure and learners’ academic achievement in a Chinese tertiary education programme. In this mainly quantitative phase, CLIL participants’ linguistic and content achievements were measured by test scores. The quantitative data collected in this phase informed the design of interview questions in the accompanying phase.

While in mainly qualitative Phase Two, both the learners and the teacher's perceptions on CLIL achievement in the 4Cs were investigated in interviews to collect participants' perceptions on their CLIL achievements and the corresponding causes. On the whole, the 'timing' of this study was sequential, with the explanatory phase first.

Weighting:

The second factor involved in mixed methods design is the weight or priority given to quantitative or qualitative research in a particular study (John W. Creswell, 2009). The weight of the two research types might be equal or with emphasis on one over the other. According to John W. Creswell (2009), the priority of one research type hinges upon the researcher's interests, the target audience, and what the researcher seeks to emphasise.

With regard to this current study, the researcher's primary interests were to test the effectiveness of the CLIL approach in the target research context and gain credence among potential stakeholders who are meanwhile the target audience of this thesis. Therefore, a deductive approach was primarily used, with the quantitative data of test scores emphasised first. The qualitative data collected in the subsequent interviews were intentionally used in a supportive role to the quantitative part. In this way, the weighting of the current research is 'quantitative'.

Mixing:

As noted by J. W. Creswell and Plano Clark (2011), the way different types of data mix in a study is of great significance. Mixing means "either that the qualitative and quantitative data are actually merged on one end of the continuum, kept separate on the other end of the continuum, or combined in some way between these two extremes" (John W. Creswell, 2009, p. 192). According to the extent of merging quantitative and qualitative data, John W. Creswell (2009) lists three types of mixing, namely

‘connected’, ‘integrating’, and ‘embedding’. ‘Connected’ means the data analysis of the first phase is connected with the data collection of the second phase in mixed research. While ‘integrating’ indicates the two types of databases are collected separately but merged in analysis. ‘Embedding’ means a secondary data form is utilised within a larger study to support another primary database.

As per the present study, the type of mixing is ‘connected’. This two-phase research started with a quantitative phase. The analysis of CLIL participants’ test scores and their related results were used to inform the interview design in the mainly qualitative follow-up phase. In this way, the data results of Phase One were ‘connected’ across phases with the data collection in Phase Two.

Sequential explanatory design:

In line with the above articulation, the three factors, timing, weight, and mixing, facilitate to shape the procedures of the present mixed-methods study, namely sequential explanatory design, as illustrated in the following figure (John W. Creswell, 2009, p. 193).

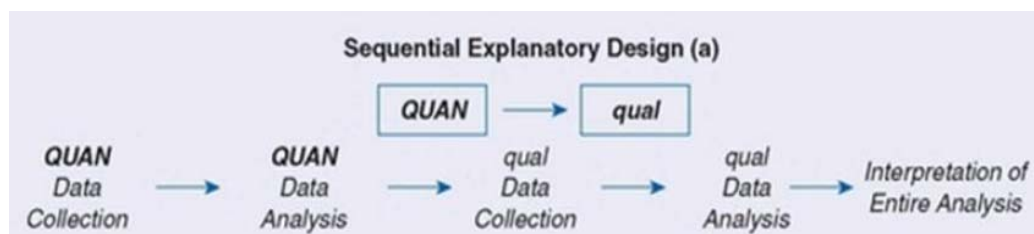


Figure 4.1 Sequential Explanatory Design

As a popular mixed-methods design strategy, the sequential explanatory design involves a two-phase project with “strong quantitative leanings” (John W. Creswell, 2009, p. 194). It is featured by the collection and analysis of quantitative data in the first phase and a qualitative second phase built on the initial results of the first counterpart. The qualitative data in this type of design is typically used to explain and interpret the initial quantitative results in more detail. According to John W. Creswell

(2009), one strength of the sequential explanatory strategy is to handle unexpected results from a quantitative phase. The qualitative follow-up phase can be used to examine noteworthy phenomena in a deep and detailed way. Another strength of this strategy lies in its straightforward nature. The whole procedure of the sequential explanatory design can be divided into explicit and separate stages, making it easy to describe, implement, and report.

Concerning the present research, the strengths and weaknesses of adopting the sequential explanatory strategy are articulated according to the following steps of data collection, data analysis, interpretation, and validity.

As per data collection in the present study, the sequential explanatory design allowed it to proceed in two distinct phases, with extensive sampling of CLIL and non-CLIL participants in the quantitative phase and potentially purposeful sampling in the following qualitative phase. The quantitative data obtained in the first phase directly informed what factors could be built on in the qualitative phase, including outlier cases, noteworthy predictors, significant results relating variables, or even demographics (John W Creswell & Creswell, 2018). Take the current study as an example. The participants' pre and post test scores in the quantitative phase showed discrepant progress in the four language competences, with receptive competences outperforming their productive counterparts. Thus, the research participants were grouped into different categories of their receptive and productive performances, with related qualitative data collected from individuals representing each category. As can be seen, the idea of explaining the mechanism, how the factors interacted to prevail CLIL participants than their controls in more depth through qualitative follow-up is one of the key strengths of the sequential explanatory design in the present study.

Similar to the way of data collection, the quantitative and qualitative databases in the sequential explanatory design are analysed separately, which might allow

researchers to accomplish their research in an easier way than some convergent designs (John W. Creswell, 2013). As aforementioned in data collection, the quantitative data in the sequential explanatory design are initially analysed to plan the qualitative follow-up. Besides informing the sampling procedure in qualitative data collection, the quantitative results might as well “point toward the types of qualitative questions to ask participants in the second phase”, which deeply influences the analysis process (John W. Creswell, 2013, p. 225). Take the current study as an example. With the employment of the sequential explanatory design, the analysis could proceed independently for each phase. This allowed the adoption of semi-structured and open-ended questions in the qualitative phase, which differed from the numerical scores in the preceding phase. This design brought the current study some advantages because the collection and analysis process in the two phases could be relatively spaced out over time and easier to administer.

From a holistic perspective, findings from the two phases complemented each other and converged to allow for a more robust analysis of whether and how CLIL affected learners’ academic performance in the target programme.

4.3 Research procedures and methods

As Coyle et al. (2010) propose, the purpose of the research determines the design of the research procedures and methods. An ethnographic study may be more appropriate for an in-depth scrutiny of a specific situation or group. A survey with stratified sampling might be more relevant for an exploration of an uncharted field or drawing generalisations. Whereas for an evaluation of the effects of an intervening manipulation, an experimental study might be more useful.

With regard to CLIL studies, a high proportion of naturalistic CLIL studies have been criticised for being biased by methodological weaknesses and selection prejudice favouring CLIL learners, especially in their initial language and content proficiency,

and other factors like motivational state and parental social status, as is shown in the empirical review (Bruton, 2013; Dallinger et al., 2016). Thus, a mixed-method approach with both quantitative and qualitative research tools is adopted to control potential confounding variables and guarantee “the gains observed are truly ascribable to CLIL practice” (Pérez-Cañado, 2012, p. 330). In detail, in Phase One, a pre-test-post-test control and experimental group design was conducted to verify whether CLIL improved Chinese tertiary students’ content and linguistic performance holistically and longitudinally to convince those key stakeholders. In Phase Two, semi-structured interviews were employed to explore how the CLIL effectiveness was realised in detail among research participants.

4.3.1 Phase One: The pre-test-post-test control and experimental group design

Considering the experimental type in the present study, a quasi-experimental design was adopted. Different from the true experiment design normally administered in laboratory conditions, a quasi-experiment is of the field experimentation generally conducted in dynamic, evolving and open situations like the target research context. A quasi-experiment allows the researcher “to employ something approaching a true experimental design” where non-contamination between the experimental and control groups is impractical throughout the research process (Cohen et al., 2017, p. 282). As per experimental designs, the pre-test-post-test control and experimental group design is regarded as a ‘good’ design to identify the essential features of an experimental study and provide the comprehensive treatment of the subjects (Bloomfield & Fisher, 2019; Shek & Sun, 2012). Thus, in the present study, pre-test-post-test control and experimental group design was adopted at the first phase to verify the causal relationship between the effectiveness of CLIL and learners’ academic achievement. The principles are listed below, including ‘experimental and control group’, ‘randomisation’ and ‘pre-test-post-test’.

Experimental and control group

In the experiment, the target course for both experimental and control groups was Advanced Financial Management (AFM), a compulsory course for the participants in the present study. For the experimental group, the course was taught in L2 (English) for one semester (18 weeks), four lessons per week. As illustrated in Section 3.1 and 3.2, related cognitively-challenging and linguistically-demanding classroom activities and tasks were implemented to provide CLIL participants a student-centred interactive learning scenario for effective content learning, in which content and language related cultures from both native and foreign perspectives are fused to foster learners' intercultural understanding.

With regard to the control group, the content course was taught in L1 (Chinese), together with a mainstream English course, two lessons per week, respectively. All the three courses, AFM in English, AFM in Chinese and English course, were taught by the same teacher. In this way, both groups' total exposure to content and language remained the same to hold constant potential confounding factors of extra content/language/teacher investment and class size effect (Bruton, 2015). Both the content and language courses for the control group were carried out in a mainstream conventional teacher-centred "banking model", in which content knowledge and linguistic knowledge was largely deposited through the teacher's lectures. Most content-related examples in the content class were given from the native perspective whereas in the language class, examples were more from the native and target language cultures.

Randomisation

The second principle in conducting the experimental design is 'randomisation', which ensures the greater likelihood of equivalence. Randomisation is regarded as a gold standard to achieve controllability, causality and generalisability (Coyle et al., 2010). In the present study, the participants were randomised into the experimental and the control groups, 31 and 32, respectively, attempting to produce equivalence over a

whole range of potential variables. With more possibility to establish ‘controllability, causality and generalisability’, the present study’s findings might not be confined to a specific CLIL research context, but replicated to other educational settings to provide a broad description and understanding of the effectiveness of CLIL on learners’ academic performance.

Pre-test-post-test

Concerning the pre-test-post-test design, two groups of tests were given before and post the experiment, including the pre- and post-tests of linguistic attainments and the pre- and post-tests of content attainments. These tests were administered to provide a mainly quantitative overview of CLIL effectiveness in participants’ content and language performance in this given context.

With regard to the evaluation of linguistic attainment, all participants were tested by simulated College English Tests, Band 6 (CET-6), a national standard proficiency test whose scoring reliability and validity have been widely studied (Chen, 2013; C. Wang, 2014; J. Wang & Sun, 2012). This nationally recognised testing framework has been regarded as the dominant assessment criteria by mainstream EFL learners for decades and has been taken by millions of Chinese college students annually. The adoption of CET-6 in this study was expected to establish parity between experimental and control subgroups in English linguistic proficiency and reveal CLIL effectiveness to the public in a straightforward way. The four parts of CET-6, reading, listening, writing, and speaking, were tested before and after the CLIL experiment and rated blindly by the teacher who is a qualified and experienced rater of the test. In this way, learners’ progress in the receptive and productive language is scrutinised respectively.

Compared to linguistic competence, assessing subject content is more challenging as there is no available national standard test. Both the pre- and post- content tests in this study were designed and rated blindly by the teacher to reduce instrumentation and

calibration errors. The pre-test was devised in L1 (Chinese) for both subgroups to guarantee greater equivalence in initial subject proficiency and reduce participants' sensitisation to the role of the English language in content learning in the CLIL experiment. With regard to the post-test, both subgroups were tested in L1 Chinese as well. Its justification lies in evaluative, practical, and theoretical reasons. Firstly, the adoption of the same language and content in pre-test and post-test ensured all participants were evaluated at the same level of difficulty in content knowledge and language for better cross-sectional and longitudinal comparisons. Secondly, assessing content in L1 might disadvantage CLIL participants, but the results might facilitate to defuse stakeholders' largest practical concern that CLIL learners' content learning might be discounted in an imperfectly known CLIL language. Thirdly, the use of L1 in assessment was in line with the aforementioned Cognitive Framework. CLIL learners' higher-order thinking skills developed in L2 English might allow a deep understanding to be communicated in L1 tests.

With the methodological considerations mentioned above, it was expected that the pre-test-post-test control and experimental group design could generate valid and reliable data to answer the research questions and scrutinise the statistical relationship between CLIL exposure and learners' attainment in language and content. Meanwhile, as discussed in Literature Chapter, the multi-faceted nature of CLIL demands an exploration of its effects in intangible achievements, which has not been well supported in the literature review. Therefore, semi-structured open-ended interviews were subsequently administrated as articulated in the following section.

4.3.2 Phase Two: The semi-structured open-ended interviews

Built on the mainly quantitative overview of CLIL effectiveness in Phase One, the purpose of Phase Two was to answer how the CLIL experimental conditions changed learners' outcomes from participants' insights in a Chinese tertiary education context. In this phase, CLIL learners' viewpoints were collected and analysed to explain and

interpret the quantitative results from the preceding phase. The detailed and nuanced descriptions of learners' perceptions and experience were handled to suggest inductively whether CLIL improved learners' academic performance in subject content, linguistic and communicative competence, intellectually cognitive proficiency and intercultural understanding, and how such progress occurred. Besides, the teacher was also interviewed to show her insights and evaluations of CLIL effectiveness in the target context. To realise this aim, the research tool of semi-structured interviews was employed.

In Phase Two, ten CLIL learners from the experimental group were randomly selected for semi-structured open-ended interviews in L1 Chinese in congenial places. The rationale is given to justify respectively 'the semi-structured format', 'open-ended response', and 'the preferred language'.

Semi-structured format

With the principle of 'fitness for purpose', the semi-structured interview format was adopted in the second phase of the present study. In line with the main characteristics of the semi-structured interview, topics and issues of interview questions were specified in advance based on the related literature and results from the first phase, highlighting the outline form of interview questions. Besides, the sequence and wording of interview questions were decided by the researcher in the course of the interview.

The strengths of this format are as follows. Firstly, the outline of the interview questions increased the breadth and comprehensiveness of data and made data collection relatively systematic. Besides, the logical gaps could be anticipated and covered by synchronous communication with continuous clarification. In the process of the interview, respondents could be kept deeply involved and motivated. Thus, the semi-structured interview enabled the researcher to chart respondents' understanding of CLIL effectiveness and disclose unforeseen but influential issues.

Open-ended response

One of the weaknesses of semi-structured format, as Coyle et al. (2010) state, lies in its limit to relate the interview to specific individuals and settings, because the exact wording and sequence of the questions are determined in advance. In an effort to increase its naturalness and flexibility and circumvent its weaknesses, open-ended feedback was preferred in the interview, allowing respondents to speak for their true perceptions of CLIL achievement and exhaust possible issues not covered in the outline form.

In addition, the aspect that needs to be taken into account is the possible distorting effects of power relation between the interviewer and respondents, especially when the interviewer is a teacher in the same university. To circumvent the distorting effects, an open-ended response was hence adopted to reduce interviewer effects and bias when respondents felt free to answer the questions in their own way.

The preferred language

According to Coyle et al. (2010, p. 411), the research interview is defined as “a two-person conversation initiated by the interviewer for the specific purpose of obtaining research-relevant information, and focused by him on content specified by research objectives of systematic description, prediction or explanation”. Thus, as a direct verbal interaction between the respondents and the researcher, respondents’ preferred language might make the interview non-threatening, acceptable and sometimes enjoyable. In the present study, The L1, Chinese, was adopted in the interview as the mediate language for increased validity. The use of L1 ensured that respondents could understand abstract questions and encouraged them to convey their perceptions more accurately and freely, reducing potential misunderstandings.

To sum up, the research design is a blueprint for conducting a study that

“maximises control over factors that could interfere with the designed outcomes from the study” (Burns & Grove, 2005, p. 227). In the present study, to investigate CLIL effects on learners’ academic performance in a Chinese tertiary education programme, a pragmatic mixed-method approach was employed, covering various methods within two successive stages. As Robson (2002) addresses, threats to research validity and reliability can never be completely erased, but may be attenuated by full consideration and appropriate measures. Informed by the related literature on CLIL research, thoughtful research was designed to enhance its validity and reliability, including the pre-test-post-test control and experimental group design in the first phase and the semi-structured open-ended interviews in the second phase. The data collected via different stages, methods and types could triangulate and converge to investigate the effectiveness of CLIL in a more valid and reliable way.

4.4 Ethical issues

Throughout the entire research design, one unavoidable aspect is the ethical issue. As is noted by Johnson and Onwuegbuzie (2004, p. 17), “pragmatism takes an explicitly value-oriented approach to research”. Hence ethical issues cannot be circumvented by any means. A major ethical dilemma of this study is the costs/benefits ratio, namely the balance between the pursuit of truth for public interests and the potential costs to participants’ individual rights (Frankfort-Nachmias & Nachmias, 1992). According to the ontologically and epistemologically dualistic stance of pragmatism, judgments in this type of research are made in a continuum between clearly ethical on one end, and clearly unethical on the other, together with the principle of fitness for purpose in different research settings and other values at stake (Kimmel, 1988).

The primary focus of ethical issues in this study is informed consent. In the methods of testing and interview, intrusions and interruptions inevitably occurred to all participants. Hence, their permissions had to be gained in advance. Before data collection in each stage, all participants were informed that they had full rights to refuse

or withdraw participation at any stage of this research. Participants were given full information about potential consequences and threats, like conceivable difficulties with lectures in L2, possible disadvantaged summative assessments including underachievement in the content test and potential slower progress in English proficiency. However, as to research procedures and purpose, merely a general description was given in fear of participants' alteration in behaviours and elicitation of reactivity effect. For instance, experimental participants in this study were informed of the general research purpose of investigating CLIL effects on learner attainment instead of the detailed objective of CLIL impact on the four different competences. It was expected that such reasonably informed consent would minimise bias in the research outcomes. Besides participants' consent, permissions were also gained from related gatekeepers like the programme director and the university administrative.

Another ethical issue is participants' right to privacy. The informed consent gained does not mean the researcher is authorised to violate participants' privacy (E. Murphy & Dingwall, 2001). In this study, two approaches were employed for this purpose, namely anonymity and confidentiality. The aim of anonymity is to guarantee a participant's identity will not be disclosed by the information provided (Grinyer, 2009). Certainly, respondents in the interview could be anonymised. Therefore, aliases or codes were used in the study to identify respondents. In fact, the protection of respondents' privacy in interviews relied more on confidentiality than anonymity. Although the researcher generally knows by whom the information is provided, promises should be given that this connection will not be publicly known (Gray, 2009). In this study, both quantitative and qualitative data would be accessed publicly for general benefits, which means the promise of confidentiality might be betrayed. Therefore, some techniques for confidentiality advised by John W. Creswell (2013) was adopted. For instance, all personal information and other identifying marks of participants were deleted from the data released. As to the information gathered from the interview, reports were made upon micro-aggregation, in which information was

disseminated via the construction of ‘average persons’ from data on individuals (Cohen et al., 2017).

In short, ethical considerations are indispensable to this study. More comprehensive precautions against potential ethical difficulties might gain more trust from participants and encourage them to provide more authentic information to enhance research validity and reliability.

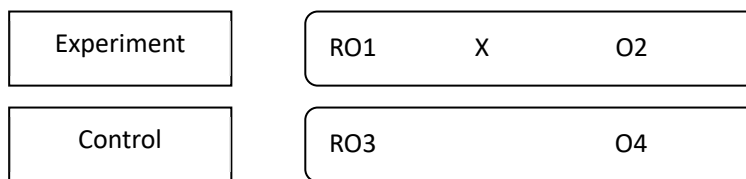
4.5 Data collection and analysis

By recording and selecting materials to inform the research study, relevant information might be turned into data to answer research questions (Richards, 2005). In this study, as presented in the aforementioned sections, data were collected within different forms through various research methods, including those quantitative data from the tests at the first phase, and qualitative ones from interviews at the second phase. Collecting multiple forms of data in this study enabled the information to be interpreted through diverse perspectives and different lenses, making the findings more comprehensive and solid. In addition, it also provided more robust reliability and validity to the data source by triangulation. In this sense, with a wide range of data, both the design of the data collection process and the detailed tools for analysing data were fully considered.

4.5.1 Quantitative data from the tests

As Cohen et al. (2017) recommend, quantitative data analysis is a powerful way of research that could not only be related to small-scale study like case study and action research, but serve large-scale surveys such as nationwide investigations. By analysing quantitative data, researchers could describe trends in the data to address research questions, test hypotheses or verify relationships of variables. In the present study, in line with the research question to investigate the effectiveness of CLIL on Chinese college students, the pre-test-post-test design was adopted, which provided a range of numerical data. As proposed by Coyle et al. (2010), the pre-test-post-test design could

be represented as the following diagram:



This simple and direct diagram makes the calculation of the causal effect easy and researchable, including three steps below:

Step One: Subtract the pre-test score from the post-test score for the experimental group to produce score 1;

Step Two: Subtract the pre-test score from the post-test score for the control group to produce score 2;

Step Three: Subtract score 2 from score 1, to produce score 3.

According to Campbell and Stanley's (2015) terminology, these three steps could be summarised into the following formula: $(O2-RO1)-(O4-RO3)$. The above formula was taken in the present study to analyse and compare the numerical data collected from the two periods of tests to answer the first research question.

As Cohen et al. (2017) suggest, numerical analysis can be performed by adopting software, like the Statistical Package for Social Science (SPSS), to apply formulae and conduct computations. Consequently, the numerical data extracted from the first phase tests were processed by SPSS 20.0, and reliability coefficients of the data were fully considered. When dealing with the data through SPSS, various tools of data analysis were applied to present the results of the tests in the study. For example, SPSS was used to compute the mean scores and standard deviations in a descriptive pattern to provide a general picture of participants' language and content achievements at pre- and post-tests. In addition, Pearson Correlation Coefficient was used to identify relationships among these performances in terms of CLIL effectiveness in the Chinese educational context.

4.5.2 Qualitative data from the interviews

Analysis of qualitative data involves the process of organising, justifying and explaining the data in order to make sense of participants' perceptions, viewpoints, and definitions of certain issues (D. L. Morgan, 1988). One of the potential complications of analysing qualitative data is the reduction of abundant volume of oral and written data into manageable and comprehensible components. In this sense, the choice of appropriate ways of data analysis should be taken into consideration in advance. Among various ways to analyse qualitative data, researchers frequently resort to content analysis, which is defined as "the process of summarising and reporting written data--the main contents of data and their messages" (Cohen et al., 2017, p. 475). By classifying textual data into some categories, the content analysis aims to code open-ended questions, reveal the focus of correlated matters, and describe trends in communicative contexts.

In the present study, under the principle of fitness for purpose, content analysis was taken as the primary means to analyse qualitative data from interviews in the second phase. The interviews were conducted in Mandarin Chinese and audio-taped in digital format, which were then transcribed as a precise record of participants' responses, including laughter, giggles, and pauses. The transcription of interviews comprised more than 15,000 Chinese words, and all data were translated into English immediately after the interviews. During the translation process, to reduce or overcome bias caused by the researcher and triangulate the validity and reliability of the translation, a professional Chinese-to-English translator was invited to do the translation. The results showed that the use of wording, sentence structure and comments in the translator's version was almost the same as the researcher's own translation, indicating the accuracy and appropriateness of the translation process.

After translation, all data were put into NVivo for coding and developing "connected themes" and "linkages grounded in the data" (Herbert & Herbert, 1995, p. 237). By listening to the data repeatedly in NVivo, concepts and themes were picked

out based on the words that participants “frequently use” and “label ... as an important concept”(Herbert & Herbert, 1995, p. 230). Certainly, due to the wide range of interview data, there were a number of themes and concepts discovered from the translation, which made the research results challenging to interpret. Consequently, when coding the data, the researcher grouped similar ideas, concepts and connections into categories. For example, on discussing how students perceived their language achievements in the CLIL context and put them into the same category, the researcher looked for “frequently used” words like “improvement”, “involvement”, and so on.

Summary

In accordance with the complex conceptualisation of CLIL, together with the considerations of the related CLIL studies in the literature, a pragmatic stance was taken to investigate CLIL effectiveness in a Chinese tertiary education programme. Subsequently, inspired by the philosophical underpinnings behind the research questions, a mixed-method approach was employed in this study to provide an in-depth and broad understanding of CLIL effects on participants’ academic performance. The approach gave flexibility for taking various methods within two successive stages, including the pre-test-post-test control and experimental group design in Stage One; the semi-structured open-ended interviews in Stage Two. The data collected via the different stages, methods and types might triangulate and converge to scrutinise the effectiveness of the CLIL approach in a more valid and reliable way.

CHAPTER FIVE – Data Analysis

As depicted in the former chapters, the present study adopted a mixed-methods research design with quantitative and qualitative methods to study CLIL effectiveness in the target context. To answer the research questions, a series of pre- and post-content and language proficiency tests were conducted among participants in a comprehensive university in Eastern China in September, 2017 and January, 2018, respectively. The research questions were firstly answered with the quantitative testing data. Clearly, quantitative data alone might not reveal the whole picture of its effectiveness and limitations. Hence, qualitative interviews were subsequently conducted to provide detailed and nuanced descriptions of learners' perceptions and experience and to suggest whether CLIL could improve learners' performance in content, communication, cognition and culture, and notably how such improvements occurred.

In this chapter, the researcher reports both quantitative and qualitative data collected from the tests and interviews. Meanwhile, related analysis and interpretations were conducted to understand how the CLIL approach impacted learners' achievements in content, communication, cognition and culture from a holistic 4Cs perspective.

5.1 Quantitative findings from Phase One

Before the initiation of data collection, the demographic and background information of the participants in both control and experimental groups were collected, including participants' gender, L2 learning experiences and content achievements in pre-requisite courses. In the present study, the 64 participants, including 37 female and 27 male participants, were second-year undergraduates. Concerning their language learning experiences, most of these participants had studied English for approximately ten years, since grade three in primary school. In addition, the majority of the participants (81.25%) reported having attended extracurricular language tutorials to improve their language proficiency. With regard to their content competence, all the participants had passed the course of Financial Management, a compulsory and pre-

requisite course for the participants in the present study, and learned some finance-related knowledge.

5.1.1 Statistic descriptive of quantitative tests

In order to achieve a largely quantitative understanding of CLIL effectiveness in participants' academic performance, the pre-test-post-test design was adopted in Phase One. The content pre- and post-tests were developed, in consultation with the related teacher, based on previous tests that were given to other learners in the same course. The content proficiency pre- and post-tests were conducted in the participants' regular classroom for both experimental and control groups on September 15th, 2017 and January 12th, 2018, respectively. A total of 64 participants completed the content pre-test: 32 learners in the experimental group (EG) and an equal number in the control group (CG). However, one participant from the E group quit the post-test for personal reasons, and subsequently, his score was excluded from the data set, resulting in a total of 31 learners in the experimental group.

As displayed in Section 4.5.1, the scores in the pre-test-post-test design could be represented in the following three steps: Score One (differences between EG's pre- and post-tests), Score Two (differences between CG's pre- and post-tests) and Score Three (differences between the EG and CG).

Score One:

As shown in Table 5.1, there were significant differences between the pre- and post-test scores in the experimental group's content and linguistic performance tests. For the content tests, the average score increased from 52.61 to 82.03, with a difference of 29.42, indicating a significant improvement in learners' content learning. Contrary to content attainments, the results of the linguistic tests were rather complex. Participants' four language skills were tested, including receptive skills (listening and reading) and productive skills (writing and speaking). As illustrated in Table 5.1, significant achievements were made in all four language skills in the target context.

Among these four skills, the average listening score held the most robust growth, increasing from 64.42 to 84.00, with a difference of 19.58 between the pre- and post-tests. This finding appears to disagree with the argument that CLIL merely benefits learners' productive skills (Bruton, 2011). Subsequent to the listening skill, the average scores of reading and speaking increased significantly, from 76.94 and 72.19 to 87.74 and 81.26, respectively. Though not as prominent as the above skills, participants' writing scores also increased from 73.42 to 78.16, with a difference of 4.74. These results offered evidence that participants' receptive and productive competences increased significantly via CLIL education. This argument corroborated the findings from those studies illustrated in the Literature Review Chapter, for example, the CLISS project (Sylvén & Ohlander, 2015) carried out among Swedish upper-secondary CLIL learners in Section 2.2.1 and Aguilar and Rodriguez's (2012) study on 87 CLIL learners in a Spanish university in Section 2.2.2.

Table 5.1 Descriptive statistics of pre-and post-tests for the experimental group

Score One	Content Test	Linguistic Test			
		Listening	Speaking	Reading	Writing
Pre-tests	52.61	64.42	72.19	76.94	73.42
Post-tests	82.03	84.00	81.26	87.74	78.16
Difference	+29.42	+19.58	+9.07	+10.80	+4.74

Score Two:

As shown in Table 5.2, the average score of the control group in the content post-test similarly climbed up to 81.31, with an increase of 29.49. With regard to the linguistic tests, the participants averaged 68.00 and 77.03 in receptive skills of listening and reading in the pre-test. While in the post-test, a significant advancement was witnessed, with the corresponding figures reaching 74.53 and 82.28, respectively. Compared with the receptive skills, the progress in the productive skills of the control group was relatively minor, even with a decrease in the writing skill. In detail, the average scores of the speaking skill increased slightly from 71.41 to 71.97, with a difference of .56. It was even noteworthy that the average scores of the control group

in the writing tests decreased from 72.78 to 72.53. The finding echoes the results of Tong’s (2012) study on Chinese university students’ English proficiency development, indicating that the mainstream EFL approach in Chinese tertiary education tends to benefit more on learners’ listening and reading competences, rather than their writing and speaking skills. Various reasons can be found in previous literature explaining this phenomenon (Cai, 2010, 2013; Hao, 2012; Tong & Shi, 2012). For example, in the traditional EFL classroom, the courses are generally teacher-centred, where insufficient attention has been paid to fostering learners’ productive language competence. Learners are not given adequate opportunities to practise productive skills. Furthermore, Chinese EFL learners’ traditional inclination to remain quiet in class might be a contributing factor as well.

Table 5.2 Descriptive statistics of pre-and post-tests for the control group

Score Two	Content Test	Linguistic Test			
		Listening	Speaking	Reading	Writing
Pre-tests	51.82	68.00	71.41	77.03	72.78
Post-tests	81.31	74.53	71.97	82.28	72.53
Difference	+29.49	+6.53	+0.56	+5.25	-0.25

Score Three:

As shown in Table 5.3, when comparing the differences of EG and CG’s content performances between pre- and post-tests, the difference of -0.07 showed no or slight statistical discrepancy between the two groups. These results reiterate some prior findings (Badertscher & Bieri, 2009; Pladevall-Ballester, 2015; Sylvén, 2017), indicating that the performance achieved in the content subject taught through the CLIL language was similar to that of the native, learners’ usual language of instruction. This result might add evidence to the statement that CLIL does not seem to hinder learners from successfully acquiring the same amount of subject content knowledge as those studying in traditional contexts (Bruton, 2011; Pérez-Cañado, 2017). In the literature of

CLIL studies, the possible reasons for this finding include learners’ affective filter, teachers’ focus on content teaching, and learners’ high demand for content knowledge, making content assimilation less challenging. In the present study, the causes of this finding will be further explored in the interview section in the following chapter.

Concerning the linguistic performance, EG participants’ language scores in the post-tests were relatively higher than the control group, with the difference of 13.05 (listening), 8.51 (speaking), 5.55 (reading) and 4.99 (writing), respectively. These findings have offered support for positive CLIL effects on learners’ linguistic performance, covering both receptive and productive skills. In other words, CLIL participants in this study generally outperformed their non-CLIL counterparts in linguistic competence, especially in speaking, which is worth discussing in the following chapter. In short, as can be expected, compared to those of content achievements, linguistic competencies and skills seem to enhance more from the CLIL approach, corresponding with the findings of the majority of CLIL studies.

Table 5.3 Descriptive statistics between the experimental and control group

Score Three	Content Test	Linguistic Test			
		Listening	Speaking	Reading	Writing
Experimental	+29.42	+19.58	+9.07	+10.80	+4.74
Control	+29.49	+6.53	+0.56	+5.25	-0.25
Difference	-0.07	+13.05	+8.51	+5.55	+4.99

5.1.2 Relationships between CLIL learners’ language competence and content performance in pre- and post-tests

In the CLIL literature, a number of studies investigated learners’ language competence and content performance separately, with little research addressing both in an integrated way, nor exploring the complex relationships between these achievements in the CLIL learning process (Pérez-Cañado, 2020; W. Yang, 2015). However, as discussed in the above sections, CLIL “takes account of integrating content learning

and language learning within specific contexts and acknowledges the symbiotic relationship between these elements” (Coyle et al., 2010, p. 14) . Consequently, this study attempted to examine learners’ content and language performance in the CLIL environment and identify relationships between these performances in terms of CLIL effectiveness in a Chinese educational context.

Concerning the control group, fewer correlations were found between learners’ content and linguistic achievements, including the multiplex relationships among speaking in pre-test (PrS), reading in pre-test (PrR), listening in pre-test (PrL) and writing in pre-test (PrW), together with two uniplex relationships between reading in post-test (PoR) and listening in post-test (PoL), speaking in post-test (PoS) and reading in post-test (PoR), as presented in Table 5.4.

Table 5.4 The results of Pearson Correlation Coefficient of Control Group (N=32)

		PrC	PoC	PrL	PrR	PrS	PrW	PoL	PoR	PoS	PoW
PrC	Pearson	1	.247	.070	.198	-.133	.161	.263	.231	.272	.015
	Sig		.173	.705	.278	.469	.379	.146	.203	.133	.933
PoC	Pearson	.247	1	.249	.163	.062	.081	.231	.256	.150	.058
	Sig	.173		.170	.374	.735	.660	.204	.157	.413	.751
PrL	Pearson	.070	.249	1	.817**	.537**	.713**	.263	.105	.105	.102
	Sig	.705	.170		.000	.002	.000	.145	.566	.568	.580
PrR	Pearson	.198	.163	.817**	1	.457**	.640**	.350	.141	.018	.005
	Sig	.278	.374	.000		.009	.000	.049	.440	.924	.980
PrS	Pearson	.133	.062	.537**	.457**	1	.745**	.007	.127	.205	.049
	Sig	.469	.735	.002	.009		.000	.971	.490	.261	.790
PrW	Pearson	.161	.081	.713**	.640**	.745**	1	.199	.035	.238	.113
	Sig	.379	.660	.000	.000	.000		.274	.851	.189	.537
PoL	Pearson	.263	.231	.263	.350*	.007	.199	1	.679**	.356*	.228
	Sig	.146	.204	.145	.049	.971	.274		.000	.046	.209
PoR	Pearson	.231	.256	.105	.141	.127	.035	.679**	1	.642**	.367
	Sig	.203	.157	.566	.440	.490	.851	.000		.000	.039
PoS	Pearson	.272	.150	.105	.018	.205	.238	.356	.642**	1	.438
	Sig	.133	.413	.568	.924	.261	.189	.046	.000		.012
PoW	Pearson	.015	.058	.102	.005	.049	.113	.228	.367	.438	1
	Sig	.933	.751	.580	.980	.790	.537	.209	.039	.012	

For the experimental group, a number of uniplex and multiplex relationships were

found statistically from the results of Pearson Correlation Coefficient, revealing diversified interactions within learners' knowledge acquisition in the CLIL context. These correlations occur among the achievements of content in post-test (PoC), speaking in pre-test (PrS), reading in pre-test (PrR), listening in pre-test (PrL), as well as speaking in post-test (PoS), reading in post-test (PoR), listening in post-test (PoL), writing in post-test (PoW), as shown in Table 5.5.

Table 5.5 The results of Pearson Correlation Coefficient of Experimental group (N=31)

		PrC	PoC	PrL	PrR	PrS	PrW	PoL	PoR	PoS	PoW
PrC	Pearson	1	.040	.284	.114	.440*	-.017	.237	.090	.205	.161
	Sig		.832	.122	.542	.013	.927	.199	.629	.268	.386
PoC	Pearson	.040	1	.849**	.796**	.514**	.233	.900**	.841**	.879**	.822**
	Sig	.832		.000	.000	.003	.206	.000	.000	.000	.000
PrL	Pearson	.284	.849**	1	.744**	.606**	.134	.972**	.839**	.870**	.804**
	Sig	.122	.000		.000	.000	.471	.000	.000	.000	.000
PrR	Pearson	.114	.796**	.744**	1	.359*	.290	.777**	.873**	.792**	.679**
	Sig	.542	.000	.000		.047	.113	.000	.000	.000	.000
PrS	Pearson	.440*	.514**	.606**	.359*	1	.061	.635**	.446*	.586**	.613**
	Sig	.013	.003	.000	.047		.744	.000	.012	.001	.000
PrW	Pearson	-.017	.233	.134	.290	.061	1	.127	.264	.198	.198
	Sig	.927	.206	.471	.113	.744		.497	.151	.286	.287
PoL	Pearson	.237	.900**	.972**	.777**	.635**	.127	1	.878**	.908**	.839**
	Sig	.199	.000	.000	.000	.000	.497		.000	.000	.000
PoR	Pearson	.090	.841**	.839**	.873**	.446*	.264	.878**	1	.857**	.802**
	Sig	.629	.000	.000	.000	.012	.151	.000		.000	.000
PoS	Pearson	.205	.879**	.870**	.792**	.586**	.198	.908**	.857**	1	.850**
	Sig	.268	.000	.000	.000	.001	.286	.000	.000		.000
PoW	Pearson	.161	.822**	.804**	.679**	.613**	.198	.839**	.802**	.850**	1
	Sig	.386	.000	.000	.000	.000	.287	.000	.000	.000	

Relations between listening, reading, speaking and writing skills in linguistic tests

In the literature investigating the effectiveness of a language learning approach, the correlation among learners' four linguistic skills, namely listening, reading, speaking and writing, are often examined. The majority of these studies (Coyle, 2018; A. Llinares & McCabe, 2020) prove that these skills are positively interconnected in the language learning process, contributing to the achievements in language knowledge. Thus, as presented in Table 5.4 and 5.5 in this study, these relationships among the four

skills were also scrutinised, and several positive correlations were identified in the language pre and post-tests for both the experimental and control groups, which partially confirmed the multiplex relationships among listening, reading, speaking and writing skills in learners' linguistic learning process.

For the control group, it was found that the achievements in speaking, reading, listening and writing were closely interconnected with each other in the pre-tests ($r=.817, .537, .713, .457, .640, .745$, Sig=.000; significant level at .05), indicating that at the beginning of the experiment, learners' linguistic skills interacted and interfered in an integrated manner. However, the multiplex relationships had changed, weakened and even vanished when CG participants continued learning in the mainstream English course. For example, only the reading skill was found connected with listening ($r=.679$, Sig=.000) and speaking ($r=.642$, Sig=.000) in the post-tests. This phenomenon might root in China's traditional EFL teaching context in which the main emphasis is given to the reading skill and grammatical points, whereas less attention is paid to those productive language skills.

As per the experimental group, correlations were found in pre-tests between listening and reading ($r=.744$, Sig=.000), between speaking and listening performance ($r=.606$, Sig=.000), indicating that these skills were closely related at the beginning of CLIL learning. A slight correlation was identified ($r=.359$, Sig=.047) in the relationship between participants' speaking and reading scores, implying that these two skills were moderately interconnected. However, it was surprising to find that no significant relationship was found between listening and writing ($r=.134$, Sig=.471), reading and writing ($r=.290$, Sig=.113), speaking and writing ($r=.061$, Sig=.744) in the pre-tests. This phenomenon reveals that CLIL participants' writing skills didn't affect how they performed in reading, speaking and listening at the outset of language learning in the CLIL context.

In terms of EG participants' language post-test scores, more interconnections were found among the four language skills, which was in contrast with that of their peer controls. For instance, it was shown that EG participants' receptive performance, namely listening and reading skills, were highly positively correlated ($r=.878$, $\text{Sig}=.000$), indicating that their listening and reading performance were integrated in a dynamic way, and influenced each other in a multi-directional pattern. So was EG participants' productive performance (speaking and writing skills) in the post-test ($r=.850$, $\text{Sig}=.000$). Besides the above two, other significant and positive relationships were found to reveal the close interrelationships between receptive and productive skills. For example, listening skills correlated positively with speaking skills ($r=.908$, $\text{Sig}=.000$) and writing ($r=.839$, $\text{Sig}=.000$), respectively. These positive correlations also applied to reading and speaking skills ($r=.857$, $\text{Sig}=.000$), and reading and writing skills ($r=.802$, $\text{Sig}=.000$). These findings indicated participants' four linguistic skills were closely interwoven in the CLIL context. They partially accord with Yang's (2014) study of 92 Taiwanese CLIL learners, where positive correlations among learners' language skills were found to exist within the system of knowledge acquisition in the CLIL context. Besides, it is noteworthy that the writing skills in the post-test were closely connected with other skills, which is different from those in the pre-tests. This finding implies that writing practice is an essential part in CLIL language learning, involving various aspects of understanding grammatical progression, using language in authentic and interactive settings and developing communicative skills.

Relations between pre-tests and post-tests in content and language

In this part, the relations between the pre-tests and post-tests were examined with regard to CG and EG participants' content and language performance. As per content performance, it was found that no statistically significant connections existed in the pre-test and post-test results for both the control group ($r=.247$, $\text{Sig}=.173$) and the experimental group ($r=.040$, $\text{Sig}=.832$). This implied that those participants, no matter in the mainstream EFL settings or the CLIL context, who performed better at the very

beginning of content knowledge acquisition might not have generally performed better in the post-tests, and vice versa. This finding was in accordance with Aguilar and Rodríguez's (2012) study illustrated in Section 2.2.2. It is noted that in various CLIL settings, learners' content performance may not be significantly influenced by their entry competence in content knowledge.

As per language performance, the results showed that no relationship existed between learners' linguistic performance in the pre- and post-tests in the control group, indicating that learners' entry linguistic competence might not affect their achievements after a period of learning in the mainstream EFL course. In contrast with the control group, the interrelationships of the experimental group were rather complex. A number of uniplex and multiplex relations were found statistically. Firstly, the results showed that CLIL participants' four language skills in the pre- and post-tests were closely interconnected with each other. For example, the listening performance between the pre- and post-test ($r=.972$, $\text{Sig}=.000$) had a strong correlation. This positive correlation applied as well to EG learners' reading scores ($r=.873$, $\text{Sig}=.000$), indicating those who had better achievements in the receptive skills (listening and reading) in the pre-test tended to outperform others in the post-tests, and vice versa. In terms of productive skills, the results showed that the correlation between learners' speaking performance in the pre- and post-test was relatively slight ($r=.586$, $\text{Sig}=.000$). In addition, no significant correlation was found between the scores of the two writing tests ($r=.198$, $\text{Sig}=.287$). Both results showed that CLIL learners' entry productive competence in the pre-tests had slight or even no impact on how they performed in the post-tests, which was different from receptive skills.

Furthermore, it was found that some language scores in the pre-tests were related to those in the post-tests, such as a strong correlation between listening performance in the pre-test and reading performance in the post-test ($r=.839$, $\text{Sig}=.000$), a positive correlation between reading performance in the pre-test and speaking performance in

the post-test ($r=.792$, $\text{Sig}=.000$), as well as a moderately positive correlation between reading performance in the pre-test and writing performance in the post-test ($r=.679$, $\text{Sig}=.000$). These interrelations made learners' language competence in the CLIL context complex, further revealing that CLIL learners' four language skills are closely interwoven and interacted, as articulated in the previous section.

Relations between content and language performance

According to Table 5.4, no significant relationship was found between CG participants' content performance and linguistic achievement in both pre- and post-tests. This implied that non-CLIL participants achieved their content and language outcomes separately without interfering with each other. This was due to the fact that CG participants learned content and language in two separate settings, namely the content course in L1 and the English course in L2.

Concerning the relations between language and content performance of the experimental group, two different results have been identified in the pre- and post-tests, as presented in Table 5.5. In the pre-test, no definite correlation has been identified between language proficiency and content performance (listening: $r=.284$, $\text{Sig}=.122$; reading: $r=.114$, $\text{Sig}=.342$; speaking: $r=.440$, $\text{Sig}=.015$; writing: $r=-.017$, $\text{Sig}=.927$). It is indicated that higher English proficiency might not promise a better content performance at the outset of CLIL education and vice versa. At this stage, content and language learning were seemingly independent from each other rather than integrated in the learning process.

However, the results in the post-tests exhibited a different picture. In contrast with that of CG participants, EG participants' content performance correlated closely to their language competence, namely listening skill ($r=.900$, $\text{Sig}=.000$), reading skill ($r=.841$, $\text{Sig}=.000$), speaking skill ($r=.879$, $\text{Sig}=.000$) and writing skill ($r=.822$, $\text{Sig}=.000$). It indicated that, on the one hand, those CLIL participants who performed better in the content subject might achieve better performance in language learning; on the other

hand, those proficient language learners might get higher achievements in the content subject. These findings accord with Crandall's assertion that "learners cannot develop academic knowledge and skills without access to the language in which that knowledge is embedded, discussed, constructed, or evaluated. Nor can they acquire academic language skills in a context devoid of academic contexts" (1994, p. 256). The aforementioned results confirm what has been found in previous CLIL literature that a close correlation exists between content and language performance, as reported in Yang's (2014) study on Taiwanese university learners in Section 2.2.3, and Jäppinen's (2005) study on Finnish CLIL learners aged 7-15 in Section 2.2.1.

Along with the findings, related reasons have been explored in the literature, and so did the present study, for example, the way of assessment. In CLIL settings, learners' proficiency should generally be assessed within a multi-variable framework, taking account of domain-specific items and the use of language in a specific context like the CLIL class. Further discussion will be provided in the following sections, together with qualitative findings in the interview responses.

5.2 Qualitative findings from Phase Two

Approximately two weeks after the post-tests (around the end of January 2018), ten randomly selected participants from the experimental group were interviewed face-to-face by the researcher in a café on campus, providing descriptive and nuanced data on the effectiveness and limitations of the CLIL approach in the target context (as displayed in Table 5.6). Before administering the interviews, a list of guiding questions was prepared. The design of these questions mainly hinged on two sources. The first was the theoretical analysis of the 4Cs framework and related CLIL studies as reviewed in Chapter Three. The second source was the quantitative data collected in Phase One, namely the scores of CG and EG participants' performances in content and language in pre- and post-tests. The initial statistical analysis of the numeral information was used to inform the interview design in Phase Two. Some noteworthy phenomena and outstanding issues as presented in Section 5.1 were included in the interview questions

for further investigation, such as the impact of language proficiency on content performance. All the interview questions were designed to capture CLIL participants' perceptions of CLIL learning experience, insights of learning contexts, personal preference of CLIL learning, and related background information. The purpose of these questions was to further elicit participants' understandings and perceptions of CLIL learning, and figure out the reasons for these understandings and perceptions. Meanwhile, the CLIL teacher's reflections were also collected to investigate correlated factors from the teacher's perspective and provide insights into CLIL learning and teaching from a different angle.

Table 5.6 Background information of interview participants

Name	Gender	Pre-test					Post-test				
		Con tent	Language				Con tent	Language			
			L	R	S	W		L	R	S	W
Adam	Male	61	67	75	70	78	83	85	87	81	77
Brenda	Female	51	70	80	86	88	89	88	90	84	84
Charles	Male	43	72	91	70	84	92	90	93	88	77
Daphne	Female	65	61	72	76	67	76	79	84	78	74
Elaine	Female	34	72	86	80	79	95	92	97	89	91
Fiona	Female	54	63	71	63	66	62	77	83	77	70
Gilbert	Male	64	64	68	70	74	68	75	80	73	68
Hannah	Female	65	73	84	68	90	93	91	91	88	88
Ian	Male	60	69	82	66	70	87	87	89	83	81
Jeffery	Male	71	58	71	81	74	64	76	82	77	74

5.2.1 Analysis of learners' interviews

During the process of data analysis, various themes and concepts emerged from the qualitative data, which not only indicated participants' perceptions of their achievements in the CLIL context, but also provided some potential explanations for their achievements. Similar themes and concepts were merged and grouped into related

categories for easier management and queries on the data. Among these categories, some matched well with the concepts in the specific components of the 4Cs framework of CLIL, for example, ‘linguistic improvement’, ‘English writing’ in Communication, ‘remembering’, ‘understanding’, ‘applying’, ‘analysing’, ‘evaluating’, ‘creating’ in Cognition, ‘content assessment’ in Content, ‘cultural awareness’ and ‘otherness’ in Culture. While other categories, such as ‘the use of L1’, might not refer directly to the four key components in the CLIL framework, but they correlated closely with CLIL learning, highlighting various external and internal factors that significantly influenced learners’ CLIL learning process.

‘Linguistic Improvement’

In the interview responses, the participants generally endorsed the viewpoints that CLIL improved their language capabilities, covering learning and using, and related concepts, like ‘improvement’, ‘improves’, and ‘improving’ appeared frequently. For example:

“My linguistic performance? Yes, I have achieved a great improvement in my English performance this semester. Look at my scores in the final tests (language post-test), 91, 91, 88 and 88. A very good performance, isn’t it? Yes, I’m satisfied with my performance this term. Surely, I will make further improvement in the future” (Hannah).

“I think I improve a lot in my language, especially in how to use English in the real-world practice” (Elaine).

These comments not only confirmed what has been found in the quantitative data that CLIL facilitated learners’ language acquisition, but also pinpointed the fact that CLIL participants felt satisfied with language learning and using in the dual-focused context of content and language. Thus, the concept of ‘linguistic improvement’ in the study, closely related to language learning and using, reflected learners’ understanding of both linguistic and communicative competence in the CLIL learning process.

‘Cognitive Engagement in CLIL Learning’

The concept of ‘engagement’ was also frequently cited by the participants in the interviews, showing their perceptions and affective states of the CLIL approach that combines content learning and language learning within specific contexts.

“CLIL makes me reflective on my knowledge acquisition, especially in writing English essays about content knowledge. I invest more effort and time into learning and using language, just because it is very interesting to learn financial knowledge in an English way with different expressions, sentence patterns, and even different ways of thinking. I think it is good, er, for me”
(Charles).

“Engagement, yes, it could be a kind of engagement. I am very pleased to attend the CLIL class that makes language learning easy and interesting. This programme is very new to me, but useful, especially the dual-focused approach. I learn to use English and use English to learn. I haven’t attended this kind of content learning before, learning finance in English or learning English through finance. It’s amazing. I involve more than before in my study”
(Elaine).

These interview responses exhibited a notable phenomenon that there was a strong correlation between participants’ perceptions of the CLIL approach and their engagement in the process of knowledge acquisition. When participants perceived significant benefits of CLIL and were confident that CLIL could improve their English proficiency and content learning, they cognitively contributed their time and energy to their learning process. They attempted to develop strategic plans as well as reflective evaluation. Hence, in the CLIL context, ‘cognitive engagement in language and content learning’ is associated with participants’ cognitive processing, acknowledging the symbiotic relationship that exists between these components of CLIL.

‘Importance of English Writing Proficiency’

In addition, six of ten participants mentioned the concept related to ‘writing’ in their interview responses, such as ‘writing skill’, ‘writing ability’ and ‘content-focused writing’. For example, Elaine and Ian emphasised the importance of writing ability in the CLIL context, stating that

“Writing ability is very important when learning financial management in English. I know that. In the CLIL class, I practice essay-writing more. I have learned to write such academic essays. I think English writing is not just a simple writing task to me, but a way of thinking, a solution to solve the problem. Certainly, that’s not enough for me. I want to learn more about CLIL. This is an amazing approach for we Chinese students to learn English and finance together, in the same course” (Elaine).

“I have never written such a great number of essays in one course before. Let me count. Yes, altogether, we wrote more than 30 short essays and two long ones in English, almost two essays for one week. I have improved a lot in my writing skill. It’s great. With more language exposure and more writing tasks, I have made the greater progress in English” (Ian).

These two participants expressed higher agreement with writing tasks in the CLIL context, highlighting their linguistic and communicative competence improvement. According to interviews in this study, most participants’ qualitative responses echoed the previous quantitative data analysis (an increase from 73.42 to 78.16 in writing tests). Participants perceived significant progress in their writing competence and confirmed that CLIL could improve their English writing proficiency in a language and content dual-focused environment. Thus, ‘writing’ in this section could be understood as learners’ awareness of the importance of English writing proficiency in the CLIL context.

‘Initial Anxiety in CLIL learning’

However, not all the interview responses to the CLIL approach were positive in the study. It was reported that some participants lacked confidence in this content and language integrated learning and was anxious in engaging with the two subjects simultaneously, especially at the beginning of the intensive exposure to CLIL. The cause of this phenomenon might be explained in a two-tiered way. Firstly, in CLIL settings where using language to learn is as important as learning to use language, learners may not communicate as competently as in their native tongue, leading to a high level of anxiety in language learning and using.

“I was anxious in the classroom. I was not confident in speaking English in front of my classmates. They would laugh at me when I made mistakes. Sometimes, I found it difficult to learn financial management in English” (Gilbert).

“I’m anxious to use English as a tool when learning finance. I think my English proficiency is not good enough, especially my speaking ability. Sometimes, I don’t know how to express my ideas clearly. Though I know the right answers to the questions, I’m still nervous about expressing them in English. I think I still need to work hard” (Daphne).

These responses revealed that in the CLIL environment, participants might be anxious not only in linguistic competence, but also in communicative competence at the beginning of CLIL learning. English was merely an independent course to these participants, just like chemistry, geography, and business management. The target CLIL course was challenging to learn because it was related to learning content as well as reconstructing the content and related cognitive processes in knowledge acquisition. In other words, content and language learning did not appear to be integrated but split in their opinions.

In addition, some participants reported that the CLIL approach made them anxious in the process of knowledge acquisition, especially in terms of content learning. They

attributed this result to the fact that in the CLIL setting, the non-language subject was not only learned in a foreign language, but also with and through a foreign language.

“I think it is very difficult to learn finance in this way. I couldn’t understand the teacher’s instructions (in English) very well. How could I say that? I got the meaning of each word in the instructions, but I failed to get the meaning of the whole sentence when putting these words together. And even worse, I was anxious in learning those so-called terminologies. I couldn’t understand them in English” (Jeffery).

“Sometimes, I got troubles in learning content knowledge. It was too difficult for me. Though the teacher explained the knowledge issues again and again, I still felt confused, er, about content issues and language points. I don’t know why. The teacher just kept lecturing the courses in English and did not mind if my English was not good enough to understand the content. There were many technical concepts and terminologies, which I felt were very difficult when they were taught at the very beginning. And another problem is even when I understood the questions, I didn’t know how to answer the questions” (Daphne).

These responses, in turn, imply that content learning was still a challenging task for some participants to fulfil in an integrated approach due to various internal and external factors. For CLIL learners like Jeffery and Daphne in the present study, the content and language integrated approach made the learning task more demanding and somewhat painful, making them anxious in learning. Unlike other learning approaches, the CLIL participants found that content and language integrated learning is far more than simply learning a non-language subject in an additional language. It is not the same as learning in the mother tongue. In CLIL settings, acquiring content knowledge and related skills is far from enough. More importantly, learners need to understand, develop and create their own knowledge and skills, leading to cognitively demanded

subject learning. In this sense, CLIL learners' 'initial anxiety in CLIL learning' showed their awareness of their own learning process and sometimes made them cognitively engaged in the process of knowledge acquisition.

'The Use of L1 (Chinese) in CLIL Learning'

It was surprising to find that 'L1' (Chinese) was frequently mentioned in participants' responses to CLIL learning, especially about content knowledge acquisition. As follows, some participants expressed strong tendency to resort to their native language in the process of content learning, as the instruction in L2 might hinder their content learning. As Adam stated,

"I find it difficult for me. Sometimes, I needed to use Chinese to understand these concepts. Certainly, I do not mean that the CLIL approach is useless for me. But could the teacher explain those confusing and tedious concepts in Chinese, too? It seems better for me".

His statement was echoed by his classmate, Fiona, who stated that

"I don't contend that content course should be taught in our L1 only. But on some occasions, it was easier for me to learn the course in L1 since most of us are competent in the Chinese language. In addition, I couldn't explain the answers to the questions clearly and consequently achieved a rather poor performance in the content test".

From these responses, it was apparent that some participants held a more sceptical attitude towards content-learning in the CLIL context. They believed that quite often, the instruction through the CLIL language could be presented in the way of bilingual blended instructions, with code-switching between languages and systematic use of both the native and the CLIL language. For example, in the present study, sometimes the L1 Chinese might be used for delineating and summarising the key points, while the L2 English for the remaining class functions. This type of code-switching is termed as 'translanguaging', leading to a dynamic form of bilingualism in the classroom.

'The Format of Content Assessment'

The way in assessing learners' knowledge acquisition becomes one of the core concerns of evaluating the effectiveness of the CLIL approach. In the present study, the design of pre- and post-tests was adopted, and the results of these tests quantitatively revealed learners' academic performance in both content and language. Based on the qualitative data from the interview responses, it was found that seven participants showed great satisfaction with linguistic assessment and perceived the methods of assessing language could effectively and successfully evaluate their language competence. However, participants' opinions diversified on investigating their attitudes toward the tests assessing content competence. Some participants were generally satisfied in certain parts of 'content assessment', such as a paper-and-pencil test, grading standards and scores.

"I think it's fine to have these paper works (tests). I hate those so-called oral tests in which I can't get good marks' (Brenda). 'It's okay, er, about the tests. Though they were very difficult for me, they assessed my performance delicately and precisely. The marker, you mean, the teacher? Yes, it's okay, too. She gave me the appropriate score'" (Charles).

In comparison, negative attitudes towards the assessment of content knowledge also emerged.

"It was unfair of this kind of test, only paper-and-pencil test. I think it couldn't fully assess our achievements. And also, I think that the teacher marked the test paper in the wrong way. She just did it by her own subjective impression of my classmates'" (Gilbert).

In this sense, some participants questioned the paper-and-pencil test in content knowledge and considered that the teacher graded learners' content performance subjectively. According to interviews in this study, the participants noted that the grading of content tests involved some non-academic elements rather than their content

proficiency, like learners' attitudes, classroom interactions and behaviours, and the teacher's familiarity with the students. Different understanding and perceptions of assessment in CLIL might impact learners' ways of learning, leading to diversified academic outcomes.

'Cognitive Processes'--'Remembering', 'Understanding', 'Applying', 'Analysing', 'Evaluating' and 'Creating'

Based on the reinterpretation of the 4Cs framework in the Literature Review Chapter, it can be contended that, for both language and content learning to be effective, learners need to be cognitively engaged and actively think, control and articulate their own learning in the CLIL process, highlighting “a transparent connecting of thinking process to knowledge construction” (L. W. Anderson & Krathwohl, 2001, p. 67). In this study, CLIL learners tended to show a strong capability in cognitive control and ‘reflect’ their own learning process by considering both content and language learning together with the growth of various thinking and problem-solving skills. The concepts of ‘remembering’, ‘understanding’, ‘applying’, ‘analysing’, ‘evaluating’ and ‘creating’ were frequently mentioned in the interview responses. These themes accorded with Anderson and Krathwohl's (2001) conceptual framework of cognition, as discussed in Section 3.1.3.

‘Remembering’ refers to the period of cognitive learning “when memory is used to produce or retrieve definitions, facts, or lists, or to recite previously learned information” (L. W. Anderson & Krathwohl, 2001, p. 67). As one of the lower-order thinking skills, ‘remembering’ was adopted by the CLIL learners to produce appropriate information for their existing memory. Due to the exhibition of broader knowledge of language and content, CLIL participants in this study had to remember or retrieve knowledge from ‘long-term memory’, not only by recognizing or identifying the financial knowledge learnt in the classroom, but also by recalling or retrieving the knowledge and attempting to apply it to their daily learning. As Hannah stated,

“I have to update the information in my memory every day. I have to recognise those definitions taken from financial management and recall them in the classroom as soon as possible”.

A further indicator of the CLIL effect is that by combining language and content knowledge, CLIL learning might be more meaningful, facilitating the construction of meaning, namely ‘understanding’ in the cognitive process. The entwined use of two types of knowledge might help participants contemplate related but different issues. Firstly, the process of CLIL learning contributes to interpreting knowledge, an initial step leading to understanding. In the interview, Elaine’s response supported that the coexistence of the two kinds of knowledge seemed to yield to a sounder interpretation of learning content (language and content).

“It was not a repetition of the same things. I had to think about these concepts from two similar but different views, may be called two different codes. And this may be the reason that I performed better in the CLIL context”
(Elaine).

Another inherent feature of CLIL that leads to understanding is exemplifying. In the CLIL context, the participants were provided examples of related knowledge, especially from cultural perspectives.

“In class, we studied the concept of interest volatility. It was quite complex for us to understand. Therefore, the teacher introduced China’s and Britain’s national interest policies, which were quite different. It was interesting, and I asked for more examples” (Daphne).

In this way, the teacher brought the target knowledge to daily life and provided examples of cultural aspects, which helped CLIL participants analyse and explain complex terminologies. This way of transmitting knowledge in the CLIL class awakened participants’ curiosity and asked for more examples, which reflected the value of exemplifying.

The subsequent cognitive step to understanding is ‘summarising’ in Anderson and Krathwohl’s (2001) cognitive process framework. ‘Summarising’ was also revealed in this study, where CLIL participants reported in interviews their capability to summarise the main knowledge of content and language learning effortlessly.

“Studying in the CLIL context may reinforce the knowledge growth. Both types of knowledge complement each other. By revising and summarising all the knowledge learned in class, I am able to remember and understand these concepts. I realise that English helps me learn finance equally well” (Charles).

It is considered that CLIL puts the cognitive procedure into real-world practice in a given situation, namely ‘applying’. By integrating language and content knowledge, CLIL learning facilitates learners to implement related knowledge they learned and the cultural aspects they experienced in new and concrete situations. In the present study, CLIL participants appeared to excel at applying the cognitive procedure to a novel task and context, developing procedural and metacognitive knowledge in a similar way.

“I come into contact with those foreign students on campus. I pay attention to what they do in their daily life, for example, their lifestyle, food, hangouts, customs at school, holidays, everyday aspects close to them. I try to apply what I have learned in the CLIL course to daily life because I think finance could not be limited to the knowledge in the classroom, but the practice in daily life. Of course, the financial issue in foreign countries is one of my favourite topics.” (Charles).

Charles’s view suggested that within the CLIL context, learners not only had access to developing their cognitive thinking processes, but also raised their awareness to apply or implement the linguistic and content knowledge to the new environment.

CLIL learning appears to foster the thinking process of ‘analysing’, segmenting a concept into parts and explaining how the parts relate to the whole. In the present study,

in the learning process of CLIL knowledge, participants presented a solid tendency to differentiate among the different parts and organise the elements to find coherence.

“When I couldn’t understand the complex descriptions, such as the relationships between current assets and current liabilities, I had to divide these descriptions into shorter and easier parts and tried to find some similar explanations in Chinese subsequently. Then, I organised the Chinese versions of the description. It was sometimes hard for me to understand them as a whole in English.” (Elaine).

Elaine’s reflection on the cognitive thinking process of ‘analysing’ indicated the interrelated use of Chinese and English language in the CLIL context, correlating with her improvement in metalinguistic competences and knowledge. Besides, in analysing the conceptual knowledge in the CLIL context, participants seemed to be able to enhance the cognitive process of analysing by deconstructing their existing knowledge of language and content, and sometimes cultural aspects, and more importantly, deciding “bias, values or intent underlying the course they work with” (L. W. Anderson & Krathwohl, 2001, p. 68).

“The idea that there are other ways of representing financial formula is an excellent starting point. I was apparently disoriented at the beginning. The different perspectives discovered in the course allowed me to change the fixed ideas about these conceptions, definitions and schemes. Thanks to the CLIL programme” (Daphne).

Higher-order cognitive processes might come up when examining and evaluating CLIL learners’ cognitive development. ‘Understanding’, ‘applying’ and ‘analysing’ converge into ‘evaluating’, the fifth cognitive process dimension in Anderson and Krathwohl’s (2001) cognitive framework. The initial move in evaluating is to check or seek inconsistencies or fallacies. An example in the study was that some participants realised that there was no one-to-one correspondence between terms in both languages, eliciting the search for further communicative strategies and led them to rethink the

learning process.

“I found that some expressions could not be literally translated. This allows us to look for more strategies of learning and new ways of communicating” (Brenda).

Besides, learners might criticise their own learning process in the cognitive function of evaluating. As Méndez García (2014) explains, this capability of criticising is ascribed to CLIL education that makes learners generally broad-minded.

“I think, as a CLIL learner, I’m much more open. I like to question more frequently. Maybe it is not right. But I think they (non-CLIL learners) tend to be more narrow-minded. They prefer rejecting things, especially new ones, probably out of fear of the unknown” (Gilbert).

Creating, the highest order cognitive process in Anderson and Krathwohl’s (2001, p. 68) framework, enables learners to “reorganise elements into a new pattern or structure” when it comes to generating or hypothesising. In the CLIL context, learners are allowed to establish their hypotheses on cultural and language differences. For example, some participants in the present study tended to pose questions to find an answer to why interest rates varied considerably between the UK and China. In this way, language-related hypotheses were likewise generated due to the co-existence of language, content and sometimes culture.

“I came to understand how those English people calculate their interest rates. I asked some questions about why the bank interest rate in the UK was so low. The teacher didn’t answer me but encouraged me to find the answer by myself. Therefore, I made several hypotheses and strived to explain my hypotheses in my own language (Chinese) on my own” (Charles).

In sum, the CLIL learners in this study were cognitively challenged to develop their content and language knowledge, and all these concepts of ‘Remembering’,

‘Understanding’, ‘Applying’, ‘Analysing’, ‘Evaluating’ and ‘Creating’ could be grouped into ‘cognitive processes’.

‘Intercultural Understanding’

In the CLIL contexts as discussed in Section 3.1.4, the focus of cultural awareness may shift from knowledge of different cultures to intercultural understanding involving different cultural experiences. The key objective of CLIL is to provide an effective incentive for real-world intercultural experiences, entrenching learners’ understanding of global mobility and citizenship. In this study, the concept of ‘awareness’ was cited frequently, highlighting participants’ acquisition of a deepening ‘cultural awareness’, as illustrated in Section 3.1.4.

“It is really good for us to use these authentic materials. Look at these examples. I am aware that it’s not only a financial issue in the textbook, but a real-world example from an English-speaking country.” (Daphne).

From these points of view, it is noted that the adoption of suitable authentic materials and cross-cultural curricular linking in the CLIL classroom can result in a more sophisticated and entrenched understanding of cultural differences and similarities.

In addition, an introduction to a variety of foreign cultures can be considered by CLIL learners as an attractive source of information, which might motivate their cultural awareness in the classroom.

“In the class, the teacher showed us some American and British coins. En, Canadian, maybe. Penny, nickel, dime and quarter. It was the first time for me to notice these terms. It was rather different from what I experienced in the former class (in Chinese). Cultural awareness? Yes, I think so, because different coins told me different stories of their own culture” (Jeffery).

“The teacher showed us a number of videos in which international guest

speakers would introduce many funny things about the world. It was interesting to hear their viewpoints about their countries, and sometimes the whole world. It deepened my awareness of these different countries and their cultures” (Charles).

According to the two excerpts, CLIL participants revealed their viewpoints that the culture-related learning experience enhanced their recognition of the latest developments in the globalised world. According to their reflections, this CLIL course granted CLIL learners opportunities to look into discrepant world cultures and an outside view of their routine context. Consequently, CLIL, with the dual foci nature of language and content, provides learners with intercultural experiences that might not be easily obtained in a monolingual context. The CLIL education provides learners with abundant intercultural experiences, deepening their understanding of global mobility and citizenship. In CLIL settings, fostering learners’ intercultural understanding is initiated by experiencing and understanding the CLIL language and content in a dual-focus way. In this sense, the term ‘cultural awareness’ was modified into ‘intercultural understanding’ in the present study, relating to the willingness to alter one’s viewpoints about diversified cultures, to discover and understand otherness, and to develop intercultural communicative skills which support CLIL learners into becoming “competent global citizens” (Byram & Wagner, 2018, p. 141).

‘Otherness’

As discussed in Section 3.1.4, CLIL possesses the potential to better enhance learners’ intercultural understanding than conventional content or language learning. Its combination of language, content and cognition offers an ideal context to promote reflection, insight and self-awareness. In this way, CLIL allows the learners to transcend their own experience and foster a perspective of ‘otherness’. However, this claim may not be as solid as it was expected in the target classroom practice. For example, Gilbert stated that

“I don’t even know enough about my own culture, so it’s hard for me to compare with others. I know the word ‘otherness’, referring to people in different cultures. But I cannot make clear about ‘self’ and ‘otherness’. This may be because we don’t have enough intercultural-related materials to contact with otherness. We only have a textbook”.

“I’ve been told that CLIL could enrich our understanding of our own culture and others. Yes, I encounter different cultures that enhance my understanding of so-called inter-culture or multi-culture. But to tell the truth. I only know these cultures but cannot get into them. It’s like a picture on the wall that you can touch but cannot get into it” (Adam).

“Unfortunately, everything is measured in exam success, so there’s little time for us to develop a love or appreciation for the culture of other countries” (Hannah).

These responses on ‘otherness’ show some participants’ unfavourable attitudes towards classroom practice in CLIL. As noted by Coyle (2013), the journey to successful CLIL is not always straightforward or immediate. Further discussions and implications will be given in the following chapter.

5.2.2 Teacher’s feedback

To triangulate the validity and reliability of the findings from learners’ interviews, the teacher’s feedback on the effectiveness of the CLIL approach was collected and analysed. The teacher, female, is around forty years old. She received her postgraduate education in a British university and has taught finance-related courses for more than a decade in a Chinese public university. In this sense, she is capable and professional in implementing the target CLIL course. Four major themes emerged in the teacher’s interview feedback, indicating her perceptions about CLIL and her professional roles. The themes will be presented in the following sections, beginning with the teacher’s teaching orientation.

‘Content Orientation or Language Orientation’

In line with the CLIL definitions articulated in Section 2.1, content goals are prioritised by both CLIL policymakers and practitioners (Coyle et al., 2010; Christiane Dalton-Puffer et al., 2014; David Marsh & Frigols, 2013). In this way, most CLIL teachers agree that content learning is in advance of L2 learning in CLIL contexts (Skinnari & Bovellan, 2016, p. 148). In the present study, this priority was also verified in the teacher’s feedback, who stressed that content learning was necessary for all learners irrespective of their L1 or L2. As expressed by the CLIL teacher:

Content is stated in the syllabus as the primary goal. Thus, according to the syllabus, it’s very much a sort of content-driven. It’s like that they (learners) have to have the same (content) knowledge at the end as the ones who are studying in Chinese.

‘The same knowledge’ revealed the primary objective of CLIL. Content remains to be in centre in teaching objectives, and language should be taught securely alongside content-related concepts and skills. However, in real-world practice, the priority might be compromised in the target context due to various contextual variables like teacher availability, language support, learners’ age and social demands. As claimed by the teacher:

“It is a consideration that when you use or spend some of your time to work with communicative skills, you have to take the time from other activities that you fully used to advance in the content knowledge. Thus, there’s always small fear inside me on whether I have done more or less in the CLIL class than the non-CLIL class because I may take some time to explain some language issues”.

In this feedback, the teacher expressed her concerns on not meeting the content goals in CLIL, for example, the ‘small fear inside her’ caused by the extra time needed

for ‘explaining some language issues’. This feedback reflects the teacher’s concern for her students’ L2 learning and highlights her sense of responsibility to teach language issues. Since language cannot be described as ‘a side-effect’, ‘by-product’ or ‘a spice’ in a dual-focused approach, the teacher made attempts to combine language and content in action and placed enhanced communicative skills as central goals in CLIL. As she remarked:

“For me, it was a combination of both (language and content). What I really appreciate are the means of language teaching that I can use in teaching finance.”

From the interview, it is evident that the teacher had an interest in both language and content. She drew a clear line and included both content and linguistic knowledge into her teaching. In other words, the above feedback implied that the teacher believed that integration of content and language happened in the classroom in a natural and dynamic way and therefore separating them from each other would be artificial. This is in line with the understanding of CLIL integration as a process where language and content are entwined and cannot be fruitfully distinguished from each other (Barwell, 2016).

‘The Use of L1’

“Often, you would hear translation going on in the classroom or during something. You know, if they (CLIL learners) were trying to work out something, there would be a lot of Mandarin and a lot of note-taking, and their notes were in Chinese.”

As indicated from the interview, the teacher noted the potential relevance of students’ L1 use in the learning process, either observing students making active use of their L1 or offering occasional translations when learners were incompetent by themselves. In the present study, the teacher regarded the use of L1 positive, helping students with their work, especially at the beginning of the course. In addition, she

further commented the use of L1 in classroom instruction by stating:

“To me, the L1 was inevitable when I was teaching content knowledge. I had to explain some key terminologies (in L1), introduce some cultural issues that learners couldn’t access, and make some instructions clear and acceptable.

To the teacher, the L1 was inevitable and could serve helpful purposes. It could be used to explain cognitively demanding concepts and lexis, classify instructions, deal with disciplinary issues, foster learners’ metalinguistic and cultural awareness by the contrast between the L1 and L2, teach learners with limited L2 resources.

‘Cognitive Benefits of Integration’

According to the interview, the CLIL teacher found benefits and set goals beyond the apparent content and language learning targets. There were more general pedagogical aims such as enhancing communication, cultivating understanding or employing new learning methods. In the present study, the teacher found that CLIL might enhance students’ cognitive skills, for example, by offering them new schemes for thinking.

The CLIL education can elicit a change of mind inside my students. When they learned the subject in an L2, it was a more challenging thinking structure. Their linguistic and content skills might progress at different rates. Working in different languages could improve their thinking skills, somehow the patterns and ideas behind the language.

The above feedback revealed that CLIL participants’ mental skills and cognitive thinking were challenged and liberated by using L2 in the CLIL settings and hence improved. The teacher reasoned that the challenge of using an L2 enhanced understanding because it slowed down the learning process and led to a better focus and further elaboration on the topics to be learned. Thus, with more participation in

communicative activities, learners experienced better learning by cognitively employing various learning strategies and thinking skills.

‘Cultural Influences’

In the present study, some conflicting viewpoints appeared in the teacher’s reflection, indicating her belief and worry about the intercultural component in the CLIL context. From the teacher’s point of view, learners’ intercultural understanding kept strengthening when content, language, thinking and culture were constructed through interactions in the CLIL context. In the classroom practice, within an amount of meaningful and challenging activities and tasks, cultural patterns, customs and ways of life were expressed in language and content. Subsequently, cultural-specific world views were reflected in language and content, leading to deepening intercultural awareness and understanding. Thus, the teacher held a positive attitude towards the changes in learners’ intercultural understanding in the CLIL learning process and attempted to integrate language learning and content learning at cognitive and cultural levels appropriate to the learners.

It (The CLIL course) has brought learners into talking English. It has brought in a lot of new cultures, other aspects of teaching... Looking at different countries and how they learn finance helped my students a lot to broaden their horizons. CLIL has brought a new pedagogical culture. But that’s not enough. We don’t have enough materials from the real-world practice.

The teacher believed that she brought a new, more communicative and cooperative culture to an educational setting where she had a strong tradition of working alone. She also mentioned the possibility of integrating with international or English communities outside the school, which resulted in the development of intercultural attitudes and the internationalisation of learners. However, she often complained about not having enough resources for intercultural teaching. All these perspectives convey the idea that CLIL learning works best in communities with cooperation and support from different

cultures.

Summary

This chapter has reported the findings from the pre- and post-content and language tests in Stage One, revealing the effectiveness of CLIL on learners' academic performance in both content and communication. Firstly, the quantitative data extracted from the content tests showed insignificant differences between the control and experimental groups in the pre- and post- content test scores, indicating that in terms of content achievements, CLIL does not naturally lead to good performances and may not definitely yield better results than other approaches. Concerning linguistic proficiency, the results revealed that both receptive and productive skills were enhanced in CLIL education, displaying a better language performance than non-CLIL learners. Moreover, a number of interrelationships and interconnections were found among content performance and language achievements. These findings support Coyle et al.'s (2010, p. 41) assertion that "different components of CLIL and their interrelationships should be considered in order to form a conceptual map for understanding CLIL".

In addition, this chapter has reported the findings from the interviews in Stage Two of the present study, revealing the CLIL effects on learners' academic performance from language, content, cognitive and cultural aspects. Several concepts were extracted, namely 'Linguistic Improvement', 'Cognitive Engagement in CLIL Learning', 'Importance of English Writing Proficiency', 'Initial Anxiety in CLIL learning', 'the Use of L1 in CLIL settings', 'the Format of Content Assessment', 'Cognitive Processes', 'Otherness' and 'Intercultural Understanding', all of which have laid the foundation for discussing the effectiveness of the CLIL approach in Chinese tertiary contexts. These concepts highlight practitioners' insights into the CLIL approach in the Chinese tertiary education context. Besides, the CLIL teacher's interview was included in the study to show her perceptions of CLIL effectiveness in the target context, triangulating the findings from the learners' perspective.

By combining and reflecting on both quantitative and qualitative data, the next chapter aims to shed light on the author's understandings of the effectiveness of the CLIL approach in the given context, providing some practical implications for both policymakers and practitioners.

CHAPTER SIX – Discussions, Conclusions and Implications

As articulated in the previous chapters, CLIL is proposed by some Chinese researchers and educators as a potential solution to globalisation and an alternative to the current mainstream EFL approach, given the rapid and widespread practices in European contexts and the advantage of combining foreign language enrichment measures into content teaching. However, fears and uncertainties emerged due to the paucity of empirical evidence in Chinese educational contexts. To defuse both policymakers and practitioners' suspicion, the present experimental study investigated CLIL effects on learners' academic performance in a Chinese tertiary education programme, aiming to provide some empirical and hard evidence for implementing CLIL in Chinese higher educational contexts.

Based on the literature review of CLIL studies in Chapter Two and the elaboration on the related theoretical framework in Chapter Three, two-phase mixed-method research was designed and developed, covering both quantitative and qualitative methods. The quantitative methods, namely pre- and post-tests of linguistic and content performance, were conducted to scrutinise the statistical relationship between CLIL exposure and learners' attainment in language and content. The key findings were verified through the qualitative methods of interviews in the second phase. The triangulated data provided more reliable and dependable evidence of CLIL effectiveness in a Chinese educational context. A data set of this study with the mixed-method approach led itself to a wide range of analyses. On the one hand, the researcher opted to proceed from the cross-sectional towards longitudinal to be able to characterise changes, and on the other hand, to proceed from quantitative to qualitative data to obtain a deep and detailed voice in the discussion. The aim of the above procedure of data analysis was to investigate temporal changes in learners' academic performance over one semester of CLIL learning.

This chapter provided a more extensive discussion based on findings from Chapter

Five, highlighting potential links between the literature, theoretical framework, the methodology adopted, and interpretations of findings. In addition, the chapter aims to depict the knowledge contributions of the study and discuss further the implications and recommendations accordingly.

6.1 Discussions

6.1.1 Improvements in writing skills

As illustrated in Sections 2.2.1 and 2.2.2, general statements of CLIL effects on learners' language learning outcomes are unsurprisingly positive in the CLIL research. CLIL learners are observed to obtain considerably higher proficiency in the target language than their counterparts taught in the mainstream EFL class. In a similar vein, positive effects on communicative competence are noticeable. In the present study, similar findings were spotted as well. Both quantitative numbers and qualitative words from this study revealed a significant improvement in language outcomes. Such improvement is attributed to factors in CLIL settings, including the teacher's help to nurture language growth through content learning; the usage of authentic materials through which learners can use English for both specific and general purposes; the frequent practice of English writing to facilitate learners' linguistic and communicative competence; diversifying of methods and approaches to classroom practice to increase learners' motivation and expectation in the process of linguistic acquisition.

Besides, results on writing skills in this study provided support for positive CLIL effects. As illustrated in Section 5.2.1, CLIL learners performed much better than their counterparts in the control group. This difference might result from the fact that the CLIL class provided more flexibility than the L1 class for approaching content from various perspectives, and more space for interactions. In other words, the CLIL participants in this study learned the content differently from their peer controls. They approached the knowledge, discussed it from diverse viewpoints and related it to their own experiences. To fulfil this, learners had to use the CLIL language in diverse ways to realize these functions. Such CLIL learning experiences might facilitate learners'

writing skills in the target language when they resort to various resources and strategies to study both content and language. Thus, some implications for incumbent and potential CLIL teachers are provided. Since CLIL learners have to approach the language of academic discourse in a content subject, CLIL teachers have to be equipped with related linguistic training on “the linguistic features required for the representation of content in their subject, in a variety of genres and both in the written and spoken registers” (Ana Llinares & Whittaker, 2010, p. 126). This focus on language makes use of learners’ communicative needs and bring in efforts on the lexis and grammar needed to convey their meanings. Moreover, CLIL teachers have to skilfully balance and fuse the priority of learning content and linguistic features, integrating the 4Cs, namely content and communication with cognition and culture to reach a “symbiotic relationship” as depicted in Section 3.1.

6.1.2 The role of language learning in CLIL contexts

Since both CLIL participants and the teacher showed satisfaction with their linguistic outcomes, one issue worth discussing in this section is to further understand the planned pedagogic fusion of the contextualised 4Cs into the teaching and learning practice of the CLIL language, namely the role of language learning in CLIL contexts. In this study, Coyle et al.’s (2010, p. 36) “trptych of language in CLIL contexts” was adopted to analyse the role of CLIL language using from three correlated aspects, covering “language of learning, language for learning and language through learning”.

Language of learning indicates “an analysis of language needed for learners to access basic concepts and skills relating to the content theme or topic” (Coyle et al., 2010, p. 36). This means that the linguistic focus in a Chinese college CLIL class might shift from grammatical levels of knowledge to functional levels of knowledge and skills required by the subject content. For example, CLIL participants in this study expressed their perception of ‘English writing proficiency’ (as in Section 5.2.1) in the interview responses. They regarded it as ‘a way of thinking’ or ‘a solution to solve the problem’,

demonstrating their recognition of the knowledge and skills in both spoken and written language required by the content subject. In this sense, CLIL language learning comprises not only grammar learning and text reading as emphasised within the conventional mainstream EFL contexts, but also authenticity and relevance that is essential to effective content learning. That is to say, it is not necessary for language learning in CLIL contexts to go along the same progression as its conventional EFL counterparts. Learners are encouraged to learn the language appropriate to the content in a more effective way. Certainly, this does not imply that grammar learning should be reduced to support content knowledge, but a more varied and richer approach can be developed to accommodate the interrelationship between content objectives and language objectives.

As argued in Section 3.1.2, CLIL learners need to progress cognitively and systematically in language learning and using. Using language to learn is no less significant than learning to use the language. Savignon (2007) suggests that language using needs to be included in the goal of linguistic acquisition in CLIL, instead of exclusively on grammatical issues. In CLIL contexts, learners are given more opportunities to use the vehicular language in authentic interactions to foster communicative skills and cognitive engagement. In this sense, language for learning, or using language to learn, facilitates learners to practise the kind of language needed to operate in a foreign language setting. As Elaine, a CLIL participant, stated in her interview (in Section 5.2.1), in the CLIL class, learners were required to develop skills needed for classroom activities, including debating, discussing, enquiring, thinking, memorizing, asking questions, pair work and cooperative teamwork. By this means, CLIL learners become capable of understanding and using the language to learn, support, and develop a repertoire of strategies related to content learning. These strategies facilitate CLIL participants to enquire, debate, negotiate, engage in teamwork and use the vehicular language independently, which are important for CLIL tasks and activities to be fulfilled successfully and effectively.

In CLIL contexts, it is asserted that a higher degree of learning occurs when learners are prompted to articulate their understanding of content and language learning. Different from conventional language or content classes, CLIL learners are required to develop a level of talk, interaction and dialogic activities that need the active cognitive engagement in the class. In this way, language through learning in CLIL settings is built on the tenet that successful learning can be achieved via active engagement in communicative activities. As Charles and Elaine's interview responses showed (in Section 5.2.1), they involved more efforts and time into language learning and using, and were pleased to engage in the process of language acquisition cognitively. The responses reveal that language is indispensable for supporting and advancing learners' cognitive processes in order to access, apply and acquire new knowledge. Thus, language through learning happens, leading to progression in knowledge and skills, understanding of content and language, and improvement of cognitive thinking.

From the above perspectives, it is noted that CLIL allows efficient integration of learning to use language and using language to learn. However, identifying the part of language to be learnt is challenging and needs careful and organized evaluation ahead of CLIL practice. This evaluation goes beyond grammatical knowledge of basic words or key terms. It also includes an understanding of the linguistic genre and other linguistic functions attached to diversified content themes. This means that the CLIL language needs to be appropriately learned in accordance with reconstruction of content, progression of cognitive processes, learning through the vehicular language, and other parameters in the learning settings.

6.1.3 Translanguaging

In CLIL classes in the present study, it was highly unlikely that participants' existing linguistic level could be as high as their cognitive level, eliciting mismatches between the two. These mismatches, especially on those occasions when the linguistic

requirement is too demanding, might hamper effective learning in the process of content knowledge acquisition. In other words, if learners' linguistic proficiency cannot meet the threshold demand of CLIL learning, linguistic barriers might occur, complicating content learning in the CLIL setting. Thus, to guarantee that learners are cognitively challenged yet linguistically supported in content learning, it is important to provide the language needed for CLIL participants to fulfil class activities and knowledge acquisition. As Otten (1993, p. 73) suggests, "content specific methodology would have to focus on the learner, making language and content learning explicit and transparent, defining subject specific skills and thus enabling the learners to bridge the gap between the learners' conceptual and cognitive capacities and the learners' linguistic level".

With the purpose to bridge the gap between learners' linguistic and cognitive levels, a concern is raised, relating to the code-switching between languages. Code-switching refers to a systematic shift from one language to another for specific reasons (Papaja & Wysocka-Narewska, 2020; L. Wei & Martin, 2009). In the CLIL context, classroom instruction is frequently carried out in the blend of the vehicular and native languages, giving rise to code-switching between the two (García, 2009; San Isidro & Lasagabaster, 2019). This type of code-switching, termed translanguaging, results in a flexible shift in the language of instruction in CLIL classes. For instance, one language might be adopted to explain key concepts, while the other to realize specific class functions or activities. Coyle et al. (2010) further clarify that by systematic switching between languages according to pre-set aims of content, language and cognition, CLIL learners may resort to the first language to fulfil various tasks. They may turn to a textbook in the native language in preview and review to develop individual learning strategies. They may ask for explanations from the teacher to build confidence and check comprehension. In addition, they may speak to the teacher in L1 in problem-solving to achieve higher levels of content mastery.

In the present study, translanguaging might have facilitated overcoming CLIL

participants' worries that learning through a non-native language might lead to their misunderstanding and incomprehension of some key terms in content knowledge. Most interviewees in the study welcomed teachers shifting from English to Chinese in articulating complex ideas. Some respondents further anticipated that Chinese could be adopted as the medium of instruction on some occasions, especially in introducing new terminologies and concepts at the very beginning of the course. These responses unveiled participants' preference of using the first language to support content learning in the CLIL settings. This result finds supportive evidence in W. Yang and Gosling's (2014) research. It was revealed that CLIL participants felt more confident in content learning when teachers attempted to accommodate learners' lower level of L2 proficiency via native language in classes. Despite participants' strong inclination for translanguaging that facilitated their CLIL learning, the CLIL teacher raised uncertainty of language choices in classroom settings. She was new to thinking explicitly about translanguaging. Her knowledge and perception of 'bilingual' practices may have been largely influenced by her experience in monolingual instructions in content teaching, either in L1 or L2. Thus, 'translanguaging' requires her to walk out of 'comfort zone' and explore an uncharted territory. In other words, different from mainstream EFL teachers who have a greater tendency to monitor and steer students away from L1 use, CLIL teachers could regard language alteration as a commonly occurring phenomenon, serving both as a metalanguage for language play or as a communicative strategy to explain problems and reach mutual understanding. Certainly, multilingual practices are dynamic, functional and highly context-dependent, which might bring some practical implications for future CLIL studies and practices.

6.1.4 Cognitive thinking processes in CLIL learning

As discussed in the previous chapter, CLIL is a cognitively demanding process because extra efforts are needed for learners to comprehend content knowledge taught in the vehicle language. Meanwhile, learners have to construct language and content knowledge in the same language. In the present study, the findings offer proof that CLIL

participants were intellectually challenged to raise both content and linguistic achievement levels, such as transforming information and ideas, solving problems, gaining understanding and discovering new meanings. In other words, CLIL learning might lead to learners' cognitive improvement in various aspects. Indications are that a different way of looking at knowledge and culture seemed to converge in CLIL participants' cognitive development. They learned "to know how to think, reason, make informed choices, respond creatively to challenges and opportunities, and be skilled in problem-solving and higher-order, creative thinking to construct a framework to interpret meaning and understanding" (Pérez-Cañado, 2021, p. 7).

When further investigating CLIL learners' cognitive engagement in content and linguistic knowledge learning, two issues related to cognitive thinking skills are worthy of further discussion. Firstly, it is noted that for the majority of CLIL participants in the present study, these cognitive thinking skills were not adopted separately but were used integrally to effective CLIL learning. Based on the adoption of lower-order thinking skills (such as remembering, understanding and applying), some participants developed and integrated higher-order thinking skills (such as creating, evaluating and analysing) by making judgments according to criteria and standards through checking and critiquing in the process of CLIL knowledge acquisition. Thus, these CLIL learners were cognitively challenged to develop their cognitive thinking skills in an interrelated manner, which facilitated their language and content knowledge growth.

Secondly, it is worth noticing that some participants tended to progress from the cognitive process of lower-order thinking skills (remembering, understanding and applying) to higher-order thinking skills (analysing, evaluating and creating) in their CLIL learning process. The aforementioned shift of thinking skills from lower-order to higher-order implied that CLIL participants in this study made gains in higher-order thinking skills, like problem solving, interpreting meaning and making comprehension. This finding resonated with Coyle et al.'s (2010, p. 5) evaluation of the CLIL approach

that “learners will develop higher-order thinking skills, demonstrating their ability to make observations, analyse, generalise and applying their skills to fresh contexts”. Certainly, this shift does not mean the either-or position of lower-order and higher-order thinking skills, but acknowledge a dynamic back-and-forth movement between these two kinds of skills in CLIL learning, as both lower-order thinking and higher-order thinking are integral to effective learning in CLIL settings.

In sum, for CLIL to be effective, it has to challenge learners to create knowledge and develop new skills in both language and content aspects through reflection and engagement in higher-order and low-order thinking. CLIL does not mean the conveyance of knowledge from an expert to a novice. It empowers learners to construct their own understanding and knowledge. In this way, a careful plan is needed in CLIL to consider the relationships between cognitive processing (learning) and knowledge acquisition (language and content).

6.1.5 Changes in learners’ intercultural understanding

As discussed in Section 3.1.4, the cultural aspect in the CLIL framework is expected to extend its potential to develop learners’ intercultural understanding in relation to their intercultural communication and cognitive thinking skills. In other words, CLIL offers learners with intercultural experiences that its monolingual counterpart lacks and facilitate learners to achieve a deeper understanding of global citizenship.

In the present study, it is worthy of looking through participants’ responses that revealed a significant gap between their willingness to increase intercultural awareness in the CLIL approach and the effectiveness of real-world practice. On the one hand, it was found that a majority of participants became aware of diversified cultures in their CLIL learning processes. They made efforts to use cognitive skills to moderate between native and foreign cultures, aiming to promote intercultural understanding. This process

might start with fostering learners' awareness of native cultures, followed by developing their cultural attitudes, knowledge and skills in interactive settings. Subsequently, learners' language and content knowledge and skills are improved to apply and analyse social processes and outcomes critically (Villabona & Cenoz, 2021). Thus, from a holistic perspective, by integrating diversified cultures into a whole teaching system, CLIL enhances learner's intercultural understanding by developing "an ability to see and manage the relationship between themselves and their own cultural beliefs, behaviours and meanings, as expressed in a foreign language, and those of their interlocutors, expressed in the same language – or even a combination of language" (Byram, 1997, p. 12).

On the other hand, the impact of the dual foci nature of the CLIL approach is still questionable. Though in the CLIL context as the present study, this approach seems to provide an ideal environment to foster learners' intercultural understanding in real-world classroom practice. As illustrated in Section 5.2.1, some participants expressed their negative feedback towards cultural learning in the CLIL setting and doubted the effectiveness of CLIL to foster a deepening intercultural understanding. Some noteworthy reasons were revealed in these participants' responses, as follows.

The first reason might lie in learners' habitual reliance on conventional mainstream educational approach. In CLIL settings, intercultural understanding is built on classroom interaction with peers and teachers in and through the vehicular language and extends social interactions beyond the classroom. In fact, intercultural understanding is constructed and developed via interactions with various individuals in diversified circumstances. New situations encourage learners to adapt purposefully to develop their intercultural awareness and skills. However, due to the long imbue ment in highly test-driven and teacher-centred conventions in the Chinese educational system, a considerable percentage of participants in this study tended to rely exceedingly on teachers' instruction and textbook content by rote memorisation. These conventions

crippled them from participating in classroom interactions, especially those on cultural comparisons. In such an educational context, whole-class instruction still prevailed in most classroom activities. However, the innovation to increase intercultural understanding appeared to be based more on teachers' hope than learners' practice.

The second cause might be the lack of appropriate authentic materials for CLIL education. Coyle et al. (2010) assert that culture associated with content and language knowledge may not be 'learned' in some lessons on festival rituals, traditional costumes, or local delicacies, but through the use of appropriate authentic materials contributing to a deepening awareness of cultural comparisons, which in turn affects the discovering and understanding of 'otherness'. However, in the classroom practice in this study and in China, ready-made CLIL teaching and learning materials are scarce, especially those on cultural patterns, customs and lifestyle of 'otherness'. This contrasts with the vast number of English language teaching coursebooks and resources in the market. Lack of appropriate ready-made material might hinder Chinese CLIL learners' motivation to work alongside other learners from different cultures, which is a fundamental way to understand 'otherness'. For implication, further considerations of CLIL materials are needed to ensure meaningful connections between classroom practice and adequate learning resources, rather than tokenistic reference.

Another potential reason might be the learners' prime attention to formal, summative assessment in this study. In line with China's educational convention, summative assessment serves as a decisive force throughout the curriculum and compels CLIL learners to over-rely on existing textbooks, which is often considered as an ineffective means to foster and develop intercultural awareness and understanding.

From a theoretical perspective, CLIL classes provide learners with a variety of opportunities for intercultural interaction. The adoption of authentic teaching materials and intercultural linking allows learners to distinguish features of diverse cultures.

However, in real-world practice, its cultural impact still largely hinges on learners' engagement in interactive and dialogic learning both inside and outside the class, together with the effective fulfilment of contextual parameters. The CLIL class "needs to be thought through to ensure meaningful connections rather than tokenistic reference" (Coyle et al., 2010, p. 95). In CLIL, the emphasis of cultural components could include extending the content, setting the content learning in different cultural contexts, investigating different learning patterns in Chinese and western educational contexts, comparing people's discrepant attitudes and reactions towards the same content topic, or exploring the potential approach in face of global mobility and employability. In sum, full considerations of culture-related classroom practice are of great significance, calling for further research and implication.

6.2 Implications

As mentioned at the beginning, the initial cause of the present study lies in the ongoing attempts to make sense of some of the significant changes that the researcher frequently observe in those CLIL learners who venture to study within the dual-focused educational approach. One aim here is to unpack this observation of changes in a way that might be helpful to integrate the CLIL framework more effectively into classroom practice. Consequently, the present study would be ended by considering how its findings may apply to the classroom. Based on the theoretical framework as well as real-world attempts, Coyle et al. (2010) have provided some practical recommendations for those who design, manage and participate in these related practices, including integrating the 4Cs into a holistic approach, giving orientation courses for bridging cultural differences, and offering relevant training to increase linguistic proficiency. Based on Coyle et al.'s (2010) recommendations, together with the findings from the present study, some implications are provided.

CLIL contributes to learners' career development and employability in a multilingual world through the integration of content, language, cognition and culture

acquisition in some way. Thus, at the preparation stage of a CLIL course, it is essential to consider the interrelations among the four elements. For instance, the design of classroom activities might be shaped by the awareness of integrating cognitive elements and content. In a similar vein, tasks concerning the interconnection of cognitive and communicative elements are needed to guarantee learners with the content language and communication language to fulfil classroom procedures. However, though CLIL teachers are generally required to take a holistic and inclusive approach to the integration of different elements in CLIL, the 4Cs may progress at different rates as illustrated in the teacher's feedback in Section 5.2.2, hinging on various external and internal variables in the CLIL contexts. Thus, some arrangements could be taken into account at this preparation stage. For example, linguistic support could be given to bridge the gap between content learning and linguistic knowledge, such as integrating the grammar points through different uses across CLIL lessons, analysing keywords and phrases before content learning, involving language use and language practice in the spiral of language progress. Besides, orientation courses could be provided to bridge cultural differences. In these courses, CLIL teachers could introduce a broader cultural context and encourage learners to participate in comparative studies by resorting to multimedia means favoured by learners, including short videos and the Internet.

In real-world practice, one important implication needs to be fully considered is 'translanguaging', as evidently shown in Sections 5.2.1 and 6.1.3. Based on planned content, language and cognition development, bilingual blended instruction with code-switching between the vehicular language and the native language is recommended for specific types of classroom activities. For instance, one language might be used for summarising and analysing key concepts, while the other for carrying out class procedures. Also, learners are encouraged to use two languages systematically and dynamically. For example, they might turn to materials in the native language in completing assignment to double-check comprehension. In the classroom, learners might turn to teachers in a particular language for further clarification and explanation.

When solving the problem, they might use their L1 to speak to teachers.

To ensure the effectiveness of CLIL courses, further analysis and understanding of CLIL processes and results are needed after real-world classroom practice. Thus, a set of evaluation measures is recommended to understand how the CLIL programmes have realized its aims, including what parameters have led to achievement, what roles teachers need to play, what kind of evidence could be achieved to examine the task styles and outcomes, and so on. These measures could be informal assessments, questionnaires, interviews, portfolios of work, etc.

In sum, CLIL might not only make a significant contribution to offering learners favourable and fruitful experiences of knowledge construction and sharing, but also encourage learners to become global citizens with tolerance, curiosity and responsibility. However, CLIL is not isolated from the contexts where it is administered. Full considerations of integrating cognitive, communication, cognition and culture into specific contexts are significant and essential.

6.3 Limitations

The present study, being preliminary and experimental, inevitably has limitations. First of all, the sample selection for this study was restricted. The study only targeted second-year students in a public university in Eastern China. The results might not be statistically generalizable to learners in other grades, other Chinese universities, or other universities in different countries. Since the CLIL setting is an essential contextual factor in the process of language and content knowledge acquisition and is usually unique, further studies are needed to identify contextual factors to gain deeper insights into the integrated approach.

Moreover, this study was conducted without a pilot study, which could have provided useful reference in advance. As noted by some researchers (Thabane et al.,

2010), a pilot study might reveal flaws in main research design, protocols not followed, or unsuitable research instruments. However, the present study didn't take some forms of a formal pilot study for various reasons, such as the complications in the research design, limitations in sample selection and the gatekeeper's refusal. Thus, the researcher invested a great deal of time and effort in considering and preparing for the present study before and during the process, for example, keeping frequent contact with the target learners and the teacher, reviewing the related theoretical frameworks, designing the research carefully and cautiously.

However, the researcher believes that the present study is one of the tentative steps taken into a complex and dynamic perspective of the CLIL approach. As Irie and Ryan (2015, p. 360) assert that “a discipline is at its most productive when theory and practice are in step, when theory both informs and is informed by practice”. It is expected that CLIL research could be well developed when the theory and practice manage to ‘talk’ to each other.

REFERENCE:

Adamson, J., & Coulson, D. (2014). Pathways towards success for novice academic writers in a CLIL setting: A study in an Asian EFL context. In R. Al-Mahrooqi, V. S. Thakur, & A. Roscoe (Eds.),

- Methodologies for effective writing instruction in EFL and ESL classrooms* (pp. 148-168). Hershey, PA, USA: IGI Global.
- Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation*, 12(1), 75-93.
- Aguilar, M. (2017). Engineering lecturers' views on CLIL and EMI. *International Journal of Bilingual Education and Bilingualism*, 20(6), 722-735.
- Aguilar, M., & Muñoz, C. (2014). The effect of proficiency on CLIL benefits in Engineering students in Spain. *International Journal of Applied Linguistics*, 24(1), 1-18.
- Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism*, 15(2), 183-197.
- Airey, J. (2009). Estimating undergraduate bilingual scientific literacy in Sweden. *International CLIL Research Journal*, 1(2), 26-35.
- Airey, J., & Linder, C. (2006). Language and the experience of learning university physics in Sweden. *European Journal of Physics*, 27(3), 553-560.
- Alonso, A. C. (2017). The receptive vocabulary size of Spanish 5th grade Primary School students in CLIL and Non-CLIL instruction. *ES Review Spanish Journal of English Studies*(36), 63-86.
- Anderson, C. E. (2011). CLIL for CALP in the multilingual, pluricultural, globalized knowledge society: Experiences and backgrounds to L2 English usage among Latin American L1 Spanish-users. *Latin American Journal of Content & Language Integrated Learning*, 4(2), 51-66.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, USA: Longman.
- Apfel, C. (2012). Coping with CLIL: Dropouts from CLIL streams in Germany. *International CLIL Research Journal*, 1(4), 47-56.
- Badertscher, H., & Bieri, T. (2009). *Wissenserwerb im content and language integrated learning: Empirische Befunde und Interpretationen*. Bern-Stuttgart-Wien, Switzerland: Haupt.
- Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Bristol, UK: Multilingual Matters.
- Banegas, D. L. (2014). An investigation into CLIL-related sections of EFL coursebooks: Issues of CLIL inclusion in the publishing market. *International Journal of Bilingual Education and Bilingualism*, 17(3), 345-359.
- Barroso, J. M. (2007). President of the European Commission [Press release]. Retrieved from <http://europa.eu/rapid.pressReleasesAction.do?reference=IP/07/646>
- Barwell, R. (2016). A Bakhtinian Perspective on Language and Content Integration: Encountering the Alien Word in Second Language Mathematics Classrooms. In *Conceptualising Integration in CLIL and Multilingual Education* (pp. 101-120): Multilingual Matters.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives. Vol. 1: Cognitive Domain*. New York: McKay.
- Bloomfield, J., & Fisher, M. J. (2019). Quantitative research design. *Journal of the Australasian Rehabilitation Nurses Association*, 22(2), 27-30.
- Bonnet, A. (2012). Towards an evidence base for CLIL: How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal*, 4(1), 66-78.
- Broca, Á. (2016). CLIL and non-CLIL: Differences from the outset. *ELT Journal*, 70(3), 320-331.
- Brown, H. D. (1980). *Principles of language learning and teaching*. Eaglewood Cliffs, NJ; USA: Prentice

Hall.

- Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA, USA: Harvard University Press.
- Bruner, J. S. (2009). *The Process of Education: A Landmark in Educational Theory*. Cambridge, MA, USA: Harvard University Press.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39(4), 523-532. Retrieved from <Go to ISI>://WOS:000298823200011
- Bruton, A. (2013). CLIL: Some of the reasons why ... and why not. *System*, 41(3), 587-597. Retrieved from <Go to ISI>://WOS:000326487200007
- Bruton, A. (2015). CLIL: Detail matters in the whole picture. More than a reply to J. Hüttner and U. Smit (2014). *System*, 53, 119-128.
- Bryman, M. A. H. A. (2004). *Handbook of data analysis*.
- Burns, N., & Grove, S. K. (2005). *The Practice of Nursing Research: Conduct, Critique, and Utilization*. US: Elsevier Saunders.
- Burston, J., & Kyprianou, M. (2009). *Integrating Task-based and Content and Language Integrated Learning*. Paper presented at the TBLT 2009: The 3rd Biennial Conference on Task-based Language Teaching, Lancaster University, UK.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon, UK: Multilingual Matters.
- Byram, M., & Wagner, M. (2018). Making a difference: Language teaching for intercultural and international dialogue. *Foreign Language Annals*, 51(1), 140-151.
- Cai, J. G. (2010). 关于我国大学英语教学重新定位的思考 (Thoughts on repositioning Chinese college English education). *外语教学与研究 (Foreign Language Teaching and Research)*, 42(4), 306-308.
- Cai, J. G. (2013). 大学英语生存危机及其学科地位研究(Research on the existence and disciplinary status of college English). *中国大学教学(College Education in China)*(2), 10-14.
- Campbell, D. T., & Stanley, J. C. (2015). *Experimental and Quasi-Experimental Designs for Research*: Ravenio Books.
- Cenoz, J. (2013). Discussion: Towards an educational perspective in CLIL language policy and pedagogical practice. *International Journal of Bilingual Education and Bilingualism*, 16(3), 389-394. doi:10.1080/13670050.2013.777392
- Cenoz, J. (2015). Content-based instruction and content and language integrated learning: the same or different? *Language, Culture and Curriculum*, 28(1), 8-24.
- Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics*, 35(3), 243-262. Retrieved from <Go to ISI>://WOS:000338108100001
- Chang, J. (2006). Globalization and English in Chinese higher education. *World Englishes*, 25(3 - 4), 513-525.
- Chaplier, C., & O'Connell, A.-M. (2015). Prolegomena to the epistemology of languages for non-specialists: the example of CLIL. *Journal of Language and Cultural Education*, 3(1), 69-77.
- Chen, C. (2013). A survey study on the context validity of the writing test of CET-4. *Education and Teaching Research*, 27(3), 84-89.
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). New York, USA: Routledge.
- Coonan, C. M. (2007). Insider views of the CLIL class through teacher self-observation-introspection. *International Journal of Bilingual Education and Bilingualism*, 10(5), 625-646.

- Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism*, 10(5), 543-562.
- Coyle, D. (2008). CLIL-A pedagogical approach from the European perspective. In N. H. Hornberger (Ed.), *Encyclopedia of Language and Education* (Vol. Volume 4: Second and Foreign Language Education, pp. 97-111). Colfax, NM, USA: Springer.
- Coyle, D. (2013). Listening to learners: An investigation into 'successful learning' across CLIL contexts. *International Journal of Bilingual Education and Bilingualism*, 16(3), 244-266. Retrieved from <Go to ISI>://WOS:000319322100002
- Coyle, D. (2018). The place of CLIL in (bilingual) education. *Theory into Practice*, 57(3), 166-176.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Coyle, D., & Meyer, O. (2021). *Beyond CLIL: Pluriliteracies Teaching for Deeper Learning*. Cambridge, England: Cambridge University Press.
- Crandall, J. (1994). Strategic integration: Preparing Inaguage and content teachers for linguistically and culturally diverse classrooms. In J. E. Alatis (Ed.), *Strategic integration: Preparing Inaguage and content teachers for linguistically and culturally diverse classrooms*. Georgetown University: Georgetown University Press.
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. London, UK: SAGE.
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative* (4th ed.). London, UK: Pearson.
- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, California, USA: Sage.
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Thousand Oaks, California, USA: Sage.
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice*, 39(3), 124-131.
- Creswell, J. W., & Plano Clark, V. L. (2011). Choosing a mixed methods design. In J. W. Creswell & Plano Clark (Eds.), *Designing and Conducting Mixed Methods Research* (pp. 53-106). New York: Sage.
- Cummins, J. (2005). Using information technology to create a zone of proximal development for academic language learning: A critical perspective on trends and possibilities. *Information Technology and Innovation in Language Education*, 105-126.
- Curtis, W., Murphy, M., & Shields, S. (2013). *Research and Education*. London: Routledge.
- Dafouz, E., Núñez, B., Sancho, C., & Foran, D. (2007). Integrating CLIL at the tertiary level: teachers' and students' reactions. *Diverse contexts converging goals. Content and language integrated learning in Europe*, 4, 91-102.
- Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences—Killing two birds with one stone? *Learning and Instruction*, 41(1), 23-31.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) Classrooms* (Vol. 20): John Benjamins Publishing.
- Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future perspectives for English language teaching*. Heidelberg, Germany: Carl Winter.

- Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From practice to principles? *Annual Review of Applied Linguistics*, 31, 182-204. Retrieved from <Go to ISI>://WOS:000295464300011
- Dalton-Puffer, C., Hüttner, J., & Smit, U. (2021). 6 From Voluntary to Obligatory CLIL in Upper Secondary Technical Colleges: Teacher and Student Voices from a Diverse Landscape. In *The Psychological Experience of Integrating Content and Language* (pp. 93-111): Multilingual Matters.
- Dalton-Puffer, C., Llinares, A., Lorenzo, F., & Nikula, T. (2014). "You Can Stand Under My Umbrella": Immersion, CLIL and Bilingual Education. A Response to Cenoz, Genesee & Gorter (2013). *Applied Linguistics*, 35(2), 213-218.
- Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). Language use and language learning in CLIL: Current findings and contentious issues. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in Clil Classrooms* (pp. 279-291). Amsterdam, Holland: John Benjamins.
- Dalton-Puffer, C., & Smit, U. (2013). Content and Language Integrated Learning: A research agenda. *Language Teaching*, 46(4), 545-559. Retrieved from <Go to ISI>://WOS:000324891100006
- Deller, S., & Price, C. (2005). Teaching other subjects in English (CLIL). *In English*, 29-31.
- Devos, N. J. (2016). *Peer interactions in new content and language integrated settings* (Vol. 24). London, UK: Springer.
- Dickey, R. J. (2004). Content (adj) or content (n) with your English classes. *Education International*, 1(3), 10-15.
- Eurobarometer. (2012). *Europeans and their languages: Eurobarometer 386 report*. Brussels, Belgium
- European Commission. (1995). *White paper on education and learning: Teaching and learning towards the learning society*.: European Commission Retrieved from http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf
- European Commission. (2007). *European Agenda for Culture in a Globalizing world*. European Commission
- European Commission Communication. (2003). *Promoting language learning and linguistic diversity: An action plan 2004-2006*. Retrieved from http://ec.europa.eu/education/doc/official/keydoc/actlang/act_lang_en.pdf
- European Council. (2003). *Guide for the Development of Language Education Policies in Europe*. European Council
- European Council. (2011). *Common European Framework of Reference for Languages*. European Council,
- Eurydice. (2006). *Content and Language Integrated Learning (CLIL) at School in Europe*. Brussels, Belgium: European Commission
- Fazzi, F., & Lasagabaster, D. (2021). Learning beyond the classroom: students' attitudes towards the integration of CLIL and museum-based pedagogies. *Innovation in Language Learning and Teaching*, 15(2), 156-168.
- Feng, A. W. (2009). English in China: Convergence and divergence in policy and practice. *AILA Review*, 22(1), 85-102.
- Frankfort-Nachmias, C., & Nachmias, D. (1992). *Research methods in the social sciences*. London, UK: Edward Arnold.
- Gallardo del Puerto, F., Gomez Lacabex, E., & García Lecumberri, M. L. (2009). Testing the effectiveness of content and language integrated learning in foreign language contexts: The assessment of English pronunciation. In Y. Ruiz de Zarobe & R. M. J. Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 63). Bristol, UK: Multilingual Matters.

- García, M. (2009). *Code-switching and Discourse Style in a Chicano Community*: Vervuert Verlagsgesellschaft.
- Giacobbi Jr, P. R., Poczwardowski, A., & Hager, P. F. (2005). A pragmatic research philosophy for applied sport psychology. *The Sport Psychologist*, 19(1), 18-31.
- Gierlinger, E. (2015). 'You can speak German, sir': on the complexity of teachers' L1 use in CLIL. *Language and Education*, 29(4), 347-368.
- Goris, J., Denessen, E., & Verhoeven, L. (2019). The contribution of CLIL to learners' international orientation and EFL confidence. *The Language Learning Journal*, 47(2), 246-256.
- Gray, D. E. (2009). *Doing research in the real world* (2nd ed.). Thousand Oaks, California, USA: Sage.
- Grin, F. (2005). *The value-added of CLIL: A language policy evaluation approach*. Paper presented at the Changing European Classroom: The Potential of Plurilingual Education, EU Presidency Conference, Luxembourg.
- Grinyer, A. (2009). The anonymity of research participants: Assumptions, ethics, and practicalities. *Pan*, 12(1), 49-58.
- Grix, J. (2018). *The Foundations of Research*. New York: Macmillan International Higher Education.
- Guo, Y. Q., & Wu, X. J. (2013). CLIL 模式在大学英语写作中的应用探究(On the application of CLIL in college English writing). *外国语文(Foreign Language and Literature)*, 29(2), 166-169.
- Hao, C. M. (2012). 当前大学英语教育反思--蔡基刚大学英语教育思想述略及启示(Reflection on current college English education--Introduction and implication of Cai Jigang's educational thoughts on college English education). *当代教育理论与实践(Theory and Practice of Contemporary Education)*, 4(7), 159-161.
- Heine, L. (2010). *Problem solving in a foreign language* (Vol. 41): Walter de Gruyter.
- Hellekjær, G. O. (2008). A case for improved reading instruction for academic English reading proficiency. *Acta Didactica Norge*, 2(1), 1-17.
- Hemmi, C., & Banegas, D. L. (2021). *International perspectives on CLIL*. Cham: Palgrave Macmillan.
- Herbert, R. J., & Herbert, I. S. (1995). *Qualitative Interviewing: The Art of Hearing Data*. California, USA: Sage.
- Hou, H.-I. (2013). An EAP curriculum design of a content and language integrated learning program for hospitality students in Taiwan. *Latin American Journal of Content & Language Integrated Learning*, 6(2), 72-95.
- Hu, G. W., & Lei, J. (2014). English-medium instruction in Chinese higher education: A case study. *Higher Education*, 67(5), 551-567. doi:10.1007/s10734-013-9661-5
- Hu, G. W., Li, L. N., & Lei, J. (2014). English-medium instruction at a Chinese University: rhetoric and reality. *Language Policy*, 13(1), 21-40. doi:10.1007/s10993-013-9298-3
- Hughes, N. (2013). Developing academic register in CLIL: UK-based Spanish L2 students' Latin America Political Economy writing in the UK. *Latin American Journal of Content & Language Integrated Learning*, 6(2), 42-71.
- Hughes, S. P., & Madrid, D. (2020). The effects of CLIL on content knowledge in monolingual contexts. *The Language Learning Journal*, 48(1), 48-59.
- Hüttner, J., & Rieder-Bünemann, A. (2010). A cross-sectional analysis of oral narratives by children with CLIL and non-CLIL instruction. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in Clil Classrooms* (pp. 61-80). Amsterdam: John Benjamins.
- Hüttner, J., & Smit, U. (2014). CLIL (Content and Language Integrated Learning): The bigger picture. A response to: A. Bruton. 2013. CLIL: Some of the reasons why ... and why not. *System* 41 (2013):

- 587-597. *System*, 44, 160-167.
- Infante, D., Benvenuto, G., & Lastrucci, E. (2009). The effects of CLIL from the perspective of experienced teachers. *CLIL Practice: Perspectives from the Field*, 156-163.
- Ioannou Georgiou, S. (2012). Reviewing the puzzle of CLIL. *ELT Journal*, 66(4), 495-504. Retrieved from <Go to ISI>://WOS:000309134400009
- Irie, K., & Ryan, S. (2015). Study abroad and the dynamics of change in learner L2 self-concept. In Z. Dornyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 520-552). Bristol, UK: Multilingual Matters.
- James, D. B. (2002). *Doing Second Language Research*. Oxford: Oxford University Press.
- Jäppinen, A. K. (2005). Thinking and content learning of mathematics and science as cognitional development in content and language integrated learning (CLIL): Teaching through a foreign language in Finland. *Language and Education*, 19(2), 147-168.
- Järvinen, H.-M. (2010). Language as a meaning making resource in learning and teaching content. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in Clil Classrooms* (Vol. 7, pp. 145). Amsterdam, Holland: John Benjamins.
- Jexenkicker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula, & S. Ute (Eds.), *Language Use and Language Learning in Clil Classrooms* (pp. 169-189). Amsterdam, Holland: John Benjamins.
- Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). Thousand Oaks, California, USA: Sage.
- Johnson, B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 297-319). Thousand Oaks, California, USA: Sage.
- Juan-Garau, M., & Salazar-Noguera, J. (2015). *Content-based Language Learning in Multilingual Educational Environments*. New York: Springer International Publishing.
- Karimi, P., Lotfi, A. R., & Biria, R. (2017). Efficacy of content and language integrated learning (CLIL) pedagogical techniques in improving aviation students' engagement in attaining the specified learning outcomes. *Journal of Teaching English for Specific and Academic Purposes*, 5(4), 781-791.
- Kimmel, A. (1988). *Ethics and values in applied social research*. Newbury Park, California, USA: Sage.
- Kirsch, J. (2016). *CLIL in Catalonia: learning through talk and interaction in secondary science classrooms*. (PhD). University of Nottingham, Nottingham, UK.
- Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London, UK: Longman.
- Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *TESOL quarterly*, 35(4), 537-560.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1(1), 31-42.
- Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). *CLIL in Spain: Implementation, results and teacher training*. Newcastle upon Tyne, UK: Cambridge Scholars.

- Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64(4), 367-375.
- Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity*, 43(2), 265-275.
- Lei, J., & Hu, G. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *International Review of Applied Linguistics in Language Teaching*, 52(2), 99-126.
- Leung, C. (2005). Language and content in bilingual education. *Linguistics and Education*, 16(2), 238-252.
- Li, W. (2008). Research perspectives on bilingualism and multilingualism. In W. Li & M. G. Moyer (Eds.), *The Blackwell guide to research methods in bilingualism and multilingualism* (pp. 3-17). Oxford, UK: Blackwell.
- Lin, A. M. Y. (2015). Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum*, 28(1), 74-89.
- Liu, B. (2010). CLIL 理念在专业课程双语教学中的应用(Implementing CLIL in the teaching of disciplinary bilingual courses). *湖南工业大学学报(社会科学版)(Journal of Hunan University of Technology: Social Sciences Edition)*, 15(2), 140-142.
- Liu, H. (2019). An action research into task-based CLIL applied to education majors: From Chinese students' perspective. *English Language Teaching*, 12(3), 94-107.
- Llach, M. D. P. A. (2009). The role of Spanish L1 in the vocabulary use of CLIL and non-CLIL EFL learners. In Y. Ruiz de Zarobe & R. M. J. Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 112-129). Bristol, UK: Multilingual Matters.
- Llinares, A. (2015). Integration in CLIL: a proposal to inform research and successful pedagogy. *Language, Culture and Curriculum*, 28(1), 58-73.
- Llinares, A., & McCabe, A. (2020). Systemic functional linguistics: The perfect match for content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*, 1-6.
- Llinares, A., & Whittaker, R. (2010). Writing and speaking in the history class. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in Clil Classrooms* (pp. 125-144). Amsterdam, Holland: John Benjamins.
- Lorenzo, F. (2007). The sociolinguistics of CLIL: Language planning and language change in 21st century Europe. *Revista Española de Lingüística Aplicada*, 1, 27-38.
- Lorenzo, F. (2016). Historical literacy in bilingual settings: Cognitive academic language in CLIL history narratives. *Linguistics and Education*, 37(1), 32-41.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, 31(3), 418-442.
- Luo, Q. L. (2006). CLIL 教学模式及其对中国公共外语教学的启示(CLIL and Its Implications for China's EFL Teaching). *北京航空航天大学学报(社会科学版)(Journal of Beijing University of Aeronautics and Astronautics: Social Sciences Edition)*, 19(3), 70-72.
- Lyster, R. (2007). *Learning and Teaching Languages through Content: A Counterbalanced Approach* (Vol. 18): John Benjamins Publishing.
- Lyster, R., & Ballinger, S. (2011). Content-based language teaching: Convergent concerns across divergent contexts. *Language Teaching Research*, 15(3), 279-288.

- Maljers, A., Marsh, D., & Wolff, D. (2007). *Windows on CLIL: Content and Language Integrated Learning in the Spotlight*. The Hague: European Platform for Dutch Education and Graz: European Centre for Modern Languages.
- Marsh, D. (2002). *CLIL/EMILE: the European Dimension: Actions, Trends and Foresight Potential*. Brussels, Belgium: European Commission
- Massler, U. (2012). Primary CLIL and its stakeholders: What children, parents and teachers think of the potential merits and pitfalls of CLIL modules in primary teaching. *International CLIL Research Journal*, 1(4), 36-46.
- Maurais, J. (2003). Towards a new linguistic order. In J. M. Maurais, M. (Ed.), *Languages in a Globalising World* (pp. 13-36). Cambridge: Cambridge Press.
- Maxwell, J. A., & Loomis, D. M. (2003). Mixed methods design: An alternative approach. *Handbook of Mixed Methods in Social and Behavioral Research*, 1(2003), 241-272.
- McDougald, J. S. (2016). CLIL approaches in education: Opportunities, challenges, or threats? *Latin American Journal of Content & Language Integrated Learning*, 9(2), 253-266.
- Mehisto, P. (2008). CLIL counterweights: Recognising and decreasing disjuncture in CLIL. *International CLIL Research Journal*, 1(1), 93-119.
- Mehisto, P., Frigols, M. J., & Marsh, D. (2008). *Uncovering CLIL: Content and Language Integrated Learning and Multilingual Education*. New York, USA: MacMillan ELT.
- Méndez García, M. d. C. (2013). The intercultural turn brought about by the implementation of CLIL programmes in Spanish monolingual areas: A case study of Andalusian primary and secondary schools. *The Language Learning Journal*, 41(3), 268-283.
- Méndez García, M. d. C. (2014). A case study on teachers' insights into their students' language and cognition development through the Andalusian CLIL programme. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras*, 22(1), 23-39.
- Met, M. (1998). Curriculum decision-making in content-based language teaching. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 35-63). Clevedon, UK: Multilingual Matters.
- Meyer, O. (2010). Towards quality CLIL: Successful planning and teaching strategies. *Puls*, 33, 11-29.
- Moate, J. M. (2011). The impact of foreign language mediated teaching on teachers' sense of professional integrity in the CLIL classroom. *European Journal of Teacher Education*, 34(3), 333-346.
- MOE. (2001). *关于加强高等学校本科教学工作提高教学质量的若干意见(The Guidelines for Improving the Quality of Undergraduate Teaching)*. (Higher Education, MOE (2001) No.4). Beijing, China: Chinese Ministry of Education Retrieved from http://www.law-lib.com/law/law_view.asp?id=16900
- MOE. (2015). *Daxue Yingyu Jiaoxue Zhinan [Guidelines on College English Teaching]*. Beijing, China: Chinese Ministry of Education
- Moghadam, N. Z., & Fatemipour, H. (2014). The effect of CLIL on vocabulary development by Iranian secondary school EFL learners. *Procedia-Social and Behavioral Sciences*, 98, 2004-2009.
- Morgan, D. L. (1988). *Focus Groups as Qualitative Research*. Beverly Hills, CA: Sage.
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48-76.
- Muñoz-Luna, R. (2014). From drills to CLIL: The paradigmatic and methodological evolution towards the

- integration of content and foreign language. *Profile Issues in Teachers' Professional Development*, 16(1), 167-180.
- Murphy, E., & Dingwall, R. (2001). The ethics of ethnography. In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland (Eds.), *Handbook of ethnography* (pp. 339-351). Thousand Oaks, California, USA: Sage.
- Murphy, J. P. (1990). *Pragmatism: From Peirce to Davidson*. Boulder, CO: Westview Pr.
- Navés, T., & Victori, M. (2010). CLIL in Catalonia: An overview of research studies. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 30-54). Newcastle upon Tyne, UK: Cambridge Scholars.
- Newman, I., & Benz, C. R. (1998). *Qualitative-quantitative Research Methodology: Exploring the Interactive Continuum*. Carbondale, IL: Southern Illinois University Press.
- Nikula, T. (2015). Hands-on tasks in CLIL science classrooms as sites for subject-specific language use and learning. *System*, 54, 14-27.
- Nikula, T., Dalton-Puffer, C., Llinares, A., & Lorenzo, F. (2016). More than content and language: The complexity of integration in CLIL and bilingual education. *Conceptualising Integration in CLIL and Multilingual Education*, 101(1).
- Ortega, L. (2015). Researching CLIL and TBLT interfaces. *System*, 54, 103-109.
- Otten, E. (1993). *Towards a content-specific methodology in bilingual education—Integrating language and content learning and teaching*. Paper presented at the Report of Workshop.
- Papadopoulos, I., & Griva, E. (2014). Learning in the traces of Greek Culture: a CLIL project for raising cultural awareness and developing L2 skills. *International Journal of Learning, Teaching and Educational Research*, 8(1), 76-92.
- Papaja, K. L. (2012). The impact of students' attitudes on CLIL. *Latin American Journal of Content & Language Integrated Learning*, 5(2), 28-56.
- Papaja, K. L., & Wysocka-Narewska, M. (2020). Investigating Code-switching in a Content and Language Integrated Learning (CLIL) Classroom. *Theory and Practice of Second Language Acquisition*, 6(1), 51-63.
- Paran, A. (2013). Content and language integrated learning: Panacea or policy borrowing myth? *Applied Linguistics Review*, 4(2), 317-342.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, California, USA: Sage.
- Pérez-Cañado, M. L. (2012). CLIL research in Europe: past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315-341.
- Pérez-Cañado, M. L. (2016). Are teachers ready for CLIL? Evidence from a European study. *European Journal of Teacher Education*, 39(2), 202-221.
- Pérez-Cañado, M. L. (2017). Stopping the “pendulum effect” in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review*, 8(1), 79-99.
- Pérez-Cañado, M. L. (2020). CLIL and elitism: Myth or reality? *The Language Learning Journal*, 48(1), 4-17.
- Pérez-Cañado, M. L. (2021). *Content and Language Integrated Learning in Monolingual Settings: New Insights from the Spanish Context* (Vol. 38): Springer Nature.
- Pérez-Vidal, C. (2013). Perspectives and lessons from the challenge of CLIL experiences. *Bilingual and Multilingual Education in the 21st Century. Building on Experience*, 59, 59-82.
- Pérez-Vidal, C., & Juan-Garau, M. (2010). To CLIL or not to CLIL? From bilingualism to multilingualism in

- Catalan/Spanish communities in Spain. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 115-138). Cambridge, UK: Cambridge Scholars Publishing.
- Pérez-Vidal, C., & Roquet, H. (2015). CLIL in context: Profiling language abilities. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 237-255). Cham, Switzerland: Springer.
- Pérez, A., Lorenzo, F., & Pavón, V. (2016). European bilingual models beyond lingua franca: Key findings from CLIL French programs. *Language Policy, 15*(4), 485-504.
- Piesche, N., Jonkmann, K., Fiege, C., & Keßler, J.-U. (2016). CLIL for all? A randomised controlled field experiment with sixth-grade students on the effects of content and language integrated science learning. *Learning and Instruction, 44*, 108-116.
- Pladevall-Ballester, E. (2015). Exploring primary school CLIL perceptions in Catalonia: students', teachers' and parents' opinions and expectations. *International Journal of Bilingual Education and Bilingualism, 18*(1), 45-59.
- Poarch, G. J. (2013). *Cross-linguistic Interaction, Lexical Representation, and Cognitive Control in Second Language Learners, Bilinguals, and Trilinguals*. London: Springer.
- Ragin, C. C. (1994). *Constructing Social Research*. Thousand Oaks, CA: Pine Forge Press.
- Richards, L. (2005). *Handling Qualitative Data*. London: Sage.
- Robson, C. (2002). *Real World Research (second edition)*. Oxford: Blackwell.
- Rorty, R. (1995). *Rorty & pragmatism: The philosopher responds to his critics*. Nashville, Tennessee, USA: Vanderbilt University Press.
- Roussel, S., & Tricot, A. (2017). Emile/CLIL, EMI vs TCC: des espaces de recherche si proches et si lointains. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication, 20*(3).
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque country. *International CLIL Research Journal, 1*(1), 60-73.
- Ruiz de Zarobe, Y. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula, & S. Ute (Eds.), *Language Use and Language Learning in Clil Classrooms* (pp. 191-212). Amsterdam, Holland: John Benjamins.
- Ruiz de Zarobe, Y. (2013). CLIL implementation: From policy-makers to individual initiatives. *International Journal of Bilingual Education and Bilingualism, 16*(3), 231-243.
- Ruiz de Zarobe, Y. (2015). The effects of implementing CLIL in education. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 51-68). New York, USA: Springer.
- Ruiz de Zarobe, Y., & Cenoz, J. (2015). Way forward in the twenty-first century in content-based instruction: moving towards integration. *Language, Culture and Curriculum, 28*(1), 90-96.
- Rumlich, D. (2014). Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In R. Breeze, C. Llamas Saíz, C. Martínez Pasamar, & C. Taberero Sala (Eds.), *Integration of theory and practice in CLIL* (pp. 75-95). Amsterdam, Holland: Rodop.
- San Isidro, X., & Lasagabaster, D. (2019). Code-switching in a CLIL multilingual setting: a longitudinal qualitative study. *International Journal of Multilingualism, 16*(3), 336-356.
- Sandberg, Y. (2015, November 23-26, 2015). *Subject-specific metaphors? Mixed-method analysis of CLIL teacher narratives*. Paper presented at the 4th International Conference on Language, Education and Diversity (LED 2015), Auckland, New Zealand.

- Savignon, S. J. (2007). Beyond communicative language teaching: What's ahead? *Journal of Pragmatics*, 39(1), 207-220.
- Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 189-213). Thousand Oaks: Sage.
- Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, 21(4), 328-341.
- Shek, D. T., & Sun, R. C. (2012). Promoting psychosocial competencies in university students: Evaluation based on a one-group pre-test/post-test design. *International Journal on Disability and Human Development*, 11(3), 229-234.
- Sheng, Y. L. (2012). 欧洲 CLIL 模式:外语教学法的新视角(European CLIL model: A new perspective on foreign language teaching methodology). *山东外语教学(Shandong Foreign Languages Teaching Journal)*(5), 65-69.
- Smit, U., & Dalton-Puffer, C. (2007). Current research on CLIL 2. *Special issue of VIEWS (Vienna English Working Papers)*, 17, 1-12.
- Somers, T., & Surmont, J. (2012). CLIL and immersion: how clear-cut are they? *ELT Journal*, 66(1), 113-116.
- Stake, R. (1995). *The Art of Case Study Research*. Thousand Oaks, CA: Sage.
- Sudhoff, J. (2010). CLIL and intercultural communicative competence: Foundations and approaches towards a fusion. *International CLIL Research Journal*, 1(3), 30-37.
- Surmont, J., Van de Craen, P. v. d., Struys, E., & Somers, T. (2014). Evaluating a CLIL student: Where to find the CLIL advantage. In R. Breeze, C. L. Saíz, C. M. Pasamar, & C. T. Sala (Eds.), *Integration of theory and practice in CLIL* (pp. 55-74). Amsterdam, Holland: Rodopi.
- Swain, M. (1995). Three functions of output in second language learning. *Principle and practice in applied linguistics: Studies in honour of HG Widdowson*, 2(3), 125-144.
- Sylvén, L. K. (2013). CLIL in Sweden—why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism*, 16(3), 301-320.
- Sylvén, L. K. (2017). Motivation, second language learning and CLIL. In A. Llinares & T. Morton (Eds.), *Applied linguistics perspectives on CLIL* (Vol. 47, pp. 51-66). Amsterdam, Holland: John Benjamins.
- Sylvén, L. K., & Ohlander, S. (2015). The CLISS project: Receptive vocabulary in CLIL versus non-CLIL groups. *Moderna språk*, 108(2), 80-114.
- Tedick, D. J., & Cammarata, L. (2012). Content and language integration in K–12 contexts: Student outcomes, teacher practices, and stakeholder perspectives. *Foreign Language Annals*, 45(S1), S28-S53.
- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L., . . . Goldsmith, C. (2010). A tutorial on pilot studies: The what, why and how. *BMC medical research methodology*, 10(1), 1-8.
- Ting, Y. T. (2011). CLIL... not only not immersion but also more than the sum of its parts. *ELT Journal*, 65(3), 314-317.
- Tong, F., & Shi, Q. (2012). Chinese–English bilingual education in China: A case study of college science majors. *International Journal of Bilingual Education and Bilingualism*, 15(2), 165-182.
- Ullmann, M. (1999). *History and geography through French: CLIL in a UK secondary school*. Retrieved from London: UK:
- Van de Craen, P., Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in primary schools: Teaching maths in a CLIL environment. In A. Maljers, D. Marsh, D. Wolff, F. Genesee,

- M. J. Frigols, & P. Mehisto (Eds.), *Diverse contexts—converging goals. CLIL in Europe* (pp. 185-200). Frankfurt am Main: Germany: Lang.
- Van de Craen, P., Mondt, K., Allain, L., & Gao, Y. (2007). Why and how CLIL works. An outline for a CLIL theory. *Vienna English Working Papers*, 16(3), 70-78.
- Villabona, N., & Cenoz, J. (2021). The integration of content and language in CLIL: A challenge for content-driven and language-driven teachers. *Language, Culture and Curriculum*. doi:10.1080/07908318.2021.1910703
- Vollmer, H. J., Heine, L., Troschke, R., Coetzee, D., & Küttel, V. (2006, 29 August). *Subject-specific competence and language use of CLIL learners: The case of geography in grade 10 of secondary schools in Germany*. Paper presented at the ESSE8 Conference, London, UK.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Cambridge, MA, USA: Harvard University Press.
- Wang, C. (2014). Communicative validity of the new CET-4 listening comprehension test in China. *Indonesian Journal of Applied Linguistics*, 4(1), 111-122.
- Wang, J., & Sun, C. (2012). On the causes of negative impact of CET-4. *Sino-US English Teaching*, 9(8), 1366-1372.
- Wang, X. W. (2011). CLIL 教学模式在大学外语教学中可行性探究-以西安外国语大学旅游学院旅游管理英语课程为例(Exploring feasibility of CLIL in college English teaching--taking Tourism Management English course in the Tourism College of Xi'an Foreign Language University as an example). *内蒙古师范大学学报(教育科学版)(Journal of Inner Mongolia Normal University: Educational Science)*, 24(7), 123-126.
- Wei, L., & Martin, P. (2009). Conflicts and tensions in classroom codeswitching: an introduction. *International Journal of Bilingual Education and Bilingualism*, 12(2), 117-122.
- Wei, R. N., & Feng, J. Y. (2015). Implementing CLIL for young learners in an EFL context beyond Europe. *English Today*, 31(1), 55-60.
- Wiesemes, R. (2009). Developing theories of practices in CLIL: CLIL as post-method pedagogies. In Y. Ruiz de Zarobe & R. M. J. Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 41-59). Bristol, UK: Multilingua Matters.
- Wolff, D. (2012). The European framework for CLIL teacher education. *Synergies Italie*, 8, 105-116.
- Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching*, 10(4), 116-126.
- Yan, C. M., & He, C. J. (2015). "We Are Left in Limbo!": Chinese EFL Student Teachers' Teaching Practicum Experience. *Frontiers of Education in China*, 10(2), 226-250.
- Yang, H. W., & Xu, C. Y. (2011). 浅析 CLIL 理念下大学英语教学与双语教学的衔接(Connecting college English education and bilingual education from CLIL perspective). *湖北经济学院学报(人文社会科学版) (Journal of Hubei University of Economics: Humanities and Social Sciences)*, 8(1), 179-180.
- Yang, W. (2014). Content and language integrated learning next in Asia: Evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme. *International Journal of Bilingual Education and Bilingualism*, 1-22.
- Yang, W. (2015). Tuning university undergraduates for high mobility and employability under the content and language integrated learning approach. *International Journal of Bilingual Education and Bilingualism (Online)*, 1-18. doi:10.1080/13670050.2015.1061474
- Yang, W., & Gosling, M. (2014). What makes a Taiwan CLIL programme highly recommended or not

recommended? *International Journal of Bilingual Education and Bilingualism*, 17(4), 394-409.
Retrieved from <Go to ISI>://WOS:000335846300002

- Zhao, D. T., & Wang, D. (2014). CLIL 观照下的大学英语课程体系建设(College English curriculum construction from CLIL perspective). *现代教育科学(Modern Education Science)*(2), 48-52.
- Zhou, M., & Zhang, J. K. (2014). 论 CLIL 的兴起、进展反对我国民族地区双语教育的启示(The rise and progress of CLIL and the inspiration for the biligual education for ethnic regions). *青海民族研究(Qinghai Journal of Ethnology)*, 25(2), 96-100.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Llinares, A., & McCabe, A. (2020). Systemic functional linguistics: The perfect match for content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*, 1-6.

APPENDICES:

Appendix One

Sample of Language test (Pre-test)

Part I Writing (30 minutes)

Directions: For this part, you are allowed 30 minutes to write a short essay on the use of robots. Try to imagine what will happen when more and more robots take the place of human beings in industry as well as people's daily lives. You are required to write at least 150 words but no more than 200 words.

Part II Listening Comprehension (30 minutes)

Section A

Directions: In this section, you will hear two long conversations. At the end of each conversation, you will hear four questions. Both the conversation and the questions will be spoken only once. After you hear a question, you must choose the best answer from the four choices marked A), B), C) and D). Then mark the corresponding letter on Answer Sheet 1 with a single line through the center.

Questions 1 to 4 are based on the conversation you have just heard.

1. A) Project organizer
B) Public relations officer.
C) Marketing manager.
D) Market research consultant.

2. A) Quantitative advertising research.
B) Questionnaire design.
C) Research methodology.
D) Interviewer training.

3. A) They are intensive studies of people's spending habits.
B) They examine relations between producers and customers.
C) They look for new and effective ways to promote products.
D) They study trends or customer satisfaction over a long period.

4. A) The lack of promotion opportunity.
B) Checking charts and tables.
C) Designing questionnaires.
D) The persistent intensity.

Part III Reading Comprehension (40 minutes)

Section B

Directions: In this section, you are going to read a passage with ten statements attached to it. Each statement contains information given in one of the paragraphs. Identify the paragraph from which the information is derived. You may choose a paragraph more than once. Each paragraph is marked with a letter. Answer the questions by marking the corresponding letter on Answer Sheet 2.

Can societies be rich and green?

[A] “If our economies are to flourish, if global poverty is to be eliminated and if the well-being of the world’s people enhanced—not just in this generation but in succeeding generations—we must make sure we take care of the natural environment and resources on which our economic activity depends.” That statement comes not, as you might imagine, from a stereotypical tree-hugging, save-the-world greenie(环保主义者), but from Gordon Brown, a politician with a reputation for rigour, thoroughness and above all, caution.

[B] A surprising thing for the man who runs one of the world’s most powerful economies to say? Perhaps; though in the run-up to the five-year review of the Millennium(千年的)Goals, he is far from alone. The roots of his speech, given in March at the roundtable meeting of environment and energy ministers from the G20 group of nations, stretch back to 1972, and the United Nations Conference on the Human Environment in Stockholm.

[C] “The protection and improvement of the human environment is a major issue which affects the well-being of peoples and economic development throughout the world,” read the final declaration from this gathering, the first of a sequence which would lead to the Rio de Janeiro Earth Summit in 1992 and the World Development Summit in Johannesburg three years ago.

[D] Hunt through the reports prepared by UN agencies and development groups—many for conferences such as this year’s Millennium Goals review—and you will find that the linkage between environmental protection and economic progress is a common thread.

[E] Managing ecosystems sustainably is more profitable than exploiting them, according to the Millennium Ecosystem Assessment. But finding hard evidence to support the thesis is not so easy. Thoughts turn first to some sort of global statistic, some indicator which would rate the wealth of nations in both economic and environmental terms and show a relationship between the two.

[F] If such an indicator exists, it is well hidden. And on reflection, this is not surprising; the single word “environment” has so many dimensions, and there are so many other factors affecting wealth—such as the oil deposits—that teasing out a simple economy-environment relationship would be almost impossible.

[G] The Millennium Ecosystem Assessment, a vast four-year global study which reported its initial conclusions earlier this year, found reasons to believe that managing ecosystems sustainably—working with nature rather than against it—might be less profitable in the short term, but certainly brings long-term rewards.

[H] And the World Resources Institute (WRI) in its World Resources 2005 report, issued at the end of August, produced several such examples from Africa and Asia; it also demonstrated that environmental degradation affects the poor more than the rich, as poorer people derive a much higher proportion of their income directly from the natural resources around them.

[I] But there are also many examples of growing wealth by trashing the environment, in rich and poor parts of the world alike, whether through unregulated mineral extraction, drastic water use for agriculture, slash-and-burn farming, or fossil-fuel-guzzling(大量消耗)transport. Of course, such

growth may not persist in the long term—which is what Mr. Brown and the Stockholm declaration were both attempting to point out. Perhaps the best example of boom growth and bust decline is the Grand Banks fishery. For almost five centuries a very large supply of cod(鱈魚)provided abundant raw material for an industry which at its peak employed about 40,000 people, sustaining entire communities in Newfoundland. Then, abruptly, the cod population collapsed. There were no longer enough fish in the sea for the stock to maintain itself, let alone an industry. More than a decade later, there was no sign of the ecosystem re-building itself. It had, apparently, been fished out of existence; and the once mighty Newfoundland fleet now gropes about frantically for crab on the sea floor.

[J] There is a view that modern humans are inevitably sowing the seed of a global Grand Banks-style disaster. The idea is that we are taking more out of what you might call the planet's environmental bank balance than it can sustain; we are living beyond our ecological means. One recent study attempted to calculate the extent of this "ecological overshoot of the human economy", and found that we are using 1.2 Earth's-worth of environmental goods and services—the implication being that at some point the debt will be called in, and all those services—the things which the planet does for us for free—will grind to a halt.

[K] Whether this is right, and if so where and when the ecological axe will fall, is hard to determine with any precision—which is why governments and financial institutions are only beginning to bring such risks into their economic calculations. It is also the reason why development agencies are not united in their view of environmental issues; while some, like the WRI, maintain that environmental progress needs to go hand-in-hand with economic development, others argue that the priority is to build a thriving economy, and then use the wealth created to tackle environmental degradation.

[L] This view assumes that rich societies will invest in environmental care. But is this right? Do things get better or worse as we get richer? Here the Stockholm declaration is ambiguous. "In the developing countries," it says, "most of the environmental problems are caused by under-development." So it is saying that economic development should make for a cleaner world? Not necessarily; "In the industrialised countries, environmental problems are generally related to industrialisation and technological development," it continues. In other words, poor and rich both over-exploit the natural world, but for different reasons. It's simply not true that economic growth will surely make our world cleaner.

[M] Clearly, richer societies are able to provide environmental improvements which lie well beyond the reach of poorer communities. Citizens of wealthy nations demand national parks, clean rivers, clean air and poison-free food. They also, however, use far more natural resources-fuel, water (all those baths and golf courses) and building materials.

[N] A case can be made that rich nations export environmental problems, the most graphic example being climate change. As a country's wealth grows, so do its greenhouse gas emissions. The figures available will not be completely accurate. Measuring emissions is not a precise science, particularly when it comes to issues surrounding land use; not all nations have re-leased up-to-date data, and in any case, emissions from some sectors such as aviation are not included in national statistics. But the data is exact enough for a clear trend to be easily discernible. As countries become richer, they produce more greenhouse gases; and the impact of those gases will fall primarily in poor parts of the world.

[O] Wealth is not, of course, the only factor involved. The average Norwegian is better off than the average US citizen, but contributes about half as much to climate change. But could Norway

keep its standard of living and yet cut its emissions to Moroccan or even Ethiopian levels? That question, repeated across a dozen environmental issues and across our diverse planet, is what will ultimately determine whether the human race is living beyond its ecological means as it pursues economic revival.

36. Examples show that both rich and poor countries exploited the environment for economic progress.

37. Environmental protection and improvement benefit people all over the world.

38. It is not necessarily true that economic growth will make our world cleaner.

39. The common theme of the UN reports is the relation between environmental protection and economic growth.

40. Development agencies disagree regarding how to tackle environment issues while ensuring economic progress.

41. It is difficult to find solid evidence to prove environmental friendliness generates more profits than exploiting the natural environment.

42. Sustainable management of ecosystems will prove rewarding in the long run.

43. A politician noted for being cautious asserts that sustainable human development depends on the natural environment.

44. Poor countries will have to bear the cost for rich nations' economic development.

45. One recent study warns us of the danger of the exhaustion of natural resources on Earth.

Part IV Speaking (20 minutes)

Questions:

1. What's the most difficult but rewarding thing you've ever learned? Why?
2. What do you think is the best way to improve your spoken English?
3. What's your opinion about university students' taking part-time jobs? How do you agree with the quote "we learn by doing"?

Discussions:

Reading is one of the major forms of learning and it is a good way to help you keep in touch with the outside world. Have a discussion about the benefits of reading books and recommend one of your favorite books.

Appendix Two

Sample of Language test (Post-test)

Part I Writing (30 minutes)

Directions: For this part, you are allowed 30 minutes to write a short essay on living in the virtual world. Try to imagine what will happen when people spend more and more time in the virtual world instead of interacting in the real world. You are required to write at least 150 words but no more than 200 words.

Part II Listening Comprehension (30 minutes)

Section A

Directions: In this section, you will hear two long conversations. At the end of each conversation, you will hear four questions. Both the conversation and the questions will be spoken only once. After you hear a question, you must choose the best answer from the four choices marked A), B), C) and D). Then mark the corresponding letter on Answer Sheet 1 with a single line through the center.

Questions 5 to 8 are based on the conversation you have just heard.

1. A) His view on Canadian universities.
B) His understanding of higher education.
C) His suggestions for improvements in higher education.
D) His complaint about bureaucracy in American universities.

2. A) It is well designed.
B) It is rather inflexible.
C) It varies among universities.
D) It has undergone great changes.

3. A) The United States and Canada can learn from each other.
B) Public universities are often superior to private universities.
C) Everyone should be given equal access to higher education.
D) Private schools work more efficiently than public institutions.

4. A) University systems vary from country to country.
B) Efficiency is essential to university management.
C) It is hard to say which is better, a public university or a private one.
D) Many private universities in the U.S. are actually large bureaucracies.

Part III Reading Comprehension (40 minutes)

Section B

Directions: In this section, you are going to read a passage with ten statements attached to it. Each statement contains information given in one of the paragraphs. Identify the paragraph from which the information is derived. You may choose a paragraph more than once. Each paragraph is marked with a letter. Answer the questions by marking the corresponding letter on Answer Sheet 2.

Reform and Medical Costs

[A] American are deeply concerned about the relentless rise in health care costs and health insurance premiums. They need to know if reform will help solve the problem. The answer is that no one has an easy fix rising medical costs. The fundamental fix—reshaping how care is delivered and how doctors are paid in a wasteful, abnormal system—is likely to be achieved only through trial and incremental(渐进的) gains.

[B] The good news is that a bill just approved by the House and a bill approved by the Senate Finance Committee would implement or test many reforms that should help slow the rise in medical costs over the long term. As report in *The New England Journal of Medicine* concluded. "Pretty much every proposed innovation found in the health policy literature these days is contained in these measures."

[C] Medical spending, which typically rises faster than wages and the overall economy, is propelled by two things: the high prices charged for medical services in this country and the volume of unnecessary care delivered by doctors and hospitals, which often perform a lot more tests and treatments than patient really needs.

[D] Here are some of the important proposals in the House and Senate bills to try to address those problem, and why it is hard to know how well they will work.

[E] Both bills would reduce the rate of growth in annual Medicare payments to hospital, nursing homes and other providers by amounts comparable to the productivity savings routinely made in other industries with the help of new technologies and new ways to organize work. This proposal could save Medicare more than \$100 billion over the next decade. If private plans demanded similar productivity savings from providers, and refused to let providers shift additional costs to them, the savings could be much larger. Critics say Congress will give in to lobbyists and let inefficient provider off the hook(放过). That is far less likely to happen if Congress also adopts strong "pay-go" rules requiring that any increase in payments to providers be offset by new taxes or budget cuts.

[F] The Senate Finance bill would impose an excise tax(消费税) on health insurance plans that cost more than \$8,000 for an individual or \$21,000 for a family. It would most likely cause Insures to redesign plans to fall beneath the threshold. Enrollees would have to pay more money for many services out of their own pockets, and that would encourage them to think twice about whether an expensive or redundant test was worth it. Economists project that most employers would shift money from expensive health benefits into wages, The House bill has no similar tax. The final legislation should.

[G] Any doctor who has wrestled with multiple forms from different insurers, or patients who have tried to understand their own parade of statements, know that simplification ought to save money. When the health insurance industry was still cooperating in reform efforts, its trade group offered to provide standardized forms for automated processing. It estimated that step would save hundreds of billions of dollars over the next decade. The bills would lock that pledge into law.

[H] The stimulus package provided money to convert the inefficient, paper-driven medical system to electronic records that can be easily viewed and transmitted. This requires open investments to help doctors convert. In time it should help restrain costs by eliminating redundant tests, preventing drug interactions, and helping doctors find the best treatments.

[I] Virtually all experts agree that the fee-for-service system—doctors are rewarded for that the cost of care is so high. Most agree that the solution is to push doctors to accept fixed payments to care for a particular illness or for a patient's needs over a year. No one knows how to make that happen quickly. The bills in both houses would start pilot projects within Medicare. They include such measures as accountable care organizations to take charge of a patient's needs with an eye on both cost and quality, and chronic disease management to make sure the seriously ill, who are responsible for the bulk of all health care costs, are treated properly. For the most part, these experiments rely on incentive payments to get doctors to try them.

[J] Testing innovations do no good unless the good experiments are identified and expanded and the bad ones are dropped. The Senate bill would create an independent commission to monitor the pilot programs and recommend changes in Medicare's payment policies to urge providers to adopt reforms that work. The changes would have to be approved or rejected as a whole by Congress, making it hard for narrow-interest lobbies to bend lawmakers to their will.

[K] The bills in both chambers would create health insurance exchanges on which small businesses and individuals could choose from an array of private plans and possibly a public option. All the plans would have to provide standard benefit packages that would be easy to compare. To get access to millions of new customers, insurers would have a strong incentive to sell on the exchange. And the head-to-head competition might give them a strong incentive to lower their prices, perhaps by accepting slimmer profit margins or demanding better deals from providers.

[L] The final legislation might throw a public plan into the competition, but thanks to the fierce opposition of the insurance industry and Republican critics, it might not save much money. The one in the House bill would have to negotiate rates with providers, rather than using Medicare rates, as many reformers wanted.

[M] The president's stimulus package is pumping money into research to compare how well various treatments work. Is surgery, radiation or careful monitoring best for prostate(前列腺)cancer? Is the latest and most expensive cholesterol-lowering drug any better than its common competitors? The pending bills would spend additional money to accelerate this effort.

[N] Critics have charged that this sensible idea would lead to rationing of care. (That would be true only if you believe that patients should have an unrestrained right to treatments proven to be inferior.) As a result, the bills do not require, as they should, that the results of these studies be used to set payment rates in Medicare.

[O] Congress needs to find the courage to allow Medicare to pay preferentially for treatments proven to be superior. Sometimes the best treatment might be the most expensive. But overall, we suspect that spending would come down through elimination of a lot of unnecessary or even dangerous tests and treatments.

[P] The House bill would authorize the secretary of health and human services to negotiate drug prices in Medicare and Medicaid. Some authoritative analysts doubt that the secretary would get better deals than private insurers already get. We believe negotiation could work. It does in other countries.

[Q] Missing from these bills is any serious attempt to rein in malpractice costs. Malpractice

awards do drive up insurance premiums for doctors in high-risk specialties, and there is some evidence doctors engage in "defensive medicine" by performing tests and treatments primarily to prove they are not negligent should they get sued.

36. With a tax imposed on expensive health insurance plans, most employers will likely transfer money from health expenses into wages.

37. Changes in policy would be approved or rejected as a whole so that lobbyists would find it hard to influence lawmakers.'

38. It is not easy to curb the rising medical costs in America.

49. Standardization of forms for automatic processing will save a lot of medical

40. Republicans and insurance industry are strongly opposed to the creation of a public insurance plan.

41. Conversion of paper to electronic medical records will help eliminate redundant tests and prevent drug interactions.

42. The high cost of medical services and unnecessary tests and treatments have driven up medical expenses.

43. One main factor that has driven up medical expenses is that doctors are compensated for the amount of care rather than its effect.

44. Contrary to analysts' doubts, the author believes drug prices may be lowered through negotiation.

45. Fair competition might create a strong incentive for insurers to charge less.

Part IV Speaking (20 minutes)

Questions:

1. What factors will you take into consideration when you hunt for a job in the future?
2. If you were a boss, what kind of person would you like to employ?
3. How will you keep balance between job and life in the future?

Discussions:

In fact, it is not always easy for people to find a job which both suits your interests and offers a good salary. If you face up with this situation, what do you prefer? A well-paid job you don't like or a low-paid job but meets your interest well?

Appendix Three

Sample of Content test (Pre-test)

一、单项选择题（共 10 题，每题 2 分，共 20 分）

1. 风险按其产生的根源划分为（ ）。
A. 可控制风险和不可控制风险 B. 客观风险和主观风险
C. 宏观风险和微观风险 D. 自然风险和人为风险
2. 下列指标值的组合，预示着企业可能处于过度经营状态的是（ ）。
A. 资金安全率与安全边际率均大于零
B. 资金安全率与安全边际率均小于零
C. 资金安全率大于零，而安全边际率小于零
D. 资金安全率小于零，而安全边际率大于零

二、多项选择题（共 5 题，每题 3 分，共 15 分）

1. 企业集团组建的宗旨或基本目的在于（ ）。
A. 发挥管理协同优势
B. 实现信息共享优势
C. 建立多级法人治理结构
D. 谋求资源的一体化整合优势
2. 股利相关理论的基本依据是（ ）。
A. 股利具有信息价值
B. 市场存在着不确定性
C. 股东存在着对当期收益的偏好
D. 股息对代理成本具有控制作用

三、名词解释题（共 5 题，每题 4 分，共 20 分）

1. 企业并购
2. 资本结构

四、简答题（共 5 题，每题 5 分，共 25 分）

1. 简述非营利组织财务管理的重要性。
2. 简述企业集团财务控制的主要方法。

五、计算题（共 2 题，每题 10 分，共 20 分）

1. 某公司股本为 6000 万股，全部为可流通股，每股净资产为 6 元。假设该股票市价为 8 元，公司回购本公司 40% 的股票。要求：计算分析股票回购对股价的影响。

Appendix Four

Sample of Content test (Post-test)

一、单项选择题（共 10 题，每题 2 分，共 20 分）

1. 下列各项中，符合企业相关者利益最大化财务管理目标要求的是()。

- A.强调股东的首要地位
- B.强调债权人的首要地位
- C.强调员工的首要地位
- D.强调经营者的首要地位

2. 狭义的财务管理环境是 ()。

- A.外部宏观环境
- B.内部管理体制
- C.企业决策体制
- D.企业管理体制

二、多项选择题（共 5 题，每题 3 分，共 15 分）

1. 企业集团组织结构设置的“三权”分立制衡原则的“三权”是包括 ()。

- A.人事权
- B.决策权
- C.执行权
- D.监督权

2. 行为财务学提出下列三种理论 ()

- A.噪声理论
- B.期望理论
- C.过度反映和不足反映理论
- D.随机游走理论

三、名词解释题（共 5 题，每题 4 分，共 20 分）

1.要约收购

2.安全边际率

四、简答题（共 5 题，每题 5 分，共 25 分）

1.简述投资风险的衡量标准。

2.成功的风险投资运作主要取决于哪些因素？

五、计算题（共 2 题，每题 10 分，共 20 分）

1. 2012 年底，A 公司拟对 B 企业实施吸收合并式收购。根据分析预测，购并整合后的 A 公司未来 3 年中的股权资本现金流量分别为 7000 万元、5000 万元、8000 万元，3 年后的股权资本现金流量将稳定在 5000 万元左右；又根据推测，如果不对 B 企业实施购并的话，未来 3 年中 A 公司的股权资本现金流量将分别为 5000 万元、4000 万元、4200 万元，3 年后的股权现金流量将稳定在 3300 万元左右。购并整合后的预期资本成本率为 6%。要求：采用现金流量贴现模式对 B 企业的股权现金价值进行估算。

Appendix Five

Guide for Interview Questions (Student) (English Version)

1. What do you think of doing financial management in English?
2. What do you think is the purpose of learning financial management in English?
3. Did you notice any changes: improvement or deterioration in your subject learning?
4. Did you notice any changes: improvement or deterioration in your linguistic learning?
5. Could you describe several activities you attended during the course? For example, in-class discussion, assessment, group-work?
6. What language did you speak when doing these activities with your peers and the teacher?
7. What do you find more challenging- the subject or the language?
8. What efforts have you made to overcome these challenges?
9. During the lessons, did you feel any cultural differences when learning subject in English?
10. What is your advice for the students who will study subjects in English?

Appendix Six

Guide for Interview Questions (Teacher) (English Version)

1. What do you think of teaching financial management in English?
2. What do you think is the purpose of teaching financial management in English?
3. What are the advantages and disadvantages of a CLIL classroom?
4. What benefits do you see in studying financial management in English for the students? For example, to their knowledge, future education and career.
5. What kinds of teaching skills did you adopt in classroom practice? For example, translanguaging.
6. How did you balance the teaching focus of content and language?
7. Do you put emphasis on introducing or explaining cultural differences?
8. How about learners' cognitive development in CLIL classroom when they encountered content knowledge in a foreign language environment?
9. How do you evaluate this CLIL practice?
10. Do you have any suggestions or comments on designing a CLIL classroom?

Appendix Seven

Sample of Interview (Student)

- R:** What do you think of learning financial management in English?
- S:** It will be useful in the future. For example, when we go abroad, it will be easier for us to get used to the English environment. Er, English seems like a common language.
- R:** What do you think is the purpose of learning financial management in English?
- S:** I guess, to improve my English as well as subject knowledge. At the very beginning of the course, the teacher introduced that it would improve our English and financial knowledge at the same time. I think so.
- R:** Did you notice any changes: improvement or deterioration in your subject learning?
- S:** Certainly, I did a good job in my subject learning. I got a very good score. Of course, I made great efforts to learn financial management in English. But I think that the way of assessment is not effective, because it relied too much on paper work.
- R:** Did you notice any changes: improvement or deterioration in your linguistic learning in terms of listening, speaking, reading and writing?
- S:** Er. The situation is rather complex to say. Generally speaking, I think, I made a little bit improvement. Both listening and reading is fine to me. Concerning speaking, CLIL is a good way to practice speaking because we had to do a lot of speaking tasks in the classroom, for example, topic discussions, peer collaborations and team presentations. But writing, I found

it too difficult for me to complete these finance-focused tasks.

R: Could you describe several activities you attended during the course? For example, in-class discussion, assessment, group-work?

S: We did group discussions regularly. The teacher gave us a topic, sometimes a controversial topic, and subsequently we responded in English. If someone couldn't solve the task, we helped each other, do explanation and paraphrasing. We solve interesting tasks, translate and laugh at some of the translations.

R: What language did you speak when doing these activities with your peers and the teacher?

S: It depends. Sometimes, the teacher started in English, and made some explanations in Chinese. To me, I tried to use English. But you know, sometimes, it was difficult for me to express my ideas in English, so I used Chinese on some occasions. I felt ease at that time.

R: What do you find more challenging- the subject or the language?

S: Subject, of course. Terminologies, expressions, concepts, sentences...

R: What efforts have you made to overcome these challenges?

S: It is a long story. I made great efforts. For example, When I couldn't understand the complex descriptions, such as the relationships between current assets and current liabilities, I had to divide these descriptions into shorter and easier parts, and tried to find some similar explanations in Chinese subsequently. Then, I organized the Chinese versions of description. It was really hard for me to understand them as a whole in English.

R: During the lessons, did you feel any cultural differences when learning subject in English?

S: Not often. The teacher sometimes introduced some multicultural issues.

R: What is your advice for the students who will studying subjects in English?

S: Efforts. If you do not put effort at the beginning, it will be difficult indeed. They will struggle like me. I'd like to call on them to study CLIL, because it is useful and interesting.