Universitety of South-Eastern Norway
Faculty of Humanities, Sports, and Educational Science

Doctoral thesis no. 71
2020

Karina Rose Mahan

Teaching Content and Language Integrated Learning (CLIL)

Classroom practices and student perspectives in three Norwegian classrooms

Karina Rose Mahan

Teaching Content and Language Integrated Learning (CLIL)

Classroom practices and student perspectives in three Norwegian classrooms

A PhD dissertation in
pedagogical resources and learning processes in kindergarten and school

Faculty of Humanities, Sports, and Educational Science
University of South-Eastern Norway
Horten, 2020

Doctoral dissertations at the University of South-Eastern Norway no. 71
ISSN: 2535-5244(print)
ISSN: 2535-5252 (online)

ISBN: 978-82-7860-416-8 (print)
ISBN: 978-82-7860-417-5 (online)

Print: University of South-Eastern Norway

Acknowledgements

At times, pursuing a doctorate may feel like a blundering through a haze. However, with the right people to guide you, you will end up on the right path. As for those who guided me along my path, I would first and foremost like to thank my main advisor, Associate Professor Karianne Skovholt. Her positive attitude, excellent advice, careful reading of my texts, and caring nature have helped me in more ways than I can express. I am also indebted to my co-advisor, Professor Tine Prøitz, for her many helpful comments and sharp observations. I was lucky enough to have many qualified professors help lead me down my path as well. A special thanks to Professors Glenn Ole Hellekjær, Liss Kerstin Sylvèn, Tarja Nikula, and Dina Tsagari for their great advice and guidance.

Secondly, I would like to thank Professor Yolanda Ruiz de Zarobe for inviting me to the University of the Basque Country. She was an amazing advisor, guide, and friend during my stay in the Basque Country. Her light tone and reminder to enjoy myself in Spain made me feel at home. Her many years as a CLIL researcher provided deep insights into my field. No research stay is complete without friends; a big thanks to my special Ph.D. guides and pintxos friends: Asier Calzada Lizarraga, Diego Llamazares de Miguel, Katherine Marcoux, Henri Castleberry, Sergio López Sancio, and Olympia Simantiraki.

I received support from a number of individuals at my own institution. I would especially like to thank Professor Ragnhild Elisabeth Lund. I have immense respect for her willingness to help anyone who needs it. Her patience, giving nature, and sound advice helped me develop as both a teacher and researcher. Likewise, I would like to thank Professor Tony Burner and his reading group, PASIE, for reading and supporting all of my articles. Associate Professors Henrik Bøhn, Magne Dypedahl, Maria Casado Villanueva, and Christian Carlsen taught me a great deal about being an academic through our many conversations. My thanks are particularly due to Professor Thomas Moser. Despite his many commitments, he always had time to read my papers and provide comments, advice, and a laugh or two. I would like to mention my previous department head, Doctor Piotr Garbacz, who believed in me and gave me my first step into academia. The

University of Southeast-Norway also has many wonderful Ph.D. candidates. A special thanks to Birgitte Kasin Hønsvall, Christian Bjørge Thorsen, Burak Cunbul, Andreas Larsson, Nahum Misael Torrez, Alessandra Dieudè, Helga Norheim, and Christine Rendahl Stenersen for being the world's best procrastinators. Another thanks goes to my friends Kim Elisabeth Bjørnstad, Ka Xuong Lay, and Jan Eirik Jakobsen for their continuous support throughout the years. Tønsberg felt much less lonely with you. It was also nice to come home to familiar faces. Thank you, Camilla Osmundsen, Regina Zheng, Abeera Akhtar, and Kristianne Rakstang for being my personal cheerleaders and making Stavanger still feel like home.

The research school NAFOL provided excellent courses, homework, and seminars, which proved to be invaluable to my thesis writing. I would like to thank Professors Kari Smith and Marit Ulvik for inviting me to NAFOL, their positive outlook, and sponsoring several events. Lastly, a school cannot be complete without its students. Thanks to fellow students Anne-Grete Kaldahl, Mette Helleve, Jannike Hegdal Nilssen, Åsve Murtnes, Liv Gardsjord Lofthus, Thomas Eide, Ingvild Bjørkeng Haugen, and Irina Amdal, among others, for the last drink and laugh after a long conference day.

A thank you to Professor Kirsti Klette for inviting me to work alongside the LISA/LISE projects at the University of Oslo. With their help, I collected data and wrote my first article. My thanks are also due to my co-authors Associate Professor Lisbeth Brevik and Professor Marianne Ødegaard for encouraging me and co-writing a solid first article. Thanks to Bjørn Sverre Gulheim, Kjetil Birkeland, and the Teaching Learning Video Lab for helping me collect and code data. I would also like to thank Jennifer Luoto for her invaluable input on the coding process of my second article, and being an amazing friend.

Research in the social sciences demands a lot of its participants. I would therefore like to thank all the teachers, students, and school leaders who spent their valuable time answering my questions, inviting me to their classrooms, and allowing me to observe their teaching and learning. There would be no Ph.D. without you, and you have my eternal gratitude.

Finally, and most importantly, I would like to thank those closest to me. My parents, Grethe and James, sister and brother, Sofia and Hasham, and grandparents Torgunn, John, Ragnar, and Bodil, were my rock in times of trouble. Their unwavering belief in me propelled me forward every day. I am indebted to Miguel Àngel Galicia Gorostieta for his unconditional love and support.

I dedicate this thesis to my mom. You believed in me every step of the way. I promised that the first book I ever wrote would be for you. I hope I made you proud.

Karina Rose Mahan, Vitoria-Gasteiz, April 2019

Abstract

Keywords: *Content and Language Integrated Learning (CLIL) teaching, CLIL pedagogy, English as a Second Language (ESL), Bilingual Teaching, Scaffolding, Student Perspectives*

This thesis investigates Content and Language Integrated Learning (CLIL) teaching practices with English L2 as the medium of instruction in three Norwegian secondary schools. The primary data are video observations of CLIL teaching and student questionnaires. The thesis is article-based, comprising three articles and an extended abstract. The extended abstract provides the following: the background and development of CLIL as a teaching methodology; a review of research on CLIL teaching; the methods and research design used in this thesis; a summary; and discussion of the results. This thesis is positioned within a sociocultural view of learning, emphasizing the importance of teacher-student interactions to understand CLIL teaching in practice.

Article I investigated how lower secondary CLIL teachers taught their subject in terms of content and language. By filming four hours of CLIL teaching in science and mathematics and comparing it to the students' English lessons, this study sought to characterize the observed CLIL teaching. The coding manual Protocol for Language Arts Teaching Observation (PLATO) was used to analyze the data. The findings indicated that CLIL teaching was content-driven, intellectually challenging and had clear instructional explanations and consistent language support. However, students were provided few opportunities to read and write.

Article II focused on how CLIL teachers in upper secondary school scaffolded learning during their lessons. Three CLIL teachers in science, geography, and social science were filmed for four hours each (N=12). PLATO was used to identify instances of scaffolding strategies. The findings suggested that the CLIL teachers used a wide range of scaffolding strategies to help their students comprehend material but few metacognitive scaffolding strategies to help students solve tasks. There were differences between scaffolding in the natural and social sciences. The natural sciences provided relatively more visual support;

the social sciences provided relatively more discussion time and allowed for longer student replies. This implies that subjects may provide different types of support for second language learners.

Article III examined how upper secondary students perceived their CLIL teaching. Fifty students from two CLIL programs were distributed two questionnaires: one that asked students why they chose CLIL and how they perceived it, and another that asked the students to assess their science teaching. The findings revealed that the upper secondary students mainly chose CLIL because they perceived English as important to succeeding in future studies and work. They perceived their CLIL teaching as mostly positive, citing that they improved their English, enjoyed the multicultural classroom environment, and felt motivated. However, some students found the absence of L1 (Norwegian) problematic, felt excluded from the school environment, and struggled cognitively with learning their subject through the L2. Students perceived that their CLIL science teachers clarified material, often conferred with and intellectually challenged them, but that the students had little decision-making regarding input in activities.

Overall, the three articles contribute to a deeper insight into how CLIL is taught in secondary schools in Norway. The findings show that the observed CLIL teaching was largely effective; the teachers manage to convey their subject through the L2 (English); students feel intellectually challenged; there is evidence of scaffolding; and CLIL is perceived as a positive experience by most students. The comparison across subjects also suggests that the natural sciences subjects provide a multitude of visual aids and language support for second language learners. However, the studies also point to areas of challenge for further development of CLIL in Norway. The lack of reading and writing in CLIL subjects emphasizes that the English language subject may provide an important resource to develop these language skills further. There are also a number of problems identified by CLIL students that need to be addressed, such as how to balance L1 in CLIL classrooms, to ensure that Norwegian students are equipped for future studies in Norwegian as well as English. Finally, CLIL programs are in danger of being isolated from

the rest of the school, and this is a challenge that needs to be examined by stakeholders to ensure that all students feel included in the school.

List of papers

Article 1

Mahan, K. R., Brevik, L. M., & Ødegaard, M. (2018). Characterizing CLIL teaching: New insights from a lower secondary classroom. *International Journal of Bilingual Education and Bilingualism*. doi: 10.1080/13670050.2018.1472206

Article 2

Mahan, K. R. (2020). The comprehending teacher: Scaffolding in Content and Language Integrated Learning (CLIL). The Language Learning Journal. doi: 10.1080/09571736.2019.1705879

Article 3 Under review, not included in online publication

Mahan, K. R., Norheim, H. (Under review). Something new and different: Student perceptions of Content and Language Integrated Learning (CLIL). Submitted to *ELT Journal.*

List of tables

List of figures

Abbreviations

BICS	Basic Interpersonal Communication Skills
CALP	Cognitive Academic Language Proficiency
CA	Conversation Analysis
CBI	Content-Based Instruction
CLIL	Content and Language Integrated Learning
EFL	English as a Foreign Language
ELL	English Language Learners/Learning
ELT	English Language Teaching
ESL	English as a Second Language
FFI	Form-focused instruction
IRF	Initiation, Response, Feedback
LK06	Læreplanverket for kunnskapsløftet [The Knowledge Promotion curriculum reform]
L1	First language
L2	Second or foreign language
LISA	Linking Instruction and Student Achievement (research group at the University of Oslo)
LISE	Linking Instruction and Student Experiences (research group at the University of Oslo)
MMR	Mixed methods research
PLATO	Protocol for Language Arts Teaching Observation
RQ	Research question
SLA	Second Language Acquisition
TESOL	Teaching English as a Second Language

UDIR	Utdanningsdirektoratet [Norwegian Directorate for Education and Training]
UiO	University of Oslo
USN	University of South-Eastern Norway
ZPD	Zone of Proximal Development

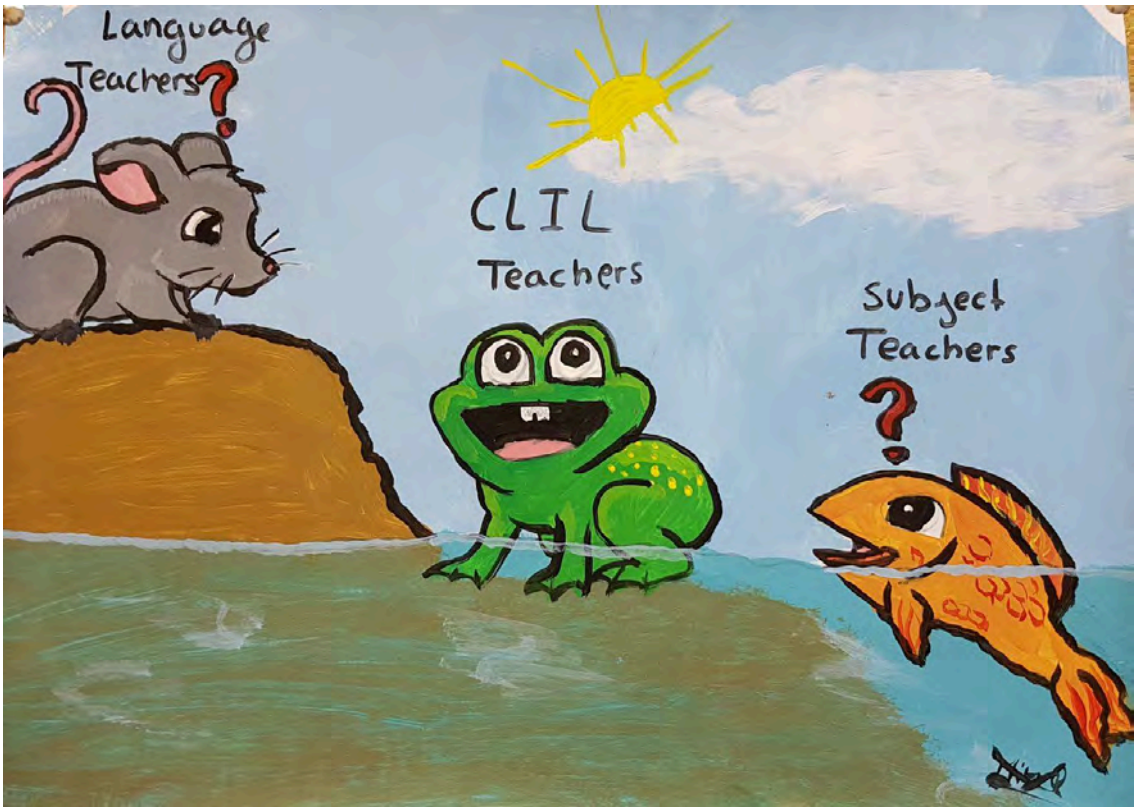
Table of contents

Part I

Extended Abstract

1 Introduction

"Language transcends subject matter and curricular boundaries, since it pervades all of education." (van Lier, 2004, p. 20)



(Art: Miren Olaizola)

The present thesis, which is article-based, investigates Content and Language Integrated Learning (CLIL) teaching practices in three Norwegian secondary schools. The aim is to identify how teachers in these schools teach CLIL in their content subject with English L2 as the primary language of instruction. The thesis is two-pronged: using a coding manual, it considers *how two lower secondary CLIL teachers integrate or balance content and language* (Article I) and *how three upper secondary CLIL teachers scaffold learning* (Article II). Then, using questionnaires, the thesis turns to students from two upper secondary CLIL classrooms to test observations from the previous studies to see *how students perceive CLIL teaching* (Article III). This chapter

provides background information to explain why these topics were chosen, some of the basic tenets of CLIL, and the Norwegian educational and linguistic context.

1.1 Background

As we move further into the 21st century, countries are finding it more and more necessary to implement bilingual education. In the European Union alone, three fifths of students are learning two or more foreign languages in upper secondary school (Eurostat, 2018), meaning that schools must somehow accommodate all of these foreign language subjects. Bilingual education provides a way to integrate language into the curriculum, minimizing the number of hours in language subjects and/or boosting language proficiency (Simensen, 2002). This thesis uses the following definition for bilingual education: when teachers and students "*include the use of multiple multilingual practices that maximize learning efficacy and communication*" (García, 2009, p. 9). There is nothing new about bilingual education; it is documented as far back two thousand years ago when the Ancient Romans taught Greek to their children (Coyle, Hood, & Marsh, 2010, p. 2). However, bilingual practices are still expanding worldwide in different forms. Each continent has its own bilingual education, from the French immersion classes in Canada (Cummins, 1998) to the English as a Medium of Instruction (EMI) approach in Africa (Brock-Utne, 2007). Bilingual programs spring from different needs and contexts. One type of bilingual education that is quickly gaining ground is **Content and Language Integrated Learning** (Cenoz, Genesee, & Gorter, 2014, p. 243). CLIL is commonly defined as "an additional language integrated in a non-language subject" (Coyle et al., 2010, p. 1). It is the European label for bilingual education (Georgiou, 2012, p. 495) and is viewed by many as a response to trends toward internationalization and globalization (Dalton-Puffer, 2007; Ruiz de Zarobe, 2013). Research is beginning to show that the CLIL methodology can give students access to aspects of language learning that language education cannot, such as authentic use of L2 to discuss other subject matter (Nikula, 2007; Nikula & Moore, 2019). The appeal of CLIL is that it can combine English L2 with a variety of content subjects to develop communicative competence in the L2. However, results have varied and are seemingly dependent on national contexts

(Sylvén, 2013, p. 301). Some countries report little success with CLIL teaching because the overall English proficiency among teachers and students is too low (e.g., Lo, 2015). In the Nordic countries, where exposure to and proficiency in English is already high, researchers are beginning to question if we "need" CLIL and what it can contribute to education (e.g., Olsson & Sylvén, 2015; Sylvén, 2019b). In Norway, there is growing concern that English is beginning to dominate as the primary language of the country, leading to "domain loss" (loss of L1 terminology in certain fields) (The Language Council of Norway, 2017a). CLIL is offered at a number of secondary schools in Norway, yet there is limited research on CLIL in Norway—especially research that emphasizes CLIL teaching. This thesis digs deeper into these issues by researching how CLIL teachers teach in terms of content, language, and scaffolding, and how students perceive these teaching practices. The current thesis is situated in the field of English didactics since CLIL presents an innovative way of learning English outside of the English language classroom (Marsh, 2002, p. 53).

In Norway, CLIL is one of the few types of bilingual education offered in secondary schools. CLIL is practiced at three to four percent of secondary schools and often presented as an international program (Svenhard, Servant, Hellekjær, & Bøhn, 2007, p. 141). When I was a CLIL student myself in 11th grade, I found it fascinating how my CLIL science teacher managed to teach his subject in English to Norwegian students. How did he manage to convey the subject matter in English? How did he teach so many students at different levels of English and science? Lastly, why did we feel it was different from "regular" science class? Researchers are beginning to recognize the precarious and important role that CLIL teachers play in bilingual education (cf. Pérez-Cañado, 2016). CLIL teachers are content teachers, but they must also possess specialized knowledge about how to teach through a second language. The frog in the picture above illustrates this. It points to a new generation of teachers who recognize linguistic as well as subject needs in education. CLIL teachers are more than content teachers but not exactly language teachers either. The field of CLIL is beginning to explore these "frogs" who must balance both content and language in bilingual education. It is the position of this thesis that by seeing CLIL teachers in action, we can understand how they overcome the many challenges

of teaching through an L2. Furthermore, by understanding students' reactions to these teaching practices, we can pinpoint what works and what does not.

1.2 An overview of Content and Language Integrated Learning (CLIL)

1.2.1 Historical and political perspectives

CLIL was practiced in diverse forms in Europe during the last half of the 1900s (for examples, see Marsh, Maljers, & Hartiala, 2001). Researchers understood these bilingual practices as very different from one another as they were taught in different ways and motivated by varying factors (Marsh, 2002, p. 50). The term *CLIL* was coined officially in 1994, providing a common term and identity for European bilingual education (Coyle et al., 2010, p. 3). CLIL received added impetus from a White Paper promoting the MT+2 (mother tongue and two additional languages) goal: that students of each European country should speak two foreign languages fluently in addition to their mother tongue (European Commission, 1995, p. 44). In order to achieve this goal, the European Union sponsored several CLIL initiatives throughout Europe (Baetens Beardsmore, 2009). Though CLIL has many grassroots programs, there has been an increase in top-down endeavors in which nations create CLIL policies and/or programs for schools (for a comprehensive overview, see Eurydice, 2006).

Researchers argue that although CLIL draws on tenets of immersion and bilingual education in North America (Coyle, 2007, 2008; Sylvén, 2019a), Europe felt the need to mark its bilingual education as unique and independent from other movements. CLIL has its own context, learners, and teachers. It is currently spreading to areas such as Latin America, Asia, and Oceania (Dalton-Puffer, 2011, p. 184) but remains European in the sense that it "has been energized by European language policy and ideology and has in turn energized implementations of these policies at local or regional levels" (Dalton-Puffer, Llinares, Lorenzo, & Nikula, 2014, p. 214).

Bilingual education in Europe was originally used to boost proficiency in marginalized languages, such as Welsh and Basque, in their respective countries or territories (Coyle, 2008, p. 98).[1] However, with the growing need for English competence, English has taken over as the most common language in CLIL (Eurydice, 2006, pp. 18–19). The particular appeal of CLIL for English language learning is that it creates naturalistic environments for using the L2, targeting communication above all else (Dalton-Puffer, 2007, 2011).

1.2.2 Defining CLIL

One of the most heated debates in CLIL research centers around being able to define what constitutes CLIL (Cenoz et al., 2014; Lasagabaster & Sierra, 2009a). There are a plethora of definitions of CLIL, ranging from narrow to broad. Researchers who expand CLIL to an umbrella term view it as inclusive of all types of bilingual education (e.g., Dalton-Puffer et al., 2014). On the one hand, since the boundaries between terms such as *immersion*, *content-based instruction*, and *CLIL* are fuzzy, it is difficult to argue how they differ under given circumstances. On the other hand, broadening CLIL to including all types of bilingual education "makes it so general as to lack practical or theoretical utility" (Cenoz et al., 2014, p. 246). In this vein, I choose to delimit CLIL to the following definition in this thesis: *teaching content subjects in another language over a period of time in a public school system* (definition taken from Mahan, Brevik, & Ødegaard, 2018, p. 4). I have chosen to delimit CLIL to the public school system as the national curriculum is followed therein. Special emphasis will be put on research in Europe as it is the most relevant educational context for this thesis. Furthermore, the primary focus will be placed on teaching through English as L2 since it is the language primarily used for instruction in CLIL programs in Norway (The Norwegian National Center for Foreign Languages in Education, 2011a, 2011b). Because this thesis is empirically grounded, CLIL is considered here a *methodology* that is realized through teaching practices.

[1] Although in the Basque context, the programs were referred to as bilingual or plurilingual education (Cenoz, 2015, p. 10)

On some points, it may be difficult to define what CLIL is, but there is a growing consensus on certain of its core characteristics. Firstly, CLIL is dual-focused on content and language (Met, 1999). Dalton-Puffer (2007) asserts that there should be concrete language learning goals in the curriculum (p. 6). This runs contrary to the immersion approach, in which it is assumed that students will learn the language of instruction through osmosis (see the theoretical elaboration in Section 2.2.2). CLIL tends to be conducted in a lingua franca (e.g., English, Spanish, German) with teachers and students who share a common L1, often the majority language of a country (Ball, Kelly, & Clegg, 2016; Dalton-Puffer, 2011). This stands in contrast to bilingual movements that attempt to revitalize endangered languages or bilingual programs for immigrant students (García, 2009). CLIL teachers in secondary schools are often bilingual in the majority language and the L2 and are content teachers (e.g., science, math) who may or may not have knowledge of language teaching methodologies (Cenoz et al., 2014, p. 252). Dalton-Puffer et al. (2014) argue that while some bilingual education programs replace language subjects with bilingual teaching, CLIL subjects tend to be framed as content subjects and do not replace the language subjects (p. 215).

Cenoz et al. (2014) highlight that the goals of CLIL students tend to be more pragmatic than ideological (p. 248). Whereas in French immersion programs in Canada, the main goal of bilingual education is for the students to become near-native speakers, CLIL practitioners will settle for communicating in the L2 (Lasagabaster & Sierra, 2009a, p. 372). Finally, CLIL tends to be taught in primary and secondary education, whereas teaching at university level often falls under the term *English as a Medium of Instruction* (Ball et al., 2016, p. 285). Section 1.2.3 will describe some of the characteristics of CLIL in Norway specifically.

1.2.3 CLIL teaching

The repertoire of strategies and methods a teacher has at her disposal is a key factor of successful learning (cf. Grossman & Morva, 2008; Hattie, 2009). Accordingly, this thesis focuses primarily on what the teacher does and how students perceive it. A general consensus exists that CLIL teaching is more challenging than non-CLIL teaching (Dalton-Puffer, 2011; Pérez-Cañado, 2016). CLIL teachers often teach through their L2, involving more work in preparation

of materials (Nikula, 2010; Pena Díaz & Porto Requejo, 2008). Teachers must also transfer L1 English materials into their own multilingual contexts and may be uncertain about their proficiency in the L2 (Pérez-Cañado, 2016, p. 206). Very few countries offer CLIL teacher programs (Eurydice, 2006, p. 42). In countries such as Norway, CLIL teachers operate without teacher training, national policies, school guidelines, or traditions of bilingual teaching. Under such conditions, researchers are beginning to ask if this is fertile ground to sow CLIL, or if there are more effective ways of cultivating English proficiency (Sylvén, 2013, p. 316).

Researchers consider CLIL to be a distinct and promising pedagogy (Ball et al., 2016; Coyle et al., 2010), but more work is needed to identify what CLIL teachers do and how they can overcome challenges in teaching through an L2 in specific contexts (Dalton-Puffer, 2011; Georgiou, 2012; Pérez-Cañado, 2012). There is limited research that describes CLIL teaching in Norway, and CLIL practitioners are calling for more research on naturalistic[2] CLIL teaching (for a full overview, see Section 3.1). This is the main area of interest in my thesis. Since its theoretical rationale for learning is sociocultural, learning is viewed as mediation between teacher and student (see Section 2.1). For this reason, it was critical to include the student perspective as well. This has been addressed by eliciting information from CLIL students as well.

[2] *Naturalistic data* refers to "video data on naturally-occurring and naturally-organized actions and their linkages" (Watson, 1992, p. 262). In the context of teaching, this would mean data stemming from classroom interaction with minimal interference from the researcher.

1.3 The Norwegian context

The current section explains the background of the language situation in Norway and the Norwegian educational system for international readers.

1.3.1 The language situation in Norway

Norway has a complex language situation. Norwegian and Sámi are the official languages of Norway. Although the total percentage of Sámi speakers in Norway is unknown, it is estimated to be between ten to twenty thousand out of a Norwegian population of five million people (Kulbrandstad, 2003). Sámi is mainly spoken in Northern Norway. Norwegian is the majority language and has two writing systems: *bokmål* (lit. "Book language") and *nynorsk* (lit. "New Norwegian"). Bokmål is the preferred main written language, used by 85–90% of the population (Vikør, 2012), yet both writing systems are obligatory subjects in schools. English is taught from 1st grade (ages 6–7), and a second foreign language is taught from 8th grade (ages 13–14). The most common foreign languages chosen by secondary school students are Spanish (37% of students), German (28%), and French (13%) (The Norwegian National Center for Foreign Languages in Education, 2019, p. 2). Additionally, there are national minority languages (Kven language, Finnish, Romani, and Yiddish), and other minority languages (e.g., Urdu, Arabic, Swedish, Danish, and Vietnamese). Norway lacks a comprehensive statistical overview of languages spoken, but approximately 18% of the Norwegian population consists of first- and second-generation immigrants (Statistics Norway, 2019a). Over 50% of immigrants are European (especially from the Nordic countries); 32% of immigrants are from Asia (including refugees from Syria and Afghanistan), and 14% are from Africa (Sandnes, 2017, p. 18).

As can be seen from the above, Norway has a variety of languages to manage. Which language should be prioritized in schools is a politically loaded question; the Language Council of Norway has set forth numerous guidelines for how and when languages should be taught (e.g., The Language Council of Norway, 2017b). How much English should be taught is a heated subject. Some Norwegians fear that English may eventually replace Norwegian (The Storting, 2008, p. 15). Minority languages, particularly from non-Western countries, are generally not prioritized.

The official policy of The Norwegian Ministry of Education and Research is only to give students with an immigrant background mother tongue instruction until they have reached a sufficient level of Norwegian (The Norwegian Directorate for Education and Training, 2016).

1.3.2 English in Norway

The status of English in Norway is another debate. A language that is not considered a native tongue is normally labeled as either a second or foreign language, depending on its status in the host country (Kachru, 1985). *English as a Second Language* (*ESL*) is a term used to describe English's status in countries where it is an official language, most likely the result of colonialism (Seargeant & Swann, 2012, p. 28). *English as a foreign language* (*EFL*) is a term that refers to contexts where English has no official status (Seargeant & Swann, 2012, p. 28). Countries where English instruction is labeled as EFL are typically countries where students are mostly exposed to English in the classroom. This sociolinguistic distinction is indispensable to this thesis since CLIL involves students in both EFL and ESL contexts. The status of English in the country needs to be taken into account when investigating L2 English proficiency, and the Nordic countries are veering toward an ESL context (Sylvén, 2019a, p. 6).

The English language is at crossroads in Norway. Norway documents one of the highest proficiency levels in the world compared to other countries where English is neither a first nor official language (Education First, 2018). Moreover, the Norwegian Directorate for Education and Training no longer classifies English as a foreign language. Norwegians use English daily in their lives (Hellekjær, 2007; Ibsen, 2002; The Language Council of Norway, 2017a), with some preferring English over Norwegian (see Brevik & Hellekjær, 2018). Despite this, English is not an official language in Norway. Thus, expectations regarding English and its role in society are unclear. Some researchers argue that labeling English a foreign language in Norway is a misnomer as it no longer feels foreign (Brevik, 2015; Rindal, 2013). Other researchers counterargue that since English holds no official status, it would be wrong to label it a second language (Burner, 2016; Simensen, 2014). In line with the unclear role of English in Norway, this thesis will consistently use the term *L2* to refer to English speakers in Norway.

1.3.3 The Norwegian school system

The Norwegian school system is divided into elementary school (Grades 1–7, ages 6–13), lower secondary school (Grades 8–10, ages 14–16), and upper secondary school (Grades 11–13, ages 17–19). Grades 1 through 10 are compulsory, and 93% of students elect to attend upper secondary school (Statistics Norway, 2019b). Which upper secondary school students enter is often based on their final grades from 10th grade. In upper secondary school, students can choose between general study programs (that grant university admission) and vocational programs (that prepare them for a vocation but not necessarily grant university admission). To date, there are five main general studies programs and eight main vocational programs in Norwegian upper secondary schools (Vilbli.no, 2019).

English has been taught as a compulsory subject in Norway since 1959 (Svenhard et al., 2007, p. 139). Currently, it is compulsory in Grades 1–10 (The Norwegian Directorate for Education and Training, 2013a). On average, students receive approximately 200 hours of English lessons a year (The Norwegian Directorate for Education and Training, 2019b). If students elect to go to upper secondary school, English is also compulsory in Grade 11 but may be an elective course in Grades 12–13, depending on the study program. It is important to note that the national curriculum emphasizes both content *and* language in the English language course of study. Students must learn about history, literature, culture, and society in English-speaking countries as well as learning English as a language (The Norwegian Directorate for Education and Training, 2006, 2013a). At the end of upper secondary school, all Norwegian students are expected to be at least at level B2 or C1 of the Common European Framework of Reference; this is a requirement for studying at university level in Norway (Norwegian Universities and Colleges Admission Service, 2019).

1.3.4 English as the language of instruction and CLIL in Norway

Norwegian, and occasionally Sami in some schools, are the official languages of instruction in Norwegian public schools. Schools must apply to the Norwegian Ministry of Education and Research if they wish to use another language of instruction, and usually only private schools do this. Schools that have English as the main language of instruction are typically private,

international schools that follow their own curricula. Alternatively, some public schools offer the International Baccalaureate (IB), which is a two-year program in upper secondary school (Grades 11–13). CLIL is another option for students who wish to stay in the Norwegian public school system but receive more English education than the average Norwegian. Unlike the international schools or the IB, CLIL follows the Norwegian national curriculum, with Norwegian teachers who may or may not have a relevant background in English language education.

The first CLIL initiative in Norway was sponsored by the Ministry of Education and Research in 1993 (Svenhard et al., 2007, p. 139). The classes taught were history, religion, tourism, and restaurant and food processing, and the language of instruction was English (Svenhard et al., 2007, p. 139). CLIL is largely a grassroots initiative in Norway, sparked by either individual teachers, the school, or the county. Schools that provide CLIL often label themselves as innovative upper secondary schools, and the majority of these programs are in English (Svenhard et al., 2007; The Norwegian National Center for Foreign Languages in Education, 2011b). The reason that CLIL is of interest in Norway now is that it provides a viable (and free) option for students who wish to learn content subjects through English, to prepare themselves for university and study abroad. However, in 2017, a new governmental policy underlined that only Norwegian and Sami should be the languages of instruction in Norwegian public schools (The Norwegian Directorate for Education and Training, 2017). A few CLIL programs have been restricted as a result of this, and its future in Norway is unclear.

CLIL in Norway comprises small groups of schools, programs, teachers, and students; therefore, I choose not to focus on the generalizability of this thesis.[3] As with most qualitative research that entails rich data on a small group of individuals (Bryman, 2012, p. 392), the current thesis emphasizes contextual uniqueness and significance, in this case focusing on CLIL in three secondary schools in Norway. By examining this phenomenon in depth, it is hoped that the findings may assist policymakers in Norway to make more informed decisions regarding the future of CLIL.

[3] Although results in Norway may very well be transferable to similar Nordic countries, such as Sweden or Finland

1.4 Overarching aims and research questions

As I will elaborate upon in the literature review, there is a need for more research on CLIL in the Norwegian context (see Section 3 for a full overview). CLIL teachers face numerous challenges teaching through the L2. There are questions revolving around the necessity of CLIL's existence in Nordic countries and what CLIL can contribute to societies where people are already proficient in English. There is limited knowledge on how CLIL teachers teach in Norway, and how students perceive the instruction (this thesis is the first to film Norwegian CLIL classrooms). For all these reasons, the thesis begins with a rather broad research question to capture the most salient aspects of CLIL teaching.

The main research question (RQ) guiding this thesis is as follows:

What characterizes CLIL teaching in three Norwegian secondary classrooms?

After a literature review (see Section 3), content, language, and scaffolding were identified as areas of CLIL teaching that could be of interest for further investigation. The main RQ is therefore divided into the following sub-research questions:

Sub-RQ1: What characterizes CLIL teaching in terms of content and language? (Article I)

Sub-RQ2: What characterizes CLIL teaching in terms of scaffolding? (Article II)

The student perspective was included as well to see how the students experienced CLIL teaching. This led to the following sub-research question:

Sub-RQ3: How do CLIL students perceive their teaching? (Article III)

Combined, the three sub-research questions answer the main RQ. By drawing on both teaching practices and student aspects of CLIL classrooms, the thesis seeks to gain a fuller understanding of how CLIL is taught in a Norwegian educational context. To answer the RQs, a mixed methods approach was used, employing both qualitative and quantitative data (Creswell, 2009). Three CLIL programs were observed and the observations video recorded, and questionnaires were handed out to the students. More information on the methods is presented in Section 4.

1.5 Structure of the thesis

This thesis consists of two parts: an extended abstract (Part I) and three journal articles (Part II). The extended abstract comprises five chapters. Chapter 1 serves to frame the theme of the research (CLIL teaching) and explain some of the educational background and context. Chapter 2 describes the theoretical lens, situating this thesis in sociocultural theory, and draws on SLA and CLIL integration theories to shed light on language use and learning. Chapter 3 provides a state-of-the-art literature review on CLIL classroom research. Chapter 4 outlines the research design and methods used in this thesis. Chapter 5 provides a summary of each article, synthesizes the findings, then discusses research contributions and, finally, possible implications for CLIL in Norway.

2 Theoretical underpinnings

In this thesis, a theory is understood in two ways: "an overall lens" that shapes the research project (Creswell, 2009, p. 249) and an "explanatory system that discusses how a phenomenon operates" (Johnson & Christensen, 2012, p. 596). This chapter provides an overview of the theoretical understandings of learning and language that underpin the thesis and articles. My theoretical lens is a combination of language use and language learning theories, as both are necessary to understand CLIL (Coyle, 2011, p. 55). The epistemological view on learning expressed in this thesis is sociocultural, which entails the belief that how CLIL students learn languages can be explained by historical, cultural, and social contexts (Mercer, 2004, p. 139). Firstly, I will clarify sociocultural theory, some of Vygotsky's main concepts applied in the thesis, and how Vygotskian thought has been interpreted in the dialogue of the CLIL classroom. Then, I will explain some of the main concepts of Second Language Acquisition (SLA) that have been considered in this thesis, namely, BICS/CALP, Form-Focused Instruction (FFI), and the Input/Output Hypotheses. I will relate how these concepts have affected applied linguists' ideas of what constitutes effective language teaching. The last section will focus on theories related to CLIL: in particular, integration theories and disciplinary literacy.

2.1 Sociocultural theory and Vygotsky

2.1.1 An introduction to sociocultural theory

The grand theory of this thesis is sociocultural theory. The research design and tools (PLATO, TRIPOD and my own questionnaire) are influenced by Lev Vygotsky and his seminal work, *Thought and Language* (I use the translated version: Vygotsky, 2012). *Thought and Language* introduced many of his groundbreaking ideas about learning and development that influence the educational sciences today (e.g., Stray & Wittek, 2014). Broadly speaking, Vygotsky believed that learning and development were mediated processes (Daniels, 2001, p. 1). While other psychologists were preoccupied with the cognitive processes of the mind (e.g., Piaget, 1926), Vygotsky emphasized development as a social as well as biological process (Alexander, 2014, p. 11). In Vygotsky's view, the student is a part of "a history, of a culture, and of a society" (Swain,

Kinnear, & Steinman, 2015, p. xi). Education is, in this sense, a dialogic process between teachers and students within cultural institutions (Mercer & Littleton, 2007, p. 4). This has inspired the "participation" metaphor: learning a subject does not mean simply acquiring knowledge but becoming "a member of a certain community" (Sfard, 1998, p. 6). The classroom must be contextualized to understand how students are socialized in the traditions of their subjects. Since sociocultural theory understands learning as interaction between humans that is situated, we need to investigate said interaction to understand how learning takes place. The emphasis on socialization and the dialogic process between teacher and student are the underpinnings of Sub-Research Question 3, which investigates student perspectives. In other words, Vygotsky's view on learning inspired this research design to include student perspectives as well in order to deepen the understanding of the relationship between teaching and learning in CLIL.

Sociocultural theory is useful for this thesis for a number of reasons. CLIL is considered a communicative approach (Richards & Rodgers, 2014, p. 117). Researchers postulate that there is a high amount of teacher-student interaction in CLIL (e.g., Nikula, 2010). Therefore, it seems fruitful to use a theoretical framework that focuses on how language is used. In CLIL, language is seen as a medium rather than an object of study (Coyle et al., 2010, p. 32) Sociocultural theory is compatible with this view on the role of language, as they both view communication as the key to learning. Furthermore, sociocultural theory is often used in CLIL research that takes a process-oriented view on language learning, such as the present thesis (Nikula, Dalton-Puffer, & Llinares, 2013, p. 75). In other words, a sociocultural approach can explain the success or failure of CLIL in terms of its context rather than by the "capability of individual students or the skill of their teachers" (Mercer, 2004, p. 139).

2.1.2 Vygotsky and language

Language occupies a special place in Vygotsky's understanding of learning (Kozulin, 1986, p. xlv). From an ontological viewpoint, Vygotsky believed that the relationship between thought and speech is a continual back and forth movement. Speech is ultimately the transformation of thought into words, and how we express ourselves is the very proof of our thought (Vygotsky,

2012, p. 231). Consequently, the dialogue that arises between the teacher and students is a reflection of their thoughts and acquired knowledge.

Early on, Vygotsky made a distinction between the type of language a child intuitively learns and the language of adults. He differentiated between the two using the terms *spontaneous* versus *scientific concepts*. Vygotsky argues that children develop functional equivalents of scientific phenomena (e.g., the sun is going up) but learn more scientific and accurate ways of expressing them as adults (e.g., sunrise) (Vygotsky, 2012, p. 109). This indicates that there is a socialization process wherein children must learn the norms and accepted terms that adults use.

2.1.3 Scientific concepts and the Zone of Proximal Development (ZPD)

According to Vygotsky, one of the objectives of development is for children to acquire the elusive *scientific concepts*. Teaching is viewed as a "powerful force" to develop students' use and understanding of scientific concepts (Vygotsky, 2012, p. 167). The acquisition of scientific concepts, according to Vygotsky, is carried out in the students' *Zone of Proximal Development (ZPD)*: "a psychological 'space' where students' experientially rich spontaneous concepts meet the teacher's systemically organized academic concepts" (Kozulin, 2012, p. xviii). The discrepancy between the child's mental age and level she reaches with assistance is the ZPD (Vygotsky, 2012, p. 198). The ZPD has played a large role in the educational sciences in recent decades (see Kinginger, 2002). Wood, Bruner, and Ross (1976) later researched how children reached learning goals with and without adult assistance, resulting in the term *scaffolding* (this will be elaborated on in Section 3.1.3). The ZPD is crucial to scaffolding theory because it is the space where scaffolding takes place (Shepard, 2005, p. 66). Scaffolding theory informs Sub-Research Question 2 and builds on Vygotsky's ZPD.

2.1.4 Vygotskian thought in Discourse Analysis

Sociocultural thought is in no way unified. Vygotsky's work is interpreted and applied in a range of different fields in the educational sciences, leading many to argue for sociocultural *theories* (Wittek, 2014, p. 134). Vygotskian thought has been applied to a variety of research in the CLIL

classroom as "it does not make much sense to argue for [CLIL] without a greater understanding of the ways in which languages are actually used in classroom interaction (Llinares, Morton, & Whittaker, 2012, p. 8). Sub-Research Questions 1 and 2 focus on dialogue in the CLIL classroom. To identify what characterizes effective dialogue, these studies are informed by some of the theoretical assumptions in Discourse Analysis. Sociocultural researchers who work with discourse analysis draw on the works of Vygotsky to understand the link between interaction and learning (e.g., Alexander, 2014; Mercer & Littleton, 2007; Wells, 1999). From this perspective, language is viewed as the teacher's main pedagogic tool (Mercer & Littleton, 2007, p. 2). Effective learning can be explained by interaction between teachers and students (Mercer, 2004, p. 139). The quintessential question posed among Vygotskian discourse researchers is what types of interaction can provide the best learning outcomes (Mercer, 2010, p. 2). The present thesis assumes that certain types of interaction between teachers and students will lead to more effective learning.[4]

Discourse analysts draw on a variety of conceptions about language use and learning that go by the name *dialogic teaching*. This refers to the crucial role of discourse in the classroom (Coyle, 2011, p. 52). One of the central concerns of dialogic teaching is the role of teachers. Followers of dialogic teaching move away from transmissionary modes of teaching in which the teacher simply "transfers" knowledge to the students (Wells & Arauz, 2006, p. 379). For effective learning to take place, dialogic teaching posits that students must be actively involved in learning, with the teacher taking on the role of facilitator instead (Ashwin & Boud, 2015, p. 225). This is highly compatible with CLIL and communicative language teaching (Richards & Rodgers, 2014, p. 122). One teacher-student interaction pattern that is prominent in literature about classroom discourse is the Initiation, Response, Feedback (IRF) pattern (Mercer, 2004, p. 7). The IRF pattern is criticized in dialogic teaching as it stifles opportunities for students to voice their own opinions (Wells & Arauz, 2006, p. 380). These criticisms are reflected in some of the basic theoretical tenets of the present thesis and in the tools used (see Section 4.4 for

[4] Discourse analysis is also a methodology, so I must underline that I use the Vygotskian theoretical rationale behind this field, but not the methodology itself (which will be explicated in Chapter 4).

more details). However, many researchers view this understanding of the role of dialogue in the classroom as normative. Dialogue is a rich and varied phenomenon, and a number of studies show there is no "blueprint" for what constitutes effective classroom dialogue (for a full overview, see Howe & Abedin, 2013).

2.2 Second Language Acquisition (SLA) and CLIL

The previous section outlined some of the main tenets of language use that may lead to effective learning. However, as Coyle (2011) underscores, we also need to understand how students acquire language to understand CLIL (p. 55). For this, we turn to SLA, a field of research that characterizes learners' underlying knowledge of the L2 (R. Ellis, 2008, p. 6). SLA identifies how students learn languages; with this knowledge, we can attempt to create optimal conditions for teaching languages. SLA is a broad branch of research that was dominated by cognitive theories for decades; however, in recent years, it has taken a turn toward social theories to explain how learners acquire languages (Swain et al., 2015, p. xi). SLA has a large influence on CLIL as most CLIL research and teaching relies, to some extent, on current understandings of how learners acquire language (Coyle, 2007, p. 548). In the following, we will unpack some of the most central concepts of SLA that are used within CLIL research. The following concepts have motivated the focus of this study (e.g., BICS and CALP), explain why some forms of teacher behavior are rewarded in the PLATO analyses (e.g., the Output Hypothesis rewards when students are permitted to speak more), and clarify the CLIL teacher's role as a teacher of both a content and language.

2.2.1 BICS and CALP

A common theoretical distinction between types of language used in CLIL is *Basic Interpersonal Communication Skills* (BICS) and *Cognitive Academic Language Proficiency* (CALP) (Angel Lin, 2016; Llinares et al., 2012). Cummins (1979) developed this framework to explain why some aspects of language are more difficult to acquire for L2 speakers. BICS is considered everyday fluency in the language, whereas CALP is defined as "the dimension of language proficiency which is strongly related to overall cognitive and academic skills" (Cummins, 1979, p. 198).

Cummins' (1981) study of language proficiency indicated that CALP language takes at least five years for L2 learners to acquire. CALP was later used as a concept in education for educators to identify potential linguistic challenges for second language learners (Cummins, 2013).

One primary reason for the appeal of CLIL is that it provides exposure to CALP language, which is not readily available to students outside of the classroom (Dalton-Puffer, 2007, p. 3). This is because lexis is central to CLIL (Richards & Rodgers, 2014, p. 120). In Norway, CALP language is of interest because it is the type of language that Norwegian students struggle with the most (Hellekjær, 2005). In the CLIL literature, there are various ways of interpreting CALP language. Some researchers use the term *academic language* as a synonym for *CALP* (e.g., Nightingale & Safont, 2019), whereas others reject the idea of a label for a generic type of language use. This is especially apparent in integration theories (see Section 2.3.4), where the term *subject-specific terminology* is used to describe the language of a specific discipline, e.g., the language of math, science, and so forth (Berger, 2016; Lorenzo & Dalton-Puffer, 2016). Since this thesis gravitates toward integration theories, the term *subject-specific terminology* is used to highlight that language differs depending on the subject. Special emphasis is put on CALP language in Articles I and II when determining the type of language to which CLIL students are exposed.

2.2.2 Form-focused instruction (FFI): How much is enough?

One of the core questions in SLA asks if we learn languages implicitly or explicitly. Investigating this question can enlighten us as to how we should teach language in the CLIL classroom. The L1 is largely learned implicitly, but an extensive body of research suggests that this is not sufficient for the L2 (N. Ellis, 2011, p. 45). The question that remains is which aspects of the L2 do CLIL teachers need to teach explicitly for effective L2 learning? The approach that deals with this question is referred to as *form-focused instruction*[5] (henceforth FFI): when teachers specify what the learners will learn and when they will learn it (R. Ellis, 2008, p. 837). Typically, FFI is associated with explicit grammar teaching and correction, but this approach also includes the identification of language learning goals. Conversely, *indirect intervention* aims to create

[5] The term *direct intervention* is also used by some scholars (e.g., R. Ellis, 2008; R. Ellis et al., 2009).

conditions for learners in which they can communicate in the L2, but the language is learned implicitly (R. Ellis, 2008, p. 837). Schmidt (1994) points out that communicative language teaching approaches, such as CLIL, tend to gravitate toward indirect intervention since stakeholders assume that learners will acquire linguistic forms through situational meaning (p. 12). Although SLA researchers generally agree that much of the L2 can be learned through indirect intervention (for instance, fluency), grammatical errors can be fossilized if students are not corrected enough (R. Ellis, 2008, p. 846). For CLIL to be called "Content *and Language Integrated Learning'*, there must be tangible language learning goals expressed in the curriculum (e.g., Dalton-Puffer, 2007, p. 10). In other words, there is tension regarding the degree to which CLIL teaching should be form-focused—or even if at all. This will be further discussed in Section 2.3. The preoccupation with this question in the CLIL literature led me to explore whether there exists any FFI in the CLIL classroom. This is reflected in Sub-Research Question 1, which examines how CLIL teachers approach language (learning) in their respective subjects.

2.2.3 Input and output for language learning

Two theories from SLA that have had a profound effect on how we understand language learning are the Input Hypothesis (Krashen, 1985) and the Output Hypothesis (Swain & Lapkin, 1995). These theories are deeply ingrained in bilingual education as they deal with a number of issues related to how CLIL should be implemented and practiced.

2.2.3.1 The Input Hypothesis

The Input Hypothesis was developed by Stephen Krashen (1985). According to Krashen, language rules are acquired in a natural order. We acquire language through receiving *comprehensible input*, moving from the elementary rules of a language to the more advanced (*i* + 1) (Krashen, 1985). The Hypothesis has particularly influenced direct approaches to language learning as it focuses on exposing students to the L1. The notion of comprehensible input has been vital to bilingual education since much research has shown that if students are taught in a language they do not understand, the bilingual program will inevitably fail (see

Brock-Utne, 2007). In other words, input must be comprehensible for the students to benefit from bilingual education. Cummins (1976) posits that to reap the benefits of bilingual education, a student needs a "threshold level" of linguistic competence in the L2 (p. 3). As a consequence, some bilingual programs require that students be at a certain proficiency level in the L2 before they can enter the program (e.g., Lo, 2015). This is also true of CLIL programs in Norway: three of the schools included in this study required high grades in English (see Section 4.2.1). Moreover, Lin (2016) suggests mapping out the linguistic requisites for each individual CLIL subject to predict what input will be incomprehensible (pp. 77–78).

2.2.3.2 The Output Hypothesis

The Input Hypothesis is not without its critics (for a comprehensive overview, see R. Ellis, 2008, p. 251). One of the most salient critiques is by Swain (1985), who researched French immersion classrooms in Canada in the 1970s and 1980s. Analyzing large-scale data of the students' linguistic outcomes, she remarked that students did not achieve native-like proficiency in French because "the teacher talks and students listen" (Swain, 1985, p. 247). What was missing was opportunities for students to use the language in meaningful contexts (Ortega, 2009, p. 62). This goes hand in hand with dialogic teaching and CLIL's communicative approach. Therefore, CLIL may be seen as different from the transmissionary teaching style of early French immersion.

In later years, Swain and Lapkin (1995) formulated the *Output Hypothesis*: in producing the L2, learners notice a linguistic problem, and this pushes them to modify their output (p. 372). *Noticing* is a key term in the Output Hypothesis; by speaking, students become aware of their own linguistic deficits. This is fundamental to SLA: not only does output give students opportunities to practice the language; it also contributes to language acquisition itself. Ellis (2014) argues that in later years, the Output Hypothesis propelled task-based instruction in the language classroom (p. 39). The Output Hypothesis has affected an array of theoretical frameworks and research articles in CLIL (see Section 3.1.2) but can be boiled down to this: the more students talk, the more opportunities they have to improve their L2. This has affected my

understanding of effective language teaching—teaching where students are encouraged to use the language reflects the Output Hypothesis, which leads to "noticing."

2.3 Content, language, and integration

In addition to drawing from SCT and SLA theories, CLIL has its own theoretical rationale, springing from French immersion in the 1960s and onwards. The following section will describe the historical and contextual development of the terms *content*, *language*, and *integration* as well as frame integration theories and disciplinary literacy. *Content*, *language*, and *integration* are central to understanding CLIL; nevertheless, their definitions are fuzzy and often used implicitly in research contexts (Banegas, 2016; Davison, 2005; Llinares, 2015). Thus, how these terms are used in the CLIL literature will be clarified.

2.3.1 Content in bilingual education

The terms *content* and *language* are separated in bilingual education to create a distinction between the subject matter that is being taught and the target language (e.g., Brinton, Snow, & Wesche, 1989; Mohan, 1986). *Content* is often embodied as whatever language is not, e.g., a "non-language" subject or "non-language" learning goals (e.g., Marsh, 2002, p. 65). The key to defining content is defining it in its context as it can range from "the delivery of elements taken directly from a statutory national curriculum to a project based on topical issues" (Coyle et al., 2010, p. 28). I use the term *content* to refer to subjects and topics where language (learning) is not the main priority. This is an important distinction for Sub-RQ1, which looks at content and language teaching.

2.3.2 The role of language

As can be seen in the introductory quote by van Lier (2004), this thesis takes the stance that language is a pervasive aspect of all education. This belief stems from the *Bullock Report* (1975), which put forth the idea that language should be reflected across the British curriculum in consideration of second language learners. Although *language* is not often defined in CLIL literature, it would appear that it is understood as referring to both language use and language

learning (Coyle et al., 2010, p. 32). One of the most critical issues in CLIL is agreeing on the role language should play in non-language subjects (Banegas, 2016; Davison, 2005; Mohan, Leung, & Slater, 2010).

To clarify the role of language in bilingual education, Met (1999) presents content and language on a continuum.

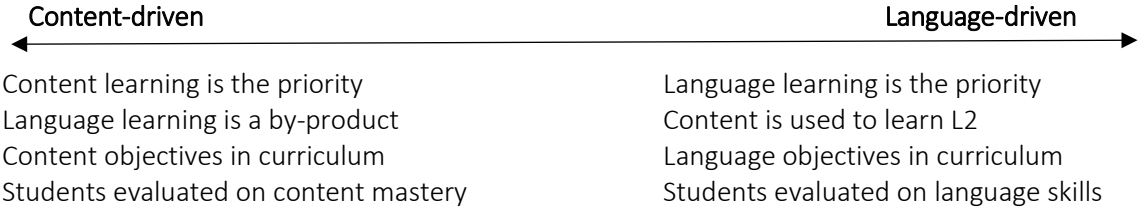


Figure 1. A continuum of Content and Language Integration, adapted from Met (1999)

Language-driven bilingual programs tend to employ FFI since there are concrete language objectives in the curriculum. Content-driven bilingual programs have a more indirect approach in which language learning is a by-product. CLIL has been labeled content-driven since it is situated in content subjects and rarely expresses language learning goals in the curriculum (Banegas, 2016; Dalton-Puffer, 2011; Georgiou, 2012). The Norwegian school system does not have a CLIL curriculum per se either, meaning that whichever language goals are included in the CLIL program already exist in the content subject in which they are taught.

Since CLIL veers toward the content-end of the scale, the general stance is that CLIL classrooms are "environments which provide opportunities for learning [language] through acquisition rather than through explicit teaching" (Dalton-Puffer, 2007, p. 3). Although the idea of approaching language learning in a naturalistic learning environment distinguishes CLIL from traditional language teaching, Dalton-Puffer (2007) warns that such an attitude may produce "a potentially rather passive notion of the language learning process" (p. 3). Researchers operating within sociocultural theory have understood language to play a functional role as the medium of instruction (Banegas, 2016; Coyle, 2008; Coyle et al., 2010; Creese, 2005; Mohan,

1986), and this manifests itself in the form of classroom conversations. This indicates that there is much focus on language use rather than on language learning.

2.3.3 The content/language dichotomy

A number of researchers have problematized the content/language dichotomy in bilingual education. Three major issues arise here: content within language, content within language subjects, and language within content. Adherents of Systemic Functional Linguistics (SFL) claim that language consists of *what we talk about* (content) and *how we talk about it* (expression) (Mohan, 1986, p. 1). In this understanding of language, content exists within language and cannot be separated from it (see Mohan et al., 2010). Dividing content and language moreover implies that the language subjects are somehow content-less. The language subject becomes reduced to "skills in speaking, reading, and writing which are readily transferable to other areas of the curriculum" (Davison, 2005, p. 221), ignoring the unique contributions the language subjects make to our understanding of literature, history, and culture (e.g., The Norwegian Directorate for Education and Training, 2013a). Conversely, all content subjects have implicit language learning goals, such as "formulate," "explain," and so forth (The Norwegian Directorate for Education and Training, 2013b). Claiming that content and language are separate gives rise to a tension in CLIL literature that "renders manifest an issue which is actually fundamental to large sections of formal education" (Dalton-Puffer, 2007, p. 6). Although the content/language divide is clearly problematic in some areas, I have still found it a useful tool to distinguish between some aspects of teaching and learning in Sub-Research Questions 1 and 2.

2.3.4 Integration and disciplinary literacy

Lastly is the term *integration*, which refers to the "fusion of language and content in learning, teaching, and research" (De Graaff, 2016, p. xiii). Dalton-Puffer et al. (2014) argue that CLIL has contributed substantially to theoretical work on integration (Dalton-Puffer et al., 2014, pp. 215–216). CLIL research postulates that integration is more than simply balancing content and language; it entails recognizing and mapping "the inherent role of language in teaching and

learning (De Graaff, 2016, p. xv). Integration is more than the sum of its parts (García, 2009, p. 8), requiring educators to rethink how to handle content and language in the classroom (Wiesemes, 2009, p. 48). This process has prompted integration theories: theories that attempt to map integration for effective teaching practices. Although there are a range of approaches to integration (see Nikula, Dafouz, Moore, & Smit, 2016), I have chosen to narrow my focus toward one integration theory, namely disciplinary literacy.[6]

Disciplinary literacy refers to an emphasis on "the knowledge and abilities possessed by those who create, communicate, and use knowledge within the disciplines" (Shanahan & Shanahan, 2012, p. 8). It rejects language skills as interchangeable and explains how students can read, write, think, and reason in various disciplines (Rainey & Moje, 2012, p. 73). Integration posits that school subjects are the "result of historical processes" and that language is historically, culturally, and socially shaped in the classroom (Nikula, Dalton-Puffer, Llinares, & Lorenzo, 2016, p. 7). Therefore, through disciplinary literacy, teachers must help students to read and write within the framework of a subject (see Shanahan & Shanahan, 2012). This includes knowing key vocabulary items, the etymologies of words, genres, and strategic ways of reading and writing in a subject.

Disciplinary literacy was applied to the thesis after the results of Article 1 came forth. The subjects of science, math, and English scored differently (e.g., science had more visuals). To explain these differences, literature reviews were conducted in different content subjects the need for a deeper understanding of the nature of individual subjects became clear (see Articles I and II). Disciplinary literacy could explain why some of these differences arose. Disciplinary literacy is a lens that nuances and explains some of the findings of this thesis. It is a way of understanding the interplay between content and language in specific contexts, in line with Vygotsky.

[6] A similar CLIL theory is "pluriliteracies" (e.g., Meyer & Coyle, 2017; Meyer, Coyle, Halbach, Schuck, & Ting, 2015). I chose disciplinary literacy because I find it more clearly conceptualized for individual subjects.

2.4 Brief summary

This section summarizes the main theoretical underpinnings of this thesis. I employ a number of language use and learning theories to explain how teachers and students can interact to create an effective space for learning. In turn, these theoretical prerequisites are used in the data analyses (Section 4.3) to determine if CLIL teaching is effective. The sociocultural view employed in this thesis calls for research that is deeply entrenched in contextual details of CLIL teaching and learning. Teaching is considered effective when the teacher creates space for students to express themselves. Content language and integration in CLIL have been conceptualized. Disciplinary literacy has been expressed as a useful tool to explain empirical differences between teaching practices in content subjects.

3 Literature review of empirical research

In this chapter, I situate the thesis in a CLIL teaching context, summarizing research on current understandings of the CLIL classroom. This review supplements and expands on the reviews of the individual articles. It is delimited to research on primary and secondary school, reflecting that there is a difference in practices in tertiary education (Dearden, 2015). The review is also delimited to research on CLIL as it is the most relevant bilingual context for the thesis. Since this is a thematic review (focusing on themes that are relevant to my research), I will be very brief regarding how the review was conducted. I searched in the databases ERIC, Google Scholar, Bibsys (the Norwegian national database), Taylor and Francis, and other major journal databases. Keywords such as *CLIL teaching*, *CLIL pedagogy*, *CLIL teacher*, and so forth were used, as well as keywords related to the respective themes of the literature review (e.g., *CLIL discourse*, *CLIL EFL*, *CLIL L1*, *CLIL questions*, *CLIL scaffolding*, and *CLIL student perspectives*). In addition to searching databases, I also used the "footnote chasing" method of finding articles and books through relevant literature lists (Flamez, Lenz, Balkin, & Smith, 2017, p. 100). Only research conducted in English or Norwegian was selected. The thesis recognizes that CLIL research is a multilingual endeavor, and there is a bulk of research, particularly in German, which could not be addressed. The literature review section is divided into international and national research contexts to reflect that national contexts can differ greatly (e.g., Sylvén, 2013).

3.1 International research on CLIL

3.1.1 A brief overview

CLIL research is a large and active field that is constantly growing (Nikula et al., 2013, p. 72). Research began in Europe in the late 90s and has expanded to Latin America, Oceania, and Asia in recent times (e.g., Turner, 2013). Spain has particularly invested in CLIL policy, research, and practices due to its many bilingual regions and growing need for English language proficiency (Lasagabaster & Ruiz de Zarobe, 2010; Sylvén, 2013). The Netherlands, Finland, and Germany have also built extensive CLIL research environments. This chapter addresses two types of CLIL

research: product-oriented research (research that focuses on learning outcomes) and process-oriented research (research that focuses on interaction and context in the classroom). First, the literature review will address product-oriented research, which is the largest and most comprehensive type of CLIL research. Then it will zoom in on process-oriented research, which is the field of research that this thesis positions itself in.

3.1.1.1 Product-oriented research and language learning outcomes

Research on CLIL is relatively new within bilingual education. Research on bilingual programs in North America has been conducted for approximately 60 years (Cummins, 1998; Nikula & Mård-Miettinen, 2014); this research has identified what currently characterizes successful bilingual programs. CLIL, however, is still struggling to justify itself as a legitimate teaching methodology in Europe (Navés, 2009, p. 36). To this end, CLIL research has, until recently, focused on (language) learning outcomes, or more specifically, how well proficient CLIL students perform in the L2 compared to non-CLIL students (Dalton-Puffer, 2011; Pérez-Cañado, 2012). This type of research has been labeled product-oriented since it focuses on the L2 learning outcomes students have achieved through CLIL (Nikula et al., 2013, p. 72). The main language gains observed in CLIL are in reading, listening, receptive vocabulary, speaking, writing, morphological phenomena, and emotive/affective outcomes (Ruiz de Zarobe, 2015, p. 56). The results remain inconclusive in areas including syntax, productive vocabulary, informal language, writing accuracy, pronunciation, and pragmatics (Ruiz de Zarobe, 2015, pp. 56–57). However, since CLIL students often take language courses in parallel with CLIL teaching, if they score higher in L2 proficiency, the results may not be due to the CLIL methodology itself but the fact that they have more opportunities to use the L2 (Dalton-Puffer, 2011, p. 186).

3.1.1.2 Research on content outcomes in CLIL

A handful of studies have also focused on content outcomes, i.e., how much CLIL students know about their content subject. The results are conflicting: some report that CLIL students score equally well in knowledge of the content subject as do non-CLIL students (e.g., Ruiz de Zarobe, 2015), whereas others report that CLIL students score lower (Fernández-Sanjurjo, Fernández-Costales, & Blanco, 2017; Jäppinen, 2005). Research on testing the content knowledge of CLIL

versus non-CLIL students is also problematic. Students who typically choose CLIL are often from higher socioeconomic backgrounds and academically strong (Bruton, 2011, 2013). This has led to skepticism concerning quantitative methods and the practice of measuring CLIL versus non-CLIL scores (Pérez-Cañado, 2012, p. 331). In order to fully understand the benefits and shortcomings of CLIL, there is a need for more qualitative and mixed methods approaches (Nikula et al., 2013, p. 73). Testing can tell us how much CLIL students have learned (to some extent) but not why CLIL students score at the levels they do.

3.1.1.3 CLIL research is context-dependent

Research on CLIL is complex for many reasons. Due to its interdisciplinary nature, researchers are needed from different fields to examine the many aspects of CLIL. CLIL has generally attracted applied linguists; however, the content side of CLIL has remained neglected (Nikula et al., 2013, p. 86). The overt focus on the language learning aspect of CLIL has created tensions between content and language. CLIL teachers are usually content teachers and given little information on how CLIL "works" in their subject. (Dalton-Puffer, 2007, p. 5). This has resulted in calls for more research by content specialists, with a focal point on teaching/learning the content subject (Georgiou, 2012; Pérez-Cañado, 2012). Results vary greatly, leading to some countries reporting much more success in their CLIL programs than others (Coyle, 2008; Georgiou, 2012). Sylvén (2013) postulates that since CLIL programs vary from country to country, the success of implementation will depend on national characteristics (p. 301). Spain has reported positive results for L2 acquisition (see Lasagabaster & Ruiz de Zarobe, 2010; Ruiz de Zarobe, 2015; Ruiz de Zarobe & Jiménez Catalán, 2009). However, in Sweden, results differ; Swedish CLIL students perform better than non-CLIL students in English L2 at the outset but do not progress more than non-CLIL students (cf. Olsson, 2016; Olsson & Sylvén, 2015, 2019). This underscores a major point of this thesis: context is key. This is compatible with the sociocultural positioning of this thesis (Vygotsky, 2012), which takes into account historical, cultural, and social aspects of teaching and learning. Although research may be conducted on CLIL teaching in other countries, "socio-cultural contexts and forms of implementation often differ" (Nikula,

2005, p. 28). Despite the research on CLIL in other countries, we scarcely know how CLIL is taught in Norway.

This concludes this brief overview of international CLIL research. This section has introduced the main body of research, which is product-oriented research. The next section will zoom in on process-oriented research, which is relevant for this thesis.

3.1.2 CLIL classroom discourse

The studies in the present section tie in with sociocultural theory as they investigate how CLIL teachers and students communicate in the classroom. This is labeled process-oriented research. An important movement within CLIL is the shift from product to process-oriented research, with an emphasis on how context can explain the learning process (Nikula et al., 2013, p. 73). CLIL classroom discourse researchers use naturalistic observation across countries, educational levels, subjects, and programs. The following studies show that data from the observation of classroom discourse is "crucial not only to understanding language and curriculum content learning, but to informing wider discussions on pedagogies and policies" (Coyle, 2007, p. 548). The majority of the studies presented reflect tenets of dialogic teaching, such as creating space for the learner (e.g., Nikula, 2007). Sub-Research Questions 1 (regarding teaching practices in content and language) and 2 (regarding scaffolding) adhere to CLIL classroom discourse. The theme "CLIL classroom discourse" was chosen in the extended abstract because Articles I and II did not have sufficient space for an overview of studies used with similar methods. The following review on classroom discourse will touch upon three major themes related to Articles I and II: differences between CLIL and EFL classroom discourse, types of questions, and the use of L1.

3.1.2.1 Differences between CLIL and EFL classroom discourse

One of the main arguments for CLIL is that CLIL teaching gives students the opportunity to use the L2 in different ways than in the traditional EFL classroom (Coyle et al., 2010, p. 10). A common pursuit in CLIL research is to compare discourse in CLIL and EFL classes to see if there are significant differences in how the teachers and students communicate (see Nikula et al.,

2013 for a comprehensive overview). This includes whole-classroom interaction, teacher-student interaction (e.g., desk talk), and, occasionally, student-student interaction. Results from studies indicate that discourse in CLIL classrooms is relatively more dialogic (see Dalton-Puffer, Nikula, & Smit, 2010; Nikula, 2005, 2010; Nikula et al., 2013). CLIL students use more varied pragmatic functions, use the language more personally, and employ more negotiation strategies when they speak English L2 than students in traditional EFL classes (García Mayo & Lázaro Ibarrola, 2015; Nikula, 2005, 2007). The discourse between CLIL students is relatively more collaborative with less turn-taking and more contributions to other speakers' turns (Moore, 2009). In terms of EFL teaching, researchers have found the lessons to be more teacher-led and to contain less scaffolding of the L1 to help students understand academic terminology (Nightingale & Safont, 2019). English is more institutionalized in EFL classrooms, centering around material and exhibiting more IRF patterns (Nikula, 2007).

Although the research explicated above has indicated that CLIL classroom discourse is more conducive to language learning than EFL classroom discourse, there is no clear consensus when taking into account various learning contexts. In Hong Kong, CLIL classrooms can be teacher-dominated; teachers talk up to 96% of the time during the lessons (Lo, 2015). Students have difficulties speaking English, and turn-taking is relatively short (Lo & Macaro, 2015). In other words, there is a range of challenges for implementing CLIL in countries where English proficiency may not be high enough or in educational contexts that are teacher-centered. Another challenge that few researchers have tackled is comparing discourse between content subjects (e.g., Nikula, 2010). From a disciplinary literacy perspective, the way we communicate depends on the subject. The question is, therefore, not if we communicate differently in content lessons versus in English lessons but whether there is a notable difference between the communication that takes place, for instance, in CLIL science and CLIL math classes or between CLIL and non-CLIL subjects. This is a highly under-researched area in CLIL classroom discourse.

To summarize, there are subtle yet noticeable differences in discourse between CLIL classrooms and EFL classrooms. The opportunities to use English L2 are different, particularly

due to the nature of the content subject. This goes hand in hand with disciplinary literacy (see Section 2.3.4). However, more research must be conducted on the content subjects themselves to explain how and why these differences emerge. This was addressed in Articles I and II, in which I compared the discourse of various CLIL subjects.

3.1.2.2 Types of questions in CLIL classroom discourse

Asking questions is a tool that teachers use to elicit information, thoughts, experiences, and opinions from students (McNeil, 2011, p. 396). Following a *Conversation Analysis* (CA) approach, a number of researchers have created typologies of questions posed in CLIL classrooms (e.g., Dalton-Puffer, 2007; Dalton-Puffer et al., 2010). The majority of these studies follow the output hypothesis (Swain & Lapkin, 1995), believing that if a question elicits longer and more syntactically complex student responses, it is more conducive to language learning (Banse, Palacios, Merritt, & Rimm-Kaufman, 2017; McNeil, 2011).

There are a number of typologies that researchers use to categorize questions. One typology is the distinction between *display questions* (questions to which the answer is known to the teacher) and *referential questions* (questions to which the teacher does not know the answers) (Long & Sato, 1983, p. 268). Studies on CLIL classrooms reveal a slight overweighting of referential questions (Dalton-Puffer, 2007; Nightingale & Safont, 2019). Another typology is *open* and *closed questions* (Barnes, 1969). Closed questions typically elicit yes/no or one-word answers, and open questions do not have a set answer, allowing the speaker to adjust the length of their answer. This typology has led to divergent findings in CLIL; Dalton-Puffer (2007) finds a majority of open questions in secondary schools in Austria, whereas Schuitemaker-King (2012) finds a majority of closed questions in primary schools in the Netherlands. Both studies conclude that CLIL classrooms have relatively few IRF sequences and that the use of open and/or referential questions did not necessarily lead to longer student answers. Circling back to the output hypothesis, it would appear that existing question typologies cannot predict the length and quality of student answers.

Regardless of the typology, the function of most questions posed by the teacher is related to checking comprehension of facts (Dalton-Puffer, 2007; Llinares & Pascual Peña, 2015; Schuitemaker-King, 2012). Llinares and Pascual Peña (2015) document that questions that elicit factual information comprise 65% of questions, whereas questions that demand students justify their answers (11%) and opinions (14%) are much fewer in number. This poses a problem as only the latter bolster argumentative and problem-solving skills. This is a crucial matter at a time when critical skills are gaining more ground in the Norwegian national curriculum (The Norwegian Directorate for Education and Training, 2019a).

Questions are a complex tool to study, and more information is needed to fully understand how teachers can elicit student speech. Existing typologies cannot predict how students will respond. CLIL teachers focus overtly on checking that their students understand material but not on how students can apply the information in other contexts. Articles I and II of this thesis investigated teacher and student questions that encouraged critical thinking by asking for clarification, elaboration, or evidence to see how often this occurred in the Norwegian context.

3.1.2.3 Use of L1

CLIL distinguishes itself from other types of bilingual education by its more flexible and strategic use of the L1 (Lin, 2015, p. 74). It is therefore relevant to note how much, how often, when, and where the L1 is used in the CLIL classroom.[7] This relates to input and output as it reveals how much students are exposed to the target language. CLIL involves mainstream students, meaning that the L1 is most likely the majority language of the country and shared by students and teachers alike (Lasagabaster & Sierra, 2009a, p. 372).

The dominance of monolingual teaching over the past century has branded the use of the L1 in foreign language teaching controversial (Gierlinger, 2015). Too much of the L1 goes against the principles of the input hypothesis (Krashen, 1985), as it means less exposure to the target

[7] Research on use of L1 goes by many names, including "translanguaging," (Nikula & Moore, 2019) "code-switching," (San Isidro & Lasagabaster, 2018) and "language alternation" (Kontio & Sylvén, 2015). This Ph.D. does not adhere to a specific branch of multilingualism, so I chose to use the term "use of L1" to include research from all branches.

language. However, research is beginning to recognize the role that L1 plays in scaffolding learning (cf. Gallagher & Colohan, 2014). The use of classroom discourse and observational methods is relatively new to this field (Gierlinger, 2015; San Isidro & Lasagabaster, 2018), and it has been a fairly under-researched area in CLIL (Pavón Vázquez & Ramos Ordóñez, 2019, p. 36). Therefore, it was of interest to research in Sub-RQs 1 and 2.

When it comes to the amount of the L1 spoken in the classroom, research results vary. The use of the L1 largely depends on the CLIL program; programs can vary from under 50 to 100% use of English L2 (Díaz Pérez et al., 2019). However, even if a CLIL program theoretically uses English L2 100% of the time, in practice, teachers still may speak in the L2 as little as 4% of the time (Lo, 2015, p. 276). In the Netherlands, Finland, and Austria, a large amount of L2 use has been observed (80% and higher) (Dalton-Puffer et al., 2010; Nikula & Moore, 2019; Schuitemaker-King, 2012). Generally, the amount of instruction in English L2 will depend on the teacher and the level of students' proficiency in the L2, but students tend to speak to each other in the L1 when working in groups or pairs (Gené Gil, Juan Garau, Salazar Noguera, Ministerio De Ciencia, & Generalitat De, 2012; Kontio & Sylvén, 2015; Yoxsimer Paulsrud, 2014). Investigating CLIL students longitudinally, San Isidro and Lasagabaster (2018) discovered that CLIL students do not use the L1 less when their L2 proficiency improves. This underscores that using the L1 does not necessarily correlate with lack of proficiency. There are other underlying factors that motivate students to use the L1, such as affective factors.[8]

Studies that focus on the teacher typically group the teacher's use of the L1 into three categories: *regulative* (The L1 is used for classroom, task and behavior management), *instructive* (scaffolded use of the L1 to help students understand content and language), and *affective* (use of the L1 to build a rapport with students) (definitions taken from Gierlinger, 2015; Lo, 2015). Teachers mainly use the L1 to fulfill a pedagogical function (instructive), but occasionally, they use the L1 to explain tasks to save time (regulative) (Lo, 2015). Gierlinger (2015) found that teachers mainly focused on re-explaining content. Students will consequently

[8] Factors that affect the learner's emotional state or attitude (Richards & Rodgers, 2014, p. 266)

get information explained in two different ways. This points to the potential of L1 for clarifying content. Nikula (2015) and Schuitemaker-King (2012) also tied L1 use to clarifying material for students or when teachers feared students did not understand. However, research calls for more planned and strategic use of the L1 by the teacher and not solely spontaneous translations or explanations (Lin, 2015).

The most under-researched topic in L1 use in CLIL is how the students use the L1 in an affective function (Nikula & Moore, 2019, p. 241). Yoxsimer Paulsrud (2014) notes that Swedish students consistently use English L2 for personal interaction, relationship-building, or identity formation (p. 189). She claims that this makes CLIL research in other contexts inapplicable since it does not "correspond to these students who have been exposed to and have learned English both in and outside of the classroom from an early age" (Yoxsimer Paulsrud, 2014, p. 188). Since Norway has a similar English context, it would be interesting to see how the L1 is used in the classroom by both teachers and students. This is a hitherto unexamined aspect of CLIL in Norway, and it is a topic in all three articles of this thesis. Regardless, as Pavón Vázquez and Ramos Ordóñez (2019) put it: "There is a need to research the presence of the L1 in CLIL with more students, with more teachers, in more contexts, and using more pedagogies" (p. 45).

This concludes the section dedicated to research on CLIL classroom discourse. There has been some process-oriented research conducted on CLIL classroom discourse (e.g., Nikula et al., 2013), but many questions remain in regard to how discourse transpires in the content classroom—particularly in the Nordic countries. The research also lacks multimodal approaches that include materials, images, computers, and so forth, since "participants rely on a range of meaning-making resources beyond that of language" (Evnitskaya & Morton, 2011, p. 113). This is why I chose to use a tool that included multimodal features: the Protocol for Language Arts Observation (see Section 4.4.1 for more information). The studies referenced above are grounded in CA and SLA theories, with less focus on the content subject and how discourse may affect the learning process. These choices led to research questions that focused on content (Sub-RQ1) and scaffolding (Sub-RQ2). This thesis examines, first and foremost, how CLIL teachers teach, which will be treated in the next section.

3.1.3 CLIL pedagogy

There is a clear overlap between this section and the previous since many classroom discourse articles concentrate on the teacher. However, the main focus of this section is on *CLIL pedagogy*, defined as "the instructional strategies CLIL teachers use to support students' learning of both content and language in subject lessons taught through English L2" (Van Kampen, Mearns, Meirink, Admiraal, & Berry, 2018, pp. 129–130). This type of research focuses on what the teacher *does*. It is holistic because it employs a variety of methods (questionnaires, interviews, observation protocols) to describe how teachers support the learning of content as well as language. CLIL pedagogy is in line with CLIL's interdisciplinary nature because it draws on tools and theories from other fields, such as SLA.

There is a growing interest in CLIL pedagogy because it is leading CLIL research toward pinpointing what type of CLIL teaching is most effective rather than more self-justifying rhetoric of whether CLIL teaching is beneficial at all (Navés, 2009, p. 36). Researchers argue that we have limited knowledge of the pedagogies employed by CLIL teachers, and it is a key area in need of further investigation (Van Kampen et al., 2018, p. 129). This section will discuss two sub-fields of CLIL pedagogy of direct interest to this thesis. The first is how CLIL teachers integrate content and language (as reflected in Sub-RQ1). The second is how CLIL teachers conduct scaffolding (reflected in Sub-RQ2).

3.1.3.1 Integrating content and language

A central issue in CLIL is the relationship between content and language and how teachers can integrate language learning in a content subject (Nikula, Dalton-Puffer, et al., 2016, p. 4). Since CLIL teachers are largely content teachers who opt to teach through an L2, researchers are beginning to ask how they pay attention to "the inherent role of language in teaching and learning" (De Graaff, 2016, p. xv). This is important to CLIL because it reflects how and if CLIL teachers manage to teach through the L2 and whether this affects learning of the content subject itself.

Van Kampen, Admiraal, and Berry (2016) found that teachers in the Netherlands report a higher focus on language, scaffolding, and input in their CLIL classes than in their non-CLIL classes. Their findings showed that students had a harder time following lectures in English, so the traditional lecture format had to be replaced by more student-centered activities. This corroborates the findings from CLIL classroom discourse (Nightingale & Safont, 2019; Nikula, 2007). Evnitskaya and Morton (2011) argue to the contrary, asserting that "CLIL classrooms are just like classrooms in which subjects are taught in the L1" (p. 124), indicating that outwardly, the differences between CLIL and non-CLIL may be quite subtle.

CLIL teachers use a range of strategies to support language learning, such as visual aids, authentic materials, vocabulary-building strategies, think-pair-share, and so forth (De Graaff, Koopman, Anikina, & Westhoff, 2007; Tavares, 2015). Although CLIL is supposedly characterized by having explicit language learning goals, few researchers have found evidence of FFI, e.g., teachers who explicitly focus on form or clearly communicate language learning goals (Dalton-Puffer, 2007; Grandinetti, Langellotti, & Ting, 2013; Schuitemaker-King, 2012; Van Kampen et al., 2016; Van Kampen et al., 2018).

CLIL literature is also shifting more toward integration theories, such as disciplinary literacy (see Section 2.3.4). A comprehensive literature review by Van Kampen et al. (2018) suggests that science is the subject that is most researched and evolved in this area, possibly due to the long tradition of scientific literacy. However, a questionnaire conducted by Van Kampen et al. (2016) found that disciplinary literacy was one of the aspects least focused on by the teachers. In other words, there is a knowledge gap for how language is shaped in the individual subjects and how teachers can create competent speakers of science, mathematics, social science, and so on. Summarizing the state of integration of content and language, the role of language still remains unclear in CLIL. Much more work is needed to discuss how CLIL teachers can tailor language needs for the individual subject and how to shape curriculum planning for language learning goals. Further empirical research should also focus more on differences rooted in the nature of the content subjects themselves, which is the focal point of Sub-RQs 1 and 2.

3.1.3.2 Scaffolding learning

There is a general consensus that students in second language learning contexts need extra support in the classroom as they have larger language deficits in the medium of instruction (Cummins & Early, 2015). In this thesis, the support that teachers provide to help students in their learning process is referred to as *scaffolding*—"a type of teacher assistance that helps students learn new skills, concepts, or levels of understanding that leads to the student successfully completing a task" (Maybin, Mercer, & Stierer, 1992, p. 188). Scaffolding takes place in the students' ZPD (Section 2.1.3). It is crucial in second language learning contexts as the success of a bilingual program can hinge on the teacher's ability to scaffold learning for second language learners (Navés, 2009, pp. 32–34).

Scaffolding is a central concept in bilingual education in North America (cf. Gibbons, 2015; Walqui, 2006). However, CLIL has only just started developing an interest in this field (see Ruiz de Zarobe & Zenotz, 2017). A number of CLIL articles and books have loosely used the term *scaffolding* without defining it (cf. Llinares et al., 2012; Nikula, Dafouz, et al., 2016). Other articles address aspects of scaffolding without referring to it as such (Dalton-Puffer, 2007; De Graaff et al., 2007; Schuitemaker-King, 2012). A space for research scaffolding needs to be carved out in CLIL literature, with an overt focus on how teachers can support learning through strategies that are directly related to content and language learning. The first step is reaching a consensus on which strategies can support learning in CLIL and how such learning can be empirically measured. This has perhaps been one of the largest contributions of this thesis: the synthesis of the second language learning research on scaffolding to explain how teachers can accommodate learning for second language learners, discussion of how to study it empirically, and its application to a CLIL context.

This concludes the section on CLIL pedagogy. It is a much smaller field of research than CLIL classroom discourse, with distinct gaps that indicate room for more research. The tools for investigating CLIL pedagogy have hailed from SLA (e.g., the observation protocols, "The Sheltered Instruction Observation Protocol") (Echevarría, Vogt, & Short, 2017) and the "penta

pie" (De Graaff et al., 2007). More research should therefore be done that includes tools that focus on aspects of content teaching (e.g., PLATO).

3.1.4 Student perspectives

The last sub-research question (3) addresses student perspectives, which represent another small branch of CLIL research. Most studies on student perspectives are cross-sectional questionnaires (e.g., Dalton-Puffer, Hüttner, Schindelegger, & Smit, 2009; Otwinowska & Forys, 2017), leaving researchers desiring more qualitative and longitudinal approaches (Lasagabaster & Sierra, 2009b; Roiha, 2019; Roiha & Sommier, 2018). The main areas of interest are general motivation (Arribas, 2015; Seikkula-Leino, 2007; Somers & Llinares, 2018) and beliefs/attitudes toward learning the L2 (De Smet, Mettewie, Galand, Hiligsmann, & Van Mensel, 2018; Sylvén, 2015). Criticizing the one-sided focus on language, Somers and Llinares (2018) assert that there is a growing need for research on student perspectives on the teaching itself (p. 13). This thesis has directed attention to this area by investigating how students perceive their CLIL science teaching, using qualitative as well as quantitative methods (Sub-RQ3).

Despite the modest amount of research into CLIL student perspectives, some conclusions can be drawn. A large percentage of students are satisfied with CLIL teaching (Pladevall-Ballester, 2013; Somers & Llinares, 2018; Yoxsimer Paulsrud, 2014). CLIL students also show more positive attitudes toward English than non-CLIL students (Lasagabaster & Sierra, 2009b), although this is likely because CLIL students are more motivated to learn English from the outset (Mearns, De Graaff, & Coyle, 2017). However, a small group of students persistently do not show high motivation in CLIL classes (e.g., Coyle, 2013). Twenty percent of students in CLIL classes in Poland reported high cognitive overload, with a direct link between achievement factors and level of cognitive struggle (Otwinowska & Forys, 2017, p. 473). Similarly, in a Finnish primary school, students had more difficulty acquiring theoretical content knowledge, showing that cognitive overload can hinder both motivation and learning of content knowledge (Jäppinen, 2005). To avoid cognitive overload, a more viable option for less proficient students might be CLIL teaching in vocational subjects that are grounded in more practical uses (e.g., Kontio & Sylvén, 2015).

A prevailing theme in research into student perspectives is *why* students choose CLIL programs (e.g., Mearns et al., 2017). Paulsrud (2019a) investigated this question in an upper secondary Swedish school through semi-structured interviews. The Swedish students were more attracted to the CLIL program due to its reputation rather than the fact that it was in English (Paulsrud, 2019a, p. 286). Learning through English was considered a "bonus," but students expressed the opinion that CLIL was a chance to improve their English in the course of their studies rather than focus on learning it per se. This highlights that Swedish students already view themselves as competent speakers of English and functionally bilingual (Paulsrud, 2019a, p. 293). However, after attending the course, many Swedish students found it to be more difficult than they initially thought and were unsure of how much English they were supposed to speak.

There is a call for more qualitative approaches that give students a voice and more research that focuses on student attitudes toward CLIL in retrospect (Roiha, 2019; Somers & Llinares, 2018). CLIL students have rarely been asked to evaluate their teaching as a whole; there is more stress on attitudes toward language learning. The way CLIL is taught in traditional content subjects is effective and enjoyable for some students, but not all. A question that CLIL research could further pursue is how to reach the students who feel there is a cognitive overload when learning difficult content matter through a second/foreign language.

3.2 Norwegian research on CLIL

The widespread nature of CLIL inhibits our ability to compile a substantial, cohesive body of research (Georgiou, 2012, p. 498). Factors that limit the application of CLIL research in other contexts include organizational and affective factors (Turner, 2013, p. 395), policy frameworks, teacher education, age of implementation, and exposure to the L2 outside of the classroom (Sylvén, 2013, p. 301). Sylvén (2013) found that CLIL factors vary widely between countries, advising the utmost caution when comparing cross-country results (p. 316). To this end, the present thesis stresses the importance of context; accordingly, we will now circumscribe this research within the Norwegian context.

Currently, CLIL is a small area of research in Norway. Glenn Ole Hellekjær is often thought of as the "father" of CLIL in Norway. He was the first to teach and research CLIL in the Norwegian context, starting in 1993 (Hellekjær, 1996; Svenhard et al., 2007). His doctoral thesis compared upper secondary CLIL and non-CLIL students, concluding that the CLIL students attained higher reading proficiency (Hellekjær, 2005). Interest in CLIL grew after this; the Norwegian National Center for Foreign Languages in Education later conducted a CLIL project in upper secondary schools (from 2009–2011), producing reports with experiences from the teachers and school leaders (Fokus på språk, 2010, 2012; Paulsen, 2010a, 2010b). A small number of articles were written in and as a result of the reports, focusing on the students' language outcomes. The studies concluded that Norwegian CLIL students of English score higher in reading proficiency (Brevik & Moe, 2012; Drew, 2013; Hellekjær & Hopfenbeck, 2012) and listening proficiency (Brevik & Moe, 2012; Moe, 2010) than non-CLIL students.

One Norwegian Ph.D. thesis has been entirely dedicated to CLIL students. Lialikhova (2019) investigated the oral competence of lower secondary history students involved in a six-week CLIL project. The students who were already highly proficient in English developed a higher competence in oral English, whereas the less proficient students did not develop significantly (Lialikhova, 2018a). A more qualitative study of the same group revealed that there was a reluctance among students to speak in English, and this was one of the main challenges for the teacher (Lialikhova, 2018b). Students spoke in English in low-risk situations (e.g., in small discussion groups or when they were not graded), but low-achieving students showed less of a willingness to speak.

There is only a handful of studies on CLIL in Norway, as opposed to our Nordic neighbors, Sweden and Finland (Sylvén, 2013, 2019b). Norwegian research has mainly been product-oriented, concluding that Norwegians CLIL students are more proficient than non-CLIL students. Yet, as Olsson and Sylvén (2015) underscore, these results can be explained by CLIL students' extra willingness and interest in English or their socioeconomic background. As it would seem natural to me that more hours in English should lead to higher English proficiency, more potent questions to pose would be what happens in these classrooms, if CLIL students

are more competent in the language and genres of the specific subject, and if they feel more equipped to use English L2 in work-related situations at a later stage in life. There is much potential for research on CLIL in Norway, and this thesis hopes to contribute to understanding the process of CLIL teaching in Norway.

3.3 Brief summary

As elaborated above, CLIL classroom research ranges from larger to smaller subfields. There are different practices and contexts. Some subfields (e.g., CLIL student perspectives) have produced more consensus in their findings, and others less (e.g., the use of L1). A variety of methods have been used to research CLIL, both qualitatively and quantitatively. From a methodological point of view, there is a call for more mixed methods (Nikula et al., 2013, p. 73) and the use of validated instruments (Mearns et al., 2017, p. 11) to enrich CLIL research. The literature review shows a clear gap in interdisciplinary research among researchers from the respective content subject fields and in the use of tools, theories, and approaches from fields of research other than SLA. Integration and scaffolding research shows that there is a need to map more theoretical understandings of these terms to ensure that researchers are investigating the same phenomena. Research shows how important context is as it provides a solid foundation for a sociocultural perspective. Due to the lack of research on CLIL in Norway, it will be of interest to see how CLIL is practiced in the Norwegian context.

4 Methods and Research Design

This chapter describes my research methods and design. The primary data consist of 24 hours of video data and two questionnaires. The investigation encompassed three CLIL classrooms and seven teachers. In this chapter, I explain the choice of methods, sample, data collection process, and instruments that led to the three articles of my thesis. This chapter is meant to provide a holistic view of this Ph.D. project, justify the relationship between the three articles, and illustrate how they build on and complement one another. Lastly, I discuss issues of validity, reliability, ethics, and the limitations of the project.

4.1 Mixed methods research approach

CLIL is a complex and contextually situated phenomenon. For this reason, I found it necessary to use more than one method and data source to answer the research questions. My Ph.D. project uses a mixed methods research (henceforth MMR) approach. MMR is a type of research that employs both qualitative and quantitative data (Creswell, 2009, 2016) or draws from a plurality of approaches/concepts that are traditionally qualitative or quantitative (Johnson & Christensen, 2012, p. 50). The fundamental principle of MMR is that collecting, integrating, and analyzing several types of data in more than one way will provide a broader picture of a phenomenon and minimize the weaknesses of any individual approach (Creswell, 2009; Johnson & Christensen, 2012). The need for mixed methods research in CLIL has also been an important factor in the choice of method as CLIL research is typically either quantitative or qualitative (Nikula et al., 2013, p. 73). The sociocultural lens led my research design to be more qualitative, including spending longer periods at CLIL schools, interviewing school leaders and teachers, and collaborating with researchers of different subjects (e.g., science and mathematics) to understand the context of teaching and learning in various CLIL subjects.

True to the MMR approach, I have not only employed qualitative and quantitative approaches in the research design and data collection; I have also integrated quantitative (student surveys) and qualitative (video data, student surveys) data. The integration of these data takes place in three data analyses: (1) quantifying the video data through the PLATO manual, (2) juxtaposing

the student survey data (qualitative and quantitative) with the video data (qualitative and quantitative), and (3) examining the CLIL classrooms within and across the three separate phases. The integration across the phases is illustrated throughout the extended abstract.

4.1.1 The multiphase mixed methods design

The research approach of this Ph.D. project falls under the category of *multiphase mixed methods*, which is a research design that combines quantitative and qualitative data in several phases with a common objective (Creswell, 2009, pp. 16, 221). There are three phases that each represent one major theme of my project. The combination of the three phases answers my main RQ. As Brevik (2015) points out, the time frame allows for influence between the phases (p. 34). Since the data were collected sequentially, data from one phase were used to improve the research design in the next phase(s). This design approach is *emergent* (Creswell, 2009, p. 186), allowing the context and needs of the phenomenon to steer the project's focus. In the following, I will explain each phase, why the phase was chosen, what it consisted of, and how it led to the next.

Figure 2 is a visualization of the Ph.D. project divided by its phases, themes, and the data.

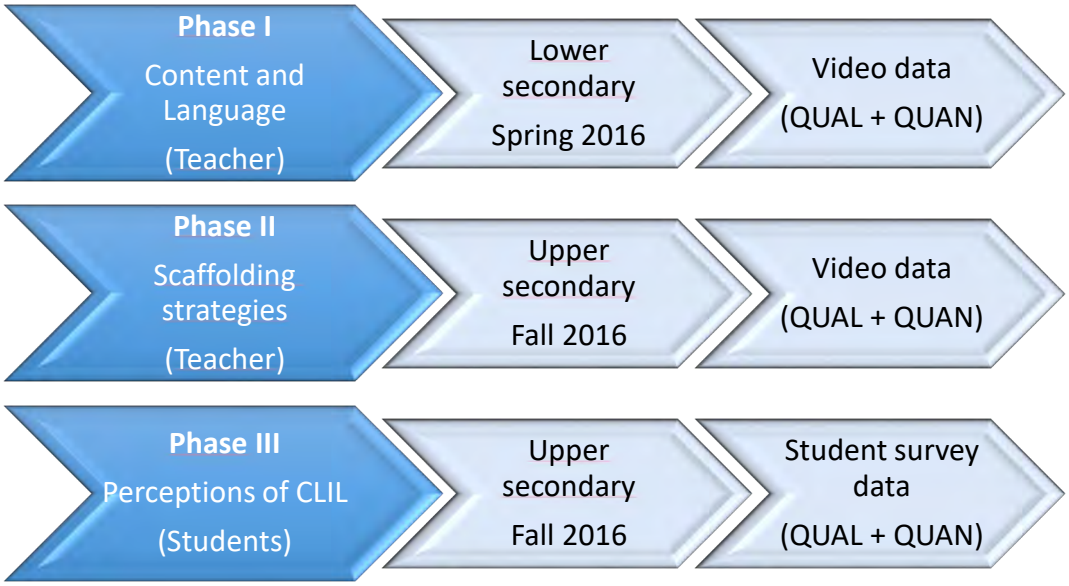


Figure 2. Multiphase mixed methods design

4.1.1.1 Phase I (Article I)

This Ph.D. project was designed to investigate how CLIL teachers in Norway teach their subject through English L2 in secondary schools. Classroom observation is a useful tool in educational research as it gleans empirical evidence of teaching practices. When I decided to investigate CLIL, I saw it as vital to be in the classrooms and observe what happened as part of a bottom-up approach (Creswell, 2009; Johnson & Christensen, 2012), which could help in describing the social context of learning as well (Mercer, 2004, p. 137). Classroom observation could determine the extent to which certain behaviors, attitudes, and teaching methods were present, generating higher reliability than teachers self-reporting their teaching practices (Ary, Walker, & Jacobs, 2014; Gall, Gall, & Borg, 2007). Video technology is becoming increasingly popular in classroom observation because the technology provides a powerful way of collecting, studying, and presenting detailed cases of practice to support teaching and learning (Derry et al., 2010, p. 4). Therefore, I used video data as my primary data source.

In the first phase, I collaborated with the LISE (Linking Instruction and Student Experiences) project at the University of Oslo since my advisor was the coordinator of the team. The LISE team contacted a CLIL program at a lower secondary school and asked if they could film CLIL teaching. This phase comprised two parts: a) filming four hours of teaching in the school subjects of English, science, and math (video data) and b) coding the video data. The data collected during Phase I were the primary data for Article 1, which were used to address Sub-RQ1 (Mahan et al., 2018).

One emerging theme in CLIL is integration and how language can be mapped and taught in different subjects (Sections 2.3.4 and 3.1.3). Phase I accordingly addressed this issue. The article was co-written with a professor of science education (Marianne Ødegaard) to gain more insight into how the science subject was taught and to reflect the interdisciplinary nature of CLIL. Phase I functioned as a pilot that was further developed in Phases II and III. Observing the teaching methods in the CLIL classes led to my interest in scaffolding as I observed variations in how the CLIL teachers scaffolded learning at the lower secondary school (leading to Phase II). The observations also emphasized for me the importance of student input since students were

active in the class. This led to my interest in how students perceive CLIL teaching (Phase III). The first phase resulted in Article I (Mahan et al., 2018).

4.1.1.2 Phase II (Article II)

Building on the findings from Article 1, I sought to understand not only how teachers supported language learning but also how teachers supported learning in general in the content subject. To answer this question, I turned to scaffolding (3.1.3). Since the lower secondary school in Phase I was the only lower secondary school to offer CLIL programs in Norway, I decided to collect data at the upper secondary level. CLIL schools in Norway are mostly in the upper secondary level (The Norwegian National Center for Foreign Languages in Education, 2011b), and collecting at different schools would add to the robustness of data. It would additionally be interesting to compare the teaching practices of the lower and upper secondary schools. In this phase, I used video equipment from the Teaching Learning Video Lab (TLVlab) at the University of Oslo to collect my own video data. I filmed three upper secondary CLIL schools for a total of 24 lessons. Using the coding manual PLATO, I analyzed the scaffolding strategies used by three CLIL teachers to support learning in their subjects. This phase resulted in Article II (Mahan, 2020).

4.1.1.3 Phase III (Article III)

Since the two previous phases emphasized teaching, I wished to include the student perspective. CLIL programs in Norway are largely voluntary; therefore, they must be considered useful by students to exist and prosper. Student perspectives have been used as a tool to improve teaching as they have been recognized as reliable and predictive of learning (Wallace, Kelcey, & Ruzek, 2016, p. 1836). To link Phase III to the previous phases, I chose to elicit information from the students regarding their perspectives on content, language, and scaffolding (Tripod Education Partners, 2015). Therefore, Phase III tested some of the observations from Phases I and II to see if student perceptions equated with the observations (e.g, instructional explanations). Additionally, some open-ended questions were added to Questionnaire I to see if there were any aspects of teaching that the video observation did not

capture (e.g, why students chose CLIL programs). The inclusion of the student perspective coincides with the sociocultural viewpoint of this thesis, as it includes more emphasis on the socialization process.

Questionnaires were the chosen method as they can be used to obtain "thoughts, feelings, attitudes, beliefs, values, perceptions . . . of research participants" (Johnson & Christensen, 2012, pp. 162–163). Additionally, there were a large number of CLIL students, so a questionnaire was a practical way of gathering data from a large set of participants. I distributed two questionnaires to the CLIL students. The first questionnaire I compiled myself. It elicited several answers about the students' background and opinions of CLIL (see Section 4.4.2). The second questionnaire was a translated version of the Tripod questionnaire and elicited responses about CLIL science teaching (Klette, Blikstad-Balas, & Roe, 2017). The final phase resulted in Article III (Mahan & Norheim, under review).

Table 1 provides an overview of the studies related to research questions, data, and articles. The next sections will describe Table 1 in detail, explaining the participants, data, analyses, and research credibility.

Table 1. Research questions, data, data analysis, and articles across the study

Phase	I Content & Language	II Scaffolding	III Student perceptions
Title of Article	Characterizing CLIL Teaching: New Insights from a Lower Secondary Classroom	The Comprehending Teacher: Scaffolding in Content and Language Integrated Learning (CLIL)	Something New and Different: Student Perceptions of Content and Language Integrated Learning (CLIL)
Methods	Mixed	Mixed	Mixed
Research question(s)	RQ1: What characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language?	RQ2: Which scaffolding strategies do three CLIL teachers use to help their L2 English students comprehend material and complete tasks?	RQ3[9]: Why do students choose CLIL programs? RQ4: What pros and cons do students experience in CLIL teaching? RQ5: How do students assess CLIL science teaching?
Goal of phase	To determine the characteristics of CLIL teaching in terms of content and language	To investigate the different scaffolding strategies CLIL teachers use to help their students in the learning process	To determine how students perceive their CLIL teaching
Data	*Video data*: 4 consecutive lessons filmed in science, mathematics, and English class (N=12)	*Video data*: 4 consecutive lessons filmed in science, geography, and social science (N=12)	*Student survey data*: Qualitative questionnaire (N=43) Quantitative questionnaire (Tripod) (N=42)

[9] Summarized as "How do CLIL students perceive their teaching" in the thesis, but three RQs in Article III.

Data analysis	*Video data:* Used the PLATO manual to code the video data.	*Video data:* Used the PLATO manual to code the video data.	*Student survey data*: Questionnaire 1: Content analysis Questionnaire 2: Used SPSS to analyze data.
Research credibility	*Reliability*: Two researchers double-coded 25% of data (inter-rater reliability)	*Reliability*: Two researchers double-coded 25% of data (inter-rater reliability)	*Reliability:* Independent sample *t*-tests were conducted for statistical significance

4.2 Sites and participants

This section will describe the different sites and participants of the Ph.D. project, relating them to the three articles. Since CLIL is contextually contingent (cf. Sylvén, 2013), I found it necessary to look for a range of participants. I chose a multilevel sample of both teachers *and* students to obtain multiple perspectives (Johnson & Christensen, 2012, p. 238). In line with the qualitative aspect of the research design (Creswell, 2009, p. 185), I also found it important to highlight the sites where the data collection took place to contextualize the various CLIL programs.

4.2.1 Sites

Three schools in three different counties in Norway participated in the Ph.D. project.[10] The primary sampling strategy was *convenience sampling*, in which "people who are available or volunteer or can be easily recruited and are willing to participate in the research study" (Johnson & Christensen, 2012, p. 230). The reason for this was the small number of CLIL programs in Norway (Svenhard et al., 2007). The most important aspect was, therefore, finding enough CLIL schools to recruit. Regarding the choice of the type of school, I chose *quota*

[10] A fourth school (upper secondary, vocational) also participated in the Ph.D. project but was ultimately not included due to an overload of data for the researcher and because the Protocol for Language Arts Teaching Observation (PLATO) instrument was not considered appropriate for project-based teaching (see Section 4.4.1).

sampling, recruiting different types of schools. Quota sampling involves identifying the major groups of interest (Johnson & Christensen, 2012, pp. 230–231). By choosing lower and upper secondary schools, as well as different types of immersion (single subjects, partial immersion, project-based), different subjects (natural and social sciences), and programs (general studies and lower secondary), the thesis represents some of the major groups of interest in CLIL programs in Norway. The participants comprised the teacher and students recruited for the Ph.D. project. Each school had one CLIL program that was recruited. Each class had a different teacher for each subject. Many students from all the CLIL classes had international backgrounds; either they had studied abroad or were first-generation immigrants. A range of languages were therefore represented: Chinese, Russian, Arabic, and so forth. All of the teachers were Norwegian except the science teacher in School 1 (Article I). Table 2 provides an overview of the schools and participants that were recruited for the Ph.D. project.

Table 2. Overview of schools and participants

School[11]	1	2	3
County	A	B	C
Type of school	Lower secondary	Upper secondary	Upper secondary
Grade	9th	11th	11th
Program	N/A	General studies program	General studies program
Type of CLIL	Partial immersion (international class)	Partial immersion (international class)	Single subject (Pre-IB)
CLIL subjects	Science, mathematics, social science, religion	Science, geography, social science	Science
Selection process	Reading proficiency test and interview	Oral and written grade in English	Oral and written grade in English
Teachers recruited	Science, mathematics, English	Science, geography, social science, English	Science

[11] The schools and counties are anonymized so the participants will not be identifiable. Schools are represented by numbers and counties by letters to avoid confusion

Students recruited (No.)	26	25	25

School 1 is a lower secondary school in county A. This CLIL class has been operating since 2009. The students are accepted into the CLIL program starting in 8th grade and continue until they graduate the 10th grade. It is labeled as an international program and has a competitive and comprehensive selection process; students must first take a reading proficiency test and then an interview with a teacher to prove their oral skills in English. This program is perhaps the closest one can come to full immersion CLIL in Norway. All subjects besides Norwegian, foreign languages, and subjects with other classes (P.E., music, swimming, and so forth) are taught in English. This site was chosen because CLIL is rarely practiced in lower secondary schools in Norway (The Norwegian National Center for Foreign Languages in Education, 2011a). Therefore, this school was of particular interest to investigate whether CLIL is suitable for students as young as 14 years old. The participants of this program (teachers and students) were recruited by the LISE project. All the teachers (science, mathematics, and English) and students (26 students) expressed willingness to participate. These participants were the basis for Article I.

School 2 is an upper secondary school in County B that has offered CLIL programs since 2014. Its structure and motivation are similar to School 2; the CLIL program is marketed as an international program that draws on the intercultural and international aspects of CLIL discussed in much of the CLIL literature (e.g., Coyle et al., 2010). This school was chosen because it offers many subjects in English (partial immersion) and has CLIL programs from Grades 11 to 13 (all of upper secondary). Additionally, studying upper secondary would allow for comparison between the teaching practices in lower and upper secondary schools. All teachers (science, geography, social science, and English) and students (N=25) expressed willingness to participate. The participants of this class were the basis for Articles II and III.

School 3 is an upper secondary school in County C. This school has offered one CLIL program from the early 1990s. The school was chosen because it offered a standard Norwegian CLIL program—an 11th grade, single-subject CLIL class in a general studies program. Most CLIL

programs in Norway are single-subject programs in Grade 11 that are tied to a school that offers the International Baccalaureate (IB). The CLIL program lasts for a year, labeled as "Pre-IB," and is meant to prepare students for the IB program in Grades 12–13. At School 3, the selection process is based on the students' grades in English from lower secondary school (10th grade), and students from international backgrounds are prioritized. The CLIL science teacher and 25 of 28 students expressed willingness to participate in the study. The participants of this class were the basis for Article III.

4.3 Data collection, instruments, and analysis

Multiple data sources are used in this Ph.D. project. They are qualitative and quantitative and integrated within and across phases, schools, and classrooms. This section will explain the data collection procedure, the data, and how they relate to the three phases of the research design.

4.3.1 Data collection

4.3.1.1 Recruitment

Three schools were recruited for my Ph.D. project in the spring and fall of 2016. All the schools their willingness and ability to host the project. The first school was recruited by the LISE project. The University of Oslo contacted the leadership and asked for permission to film on the school grounds. After this, various teachers in the CLIL program agreed to participate in the project. The students were informed orally and in writing of the project and given a consent form. Since the students were under 15, a guardian had to sign the consent form as well, following the national research ethics guidelines (from NESH, 2006). A similar procedure was conducted for Schools 2 and 3, which I recruited in collaboration with an advisor.

4.3.1.2 Filming (Phases I and II)

After the schools were contacted, I agreed with the teachers on when and how to collect data from their classrooms. The first phase was filming, which was conducted from May 2016–November 2016, approximately two weeks at each school. I filmed four consecutive hours in each subject, abiding by the LISA research design (Klette et al., 2017). Two small, discreet cameras were placed on either side of the room: one facing the teacher and the other facing the students. The teacher wore a microphone, and a microphone was placed in the middle of the classroom to capture student interactions. This design was teacher-centered, as it mostly captured what the teacher said. As a researcher, I was a *participant-as-observer* (Gold, 1958, pp. 220–221). The teacher and students were aware that they were part of a research project, but I had minimal contact with them in the classroom, sitting in the back, silently taking notes and observing. This was to counteract the *observer effect*—how the class may act differently due to a researcher's presence (Ary et al., 2014, p. 219). I wished to observe naturalistic teaching: i.e., how a CLIL classroom would normally take place. This approach allowed me to maintain objectivity and neutrality, but it was also more difficult to put myself in the classroom situation, considering that I was not highly involved with the teachers and students (Johnson & Christensen, 2012, p. 209).

In addition to filming, I collected handout sheets, PowerPoint presentations, and took pictures of the blackboard. This was to understand better the context surrounding what the students were doing (see Evnitskaya & Morton, 2011). Students who did not wish to be part of the project were strategically placed in the cameras' blind spots and their utterances not transcribed. I conducted informal interviews with each teacher to understand the context of their teaching. During these interviews, they provided information about their background and understanding of CLIL, their teaching style, and their view of language learning. These interviews were background information that supplemented the data in Articles I and II.

4.3.1.3 Distributing questionnaires (Phase III)

During the filming period, I also distributed two questionnaires to the students (see Section 4.4.2). The first questionnaire related to the students' linguistic background and included open-ended questions in response to which they could write their opinions on their CLIL program. I constructed the questionnaire myself, and it was available online. All of the students had school laptops and were instructed to access a link to complete the questionnaire online. The second questionnaire was a version of the Tripod questionnaire translated into Norwegian by the LISA team at the University of Oslo (Klette et al., 2017). It was designed to gather information on science teaching and was distributed to the students around their last filmed lesson. The Tripod questionnaire was on paper, and the students filled in their responses on a scale from 1–5 (for more information on the questionnaires, see Section 4.4.2).

4.3.2 Data

My three primary sources of data were the video data and the two questionnaires. The following table provides an overview of the amount of data and how the data relate to the articles.

Table 3. Overview of data

	Video data	Questionnaire 1 (Mine)	Questionnaire 2 (Tripod)	Use in articles
School 1[12]	12 lessons (science, mathematics, English)	Not distributed	English (N= 25) Science (N=26)	Video data used in Article 1
School 2	16 lessons (science, geography, social science, English)	N=18	Science (N=20) English (N=19)	Video data used in Article 2 (excluding English lessons); Questionnaire data used in Article 3 (excluding English)
School 3	4 lessons (science)	N=25	Science (N=22)	Questionnaire data used in Article 3

[12] All data from School 1 belong to the LISE project and may be used in other research projects.

4.4 Instruments and data analysis

This section will provide information on the three main instruments used in this study. I will present background information about each instrument, how it relates to the project, its theoretical tenets, data analysis procedures, and discuss its strengths and weaknesses. Two of the instruments (PLATO and Tripod) follow the LISA design (cf. Klette et al., 2017), and the third was a questionnaire I devised myself. More information about the data analysis processes can be found in the individual articles.

4.4.1 Protocol for Language Arts Teaching Observation (PLATO) (Phases I–II)

The main instrument was the coding manual, PLATO. It was used as the primary instrument of analysis in Articles I and II. PLATO is a coding manual that assesses 12 dimensions (also called codes) of teaching in English Language Arts classrooms: Purpose, Intellectual Challenge, Representation of Content, Classroom Discourse, Text-Based Instruction, Feedback, Strategy Use and Instruction, Modeling and Use of Models, Connections to Prior Academic Knowledge, Accommodations for Language Learning, Behavior Management, and Time Management.[13] Raters score these dimensions of teaching on a score from 1–4, from no evidence (a score of 1) to strong and consistent evidence (a score of 4) (Grossman, 2015). PLATO was developed by Pamela Grossman and her team at Stanford University, building on research on effective teaching practices in secondary school in the United States of America (Grossman, Loeb, Cohen, & Wyckoff, 2013). PLATO was originally used in the Measures of Effective Teaching (MET) project to link teaching practices to student achievement (Grossman, Cohen, & Brown, 2014, p. 305). The LISA and LISE teams at the University of Oslo piloted PLATO in the Norwegian context and applied it to a variety of subjects (Klette et al., 2017). Although PLATO claims to adhere to "best measure" practices and not any specific theories (see Grossman, 2015), there is still a clear link

[13] More information on the dimensions is available in Articles I and II. The LISA project has expressed a preference that researchers not share anything directly from the PLATO manual or Norwegian version of the Tripod questionnaire, so they are not included in the thesis or appendices.

between PLATO and tenets of dialogic teaching (e.g., Alexander, 2014). Teachers who ask questions that are open and allow for longer student responses are rewarded with higher scores. Similarly, when students are given more opportunities to discuss, it also provides higher scores for teaching. In other words, according to PLATO, teaching is perceived as effective when it is more dialogic and the students are included (Grossman, 2015; Grossman et al., 2014). There is also a preference for explicit teaching (e.g., Purpose, Strategy Use and Instruction), as the teacher will score higher for teaching that explains what an activity is and why and how it will be done. This can be linked to tenets of SLA and FFI, which treat language learning more explicitly.

PLATO was chosen for several reasons. Firstly, it is a validated and standardized observation manual that is based on previous research (Grossman, 2015, p. 2). This could heighten the validity and reliability of my findings. Secondly, PLATO gave my research a clear focus, with results that can be comparable to other studies that use the same tool. This can create a common language for educational researchers as much observational research is qualitative and disparate (Klette & Blikstad-Balas, 2018, p. 132). The CLIL literature calls for more research on content; PLATO investigates features of content teaching, such as types of tasks, how the teacher explains content, and so forth. This could enable more interdisciplinary research with involvement from content researchers (see Article I). There were also practical reasons for using the tool: the University of Oslo had several PLATO researchers who could double-code my data, heightening the reliability of the results. Lastly, several of the dimensions in the manual (e.g., Classroom Discourse, Accommodations for Language Learners, and Intellectual Challenge) were highly relevant for CLIL teaching and could be connected to previous CLIL research, such as classroom discourse, use of the L1, CLIL pedagogy, and so forth. Using a tool that primarily analyzes dialogue (PLATO), I not only investigate what is said; the dialogue becomes a manifestation of what has been thought and understood in the CLIL classroom.

Researchers who use PLATO must be certified by Stanford University. In 2016, I received certification. PLATO researchers operate with 15-minute segments. We watch 15 minutes of teaching and assign a score from 1–4 based on PLATO criteria. Scores of 1 or 2 are considered

low-end, scores of 3 or 4 are high-end.[14] I began by transcribing all utterances by teachers and students and then coded the video data. For Article I, I worked with certified PLATO researchers in science and mathematics. We discussed how to interpret the various dimensions of teaching in different subjects, e.g., what constitutes high-quality reading and writing in mathematics. We coded one hour of video data from Article I together, discussing each score to heighten the inter-rater agreement level (Heyman, Lorber, Eddy, & West, 2014). We then coded separately; I coded all of the data, and another researcher coded 25% of the same video data. Then, we went through each code to see how often we agreed. This is *double-coding*, which was conducted to check the inter-rater agreement level. We discussed each instance in which we disagreed about a code and changed the score if we both agreed to do so. For Article II, one other PLATO researcher double-coded 25% of the data, but we did not code video data together. In general, there was a high level of inter-rater agreement for both articles (we agreed between 80 and 90% of the time). However, some of PLATO's wording is ambiguous and can be open to interpretation. For example, one of the "supportive materials" for the dimension "Accommodations for Language Learning" concerned "visual displays" (Grossman, 2015). In Article II, the rater and I could not agree on whether YouTube videos that showed illustrations of the phenomenon the students were learning constituted a visual display.

As useful as the PLATO manual may be, there are drawbacks to standardized coding manuals. One of the challenges of working with PLATO in the case of my video data was that PLATO is teacher-centered. In other words, if the teacher is not at the center of teaching (such as in project-based teaching), the teacher will score low even if the students are active and learning. This is because PLATO rates how well a teacher does or says something but focuses little on what the students are doing. PLATO is, therefore, only appropriate for specific teaching formats, such as whole-classroom teaching. For this reason, I chose not to include the video data for Schools 3 and 4. These schools were filmed during projects, so the teacher would have been scored unjustly lower. The numbers might have misrepresented the teaching. Klette and Blikstad-Balas (2018) explain that coding manuals are *reductionist* (they only locate pre-defined,

[14] For more details on the data analysis, see the methods sections in Articles I and II.

decontextualized aspects of teaching). Relying on pre-defined codes may lead to *magnification* (if a teacher scores low on a code, this does not explain why and may provide misleading results) (pp. 132, 134). Furthermore, the PLATO manual was created specifically for English language teaching in the United States. Nor could I alter the tool to fit my data because it was not deemed desirable by the LISA team. Due to these factors, PLATO may not fully capture the practices of bilingual teaching or the Norwegian educational context. For these reasons, I found it important to turn to the student perceptions of CLIL teaching to supplement my findings from Articles I and II. Due to the reductionist nature of coding manuals, another tool with a more open structure might capture aspects of teaching that were not captured by the PLATO analyses.

4.4.2 Questionnaires I and II (Phase III)

To investigate the student perspective (Article III), two questionnaires were chosen; a mixed and a quantitative questionnaire. A questionnaire is a practical and quick way to gather information about a topic that is easy to distribute at the end of a lesson (Gall et al., 2007). In the following, I will describe the two questionnaires used.

4.4.2.1 Questionnaire I

The first questionnaire was an online questionnaire in Norwegian that I developed to gain more insight into the students' backgrounds, attitudes, and beliefs concerning CLIL. I compiled the questionnaire myself but based it loosely on the questionnaire compiled by Hellekjær (2005) to elicit information on how Norwegian students experience reading course material in English (pp. 263–284). The questionnaire was first piloted to a non-CLIL class in the same grade (11th grade). The questions themselves were not altered from pilot to final product, but the procedure was changed after the pilot group (e.g., due to when and where it was distributed). One example was that I needed to distribute it myself; if the teacher distributed it, the students perceived the task as being an obligatory part of the lesson, and this created problems consent-wise.

Questionnaire I consisted of 18 questions related to the student's linguistic background, their English habits (e.g., how much English they read after school and what types of texts they read), their attitudes toward English, their own assessment of English proficiency, what motivated them to apply for a CLIL class, and the pros and cons of attending a CLIL class.[15] The questions were both closed (answers on a scale or checked in a box) and open-ended (students answered however they wished) (Johnson & Christensen, 2012, p. 169). Originally, I intended to characterize CLIL students based on their language habits but decided instead to delimit the research focus to their perceptions of teaching. To analyze the data, I used content analysis: qualitative research that identifies patterns, themes, and meaning in a body of research (Berg, 2009, p. 338). I derived codes from the data inductively to formulate theories about how students perceived CLIL.[16]

The drawback of using a questionnaire to elicit opinions rather than an interview is that you cannot probe the students further (Frankfort-Nachmias & Nachmias, 1992). There is no way of asking students to elaborate or clarify an answer, which may have led to more robust data. The data from Questionnaire I were used in Article III to answer the question of how students perceived their CLIL program (RQs 3 and 4 in Table 1).

4.4.2.2 Questionnaire II

The second questionnaire was a written version of the Tripod questionnaire. The Tripod questionnaire elicits answers about how students perceive their teaching (Tripod Education Partners, 2015). It was developed, piloted, and validated at Harvard University and is one of the most widely used student questionnaires in the United States of America (Tripod Education Partners, 2015). It was translated into Norwegian and piloted in Norwegian secondary schools by the LISA research team, and it has been distributed to over 1,000 Norwegian students (Klette et al., 2017). The questionnaire uses a 5-point Likert scale with the five options *never, rarely,*

[15] See Appendix 2

[16] For additional detail, see the methods section of Article III.

sometimes, often, and *always*. Tripod covers seven dimensions of teaching: how the students report on how well teachers *care, confer, captivate, clarify* and *consolidate* material, *challenge* them as well as on their c*lassroom management* abilities. Each of the 43 questions belongs to one of these domains (Tripod Education Partners, 2015). For Article III, my co-author (Helga Norheim) and I analyzed the data quantitatively. We used SPSS 25 to calculate the mean and standard deviation and conduct independent *t*-tests to see if there were significant differences between the two CLIL schools.

I chose the Tripod questionnaire because it provided insights into specific aspects of teaching that could be comparable with the PLATO manual, e.g., the Tripod dimension "clarify" could be compared to the dimension "instructional explanations" in PLATO as they both investigate how the teacher explains material (Klette et al., 2017). Both tools were used in the Measures of Effective Teaching (MET) project and the LISA/LISE projects at the University of Oslo, illustrating the compatibility of the tools. Since Tripod has been piloted and validated in the Norwegian context, the questions were confirmed to be relevant and appropriate in the Norwegian educational context (Klette et al., 2017).

Tripod was considered relevant for characterizing CLIL teaching because student perspectives can be used as a tool to assess teaching (Wallace et al., 2016, p. 1836). Standardizing student answers allowed CLIL classes to be compared with one other to see if students experienced the teaching differently depending on the teacher. The questionnaire contained questions related to the content teaching itself (e.g., whether students understood teacher explanations), which has been sought after in CLIL student perspectives (Somers & Llinares, 2018, p. 13). The Tripod questionnaire is quick to complete (the students fill it out during the last 10 minutes of the last recorded lesson) and easy to code. However, the questionnaire consisted of closed questions. Although this is useful for quantitative information-gathering, participants are forced to answer questions they may not find to be the most relevant to the topic and to choose from a limited number of answers (Johnson & Christensen, 2012). Furthermore, Wallace et al. (2016) problematize the dimensions of teaching in the Tripod; they point out that while a student may give high scores to items for "clarify," the questions may not fully cover the construct and may,

therefore, not actually determine how well the student perceives how well the teacher clarifies material (p. 1859). Student interviews may have been more in line with a sociocultural approach since this method would provide more contextual information about the students.

The data from Questionnaire II were used in Article III to compare science teaching in two CLIL programs (RQ5 in Table 1). This addressed how students experience their content teaching, which is greatly sought after in the CLIL literature (Somers & Llinares, 2018, p. 13).

4.5 Research credibility

Research credibility refers to how trustworthy, objective, and authentic research is (Creswell, 2009; Peräkylä, 2011). Researchers commonly divide credibility between two aspects: *validity* (if the methods, tools, and theory measure the construct the researcher is investigating) and *reliability* (how consistent the results are) (Johnson & Christensen, 2012, pp. 593, 597). In a complex project with several data sources and methods, it is important to discuss the strengths and weaknesses of the design. I have taken several precautions to strengthen the credibility of my research. The next section discusses the validity, reliability, and ethical concerns of my Ph.D. project and how they have been treated to strengthen the research credibility.

4.5.1 Reliability

In quantitative research, reliability refers to "the consistency or stability of a set of test scores" (Johnson & Christensen, 2012, p. 138). In qualitative and mixed research, this definition is broadened to express other types of replicability—for instance, whether the research process is consistent and stable over time (Miles, Huberman, & Saldaña, 2014). First, I piloted the project by visiting schools and interviewing and observing different classrooms. This led to refining the tools used in the project (Miles et al., 2014).

Reliability was important for PLATO because the scores needed to be stable; i.e., the findings cannot be dependent on who is coding, when the coding takes place, and so forth. Two aspects of reliability were important: inter- and intra-rater reliability (Johnson & Christensen, 2012, p.

524). *Inter-rater reliability* refers to the degree of agreement among raters. To ensure this, I discussed codes with three other raters, who were experts in the subjects I was coding (e.g., science and mathematics), in addition to my co-authors, for Article I. I coded the first lesson of Article I with my co-authors to calibrate our coding. Then, an external rater double-coded 25% of the video data for Articles I and II. The inter-rater reliability for PLATO was high (80% for Article I, 90% for Article II), meaning that, generally, raters agreed on the same scores 80 to 90% of the time. *Intra-rater reliability* refers to whether the raters consistently score items the same way. Becoming a certified PLATO researcher raised my intra-rater reliability as I was trained to rate consistently.

4.5.2 Validity

Validity normally refers to the extent to which an instrument measures its construct (Ary et al., 2014, p. 224). For the quantitative data and analyses, validity issues could, for instance, be of concern if the PLATO codes measure scaffolding or if important variables that might have represented scaffolding more accurately are missing. Qualitative validity relies on the accuracy of procedures (Gibbs, 2003). Qualitative researchers concern themselves with whether they have an authentic portrait of the phenomenon they are researching (Miles et al., 2014, p. 313). This is particularly important when representing people. Qualitative questions of validity for my Ph.D. project might include whether the CLIL teaching is accurately portrayed. Johnson and Christensen (2012) suggest that in mixed research, it is important to address validity on both sides (pp. 275–276). They refer to this as *multiple validities*—"the extent to which the mixed methods researcher successfully addresses and resolves all relevant validity types" (p. 275). To ensure that I measured my construct (e.g., scaffolding), I took several precautions with the validity of the research design and process, which will be explained in this section.

The first step of ensuring the validity of my Ph.D. project was to address sequential validity. *Sequential validity* refers to the degree to which one step builds on an earlier stage and minimizes problems that may occur if the qualitative and quantitative components are conducted at different times (Johnson & Christensen, 2014, p. 274). As mentioned in Section

4.1.1, my project was emergent, allowing me to steer the focus of the Ph.D. project based on its context and needs (Creswell, 2009, p. 186). The choice of article themes and PLATO codes changed for each phase to complement the previous phase to ensure that I covered relevant aspects of CLIL teaching.

Additionally, I employed several *external auditors* (researchers unfamiliar with my project). My university provided three seminars in which an external auditor examined my research design and progress. This took the form of them reading my texts, giving feedback, and discussing my thesis for one and a half hours, with a final report. My research school, NAFOL (nasjonal forskerskole for lærerutdanning), held two masterclasses where prominent CLIL and TESOL (Teaching English as a Second Language) researchers assessed my articles and extended abstract. This ensured the quality of my research, as did the peer-reviews that my articles underwent and the research groups at my university that read my articles.

In qualitative research, it is important for the researcher to reflect on her position and how her background and viewpoints may affect her research (Creswell, 2009, p. 186). This is referred to as *reflexivity* and minimizes researcher bias (Johnson & Christensen, 2012, p. 265). More specifically, researchers must reflect on how they are positioned as researchers and how their cultural, social gender, class, and personal politics affect the research (Berg, 2009, p. 61). This is also highly relevant in a sociocultural approach because it forces the researcher to contextualize herself in the classroom setting. For me, this involved several steps. Being a former CLIL student, I had to reflect on my role as a researcher and what my own pre-conceived notions of CLIL were. I also had to decide on my role as a researcher. I ultimately chose to be a participant-as-observer, minimizing my presence, while still explaining my project and background to participants for ethical reasons (Johnson & Christensen, 2012, p. 209). Another additional issue was that of *bias*. CLIL researchers may be inclined to believe that CLIL is effective based on their own interests and wishes (Georgiou, 2012, p. 502). However, I wished to be objective and neutral in my findings and to base my conclusions on empirical findings. For this reason, I had to consciously frame my research while remaining honest about my results and

the shortcomings of my findings. It is also hoped that the processes of external auditing and rating limited bias.

Construct validity refers to the "validity of inferences made about a construct based on the measures, treatment, subjects, and settings used in an experimental study" (Ary et al., 2014, p. 313). Phases I and II measure high-order concepts: *content, language,* and *scaffolding*. This meant I had to be mindful of whether the tools I used to measure these concepts could accurately measure them. For example, the elements in the PLATO manual may not measure what other researchers would consider "content." To counter threats to construct validity, I defined my concepts clearly by building on previous literature and was transparent about the data analysis and results. Additionally, tools such as PLATO and TRIPOD have already been developed, piloted, and used over a period of time to ensure construct validity.

Internal validity addresses the researcher's ability to draw correct inferences from the data (Creswell, 2009; Johnson & Christensen, 2012). Internal validity examines factors such as selection (whether the participants are representative of the phenomenon) or extraneous variables. Concerning Phases I and II, I viewed it as important to control for extraneous variables that may explain why teachers taught the way they did. I strengthened the internal validity by gathering as much information from the school and the teachers as I could, often asking teachers why they chose certain themes, topics, methods, and so forth after the lesson. Despite these precautions, there were other variables that I might not have captured in my data collection, such as the teacher's choice of topic, teaching method, or the students' reaction to teaching. Therefore, I focused more on descriptive than causal research.

Finally, the three phases reflected *external validity*—the extent to which the results of my study can be generalized to other CLIL classes in Norway (Johnson & Christensen, 2012, pp. 256–257). Since CLIL in itself is context-based, it was important to be mindful of the extent to which the sample size could be generalized. For instance, the CLIL teaching could depend on the teacher's qualifications and motivations (Paulsen, 2010b) or the students and the environment in the classroom. Moreover, the sample size (three classrooms, seven teachers, and 76 students) is not sufficient to justify generalizability for the entire nation. As several methodologists point

out, generalizability is not the goal of qualitative research per se; therefore, the articles are more descriptive than causal, with disclaimers in regard to generalizability (Creswell, 2009; Gibbs, 2003; Johnson & Christensen, 2012). Each school was treated as a case study in the articles; the focus is centered on what was observed rather than the potential for generalizability.

4.6 Research ethics

Research ethics is fundamental to the educational sciences as our research subjects are often human beings. As Berg (2009) puts it: the cornerstone of research on humans is "do no harm" (p. 60). When conducting research, there are several ethical criteria to take into account. One involves rules and standards of research. My first step was, therefore, to have my Ph.D. project approved. Before data collection, I sent my project description to Norwegian Social Science Data Services (NSD) for review. NSD confirmed that all procedures, filming, and surveys followed Norway's ethical guidelines and greenlit the Ph.D. project. I developed a consent form that was given to all participants, as well as an oral explanation of my project (see Appendix 3). All participants provided written consent before data collection began. The National Committee for Research Ethics in the Social Sciences and Humanities (NESH) recommends the age of consent for participation in research projects to be 15 years old (NESH, 2006). Participants under the age of 15 must have written consent from a legal guardian. This was accounted for in Phase I.

NESH also has strict guidelines for the storage of data. I was required to store the video data on a physically isolated computer on a secure server. The University of Oslo stored the data from School 1 on their server UiO/TSD 2.0 and will store the video data until they must be deleted in 2022. I stored all the data from Schools 2 and 3 on an encrypted disc with a password that only I had access to.

Anonymity is another vital concept in research (Ary et al., 2014, p. 51). All participants have the right to anonymity before, during, and after research. To ensure this, I assigned the schools, teachers, and students pseudonyms, as well as minimized the collection of indirect personal

data that might identify the participants (such as county or any prevalent characteristics). Since it is challenging and costly to anonymize video data (Derry et al., 2010, p. 36), I restricted the number of individuals who had access to the video data. The participants were furthermore made aware that they could withdraw from the project at any time.

Once one sets following the rules and regulations aside, there only remain the reflections the researcher must make when writing up the research (Creswell, 2007). As a researcher, I must take into account how the project will impact not only the participants but also the future of CLIL. CLIL is in a fragile state in Norway. There is very little research on CLIL in Norway in general, and if CLIL research points to negative results, this may give the Department of Education more incentive to shut down CLIL. As a CLIL researcher, this poses various ethical dilemmas. Although I must be objective and truthful in my findings, there is pressure from various stakeholders (schools, teachers, parents, and students) to find positive results so that CLIL will not be shut down by the government. For me, this has been a difficult balance to maintain, but I hope the honesty of the articles and heavy reliance on tools with comprehensive processes have minimized any biases I might have.

4.7 Limitations of the Ph.D. project

Now that the methods have been accounted for, this section will discuss some of the limitations of the Ph.D. project, with emphasis on the research design.

4.7.1 Discrepancies between theory and choice of tools

My primary concern is the discrepancy between the sociocultural view (which views teachers and students as equally important) and the teacher-centric tools employed in the thesis. The LISA research design consistently focuses on the teacher; the microphone captures all of the teacher utterances, and there is a microphone in the middle of the classroom, but what is said between students in groups is unknown. Filming classroom interaction is a methodological challenge for education researchers. It would be immensely taxing to hear and transcribe what thirty people are discussing simultaneously, which begs the question of what the focal point of

the data should be (Blikstad-Balas, 2016, p. 515). PLATO is teacher-centered as it focuses primarily on what the teacher does. When students are working independently, the teacher will score low because she is not necessarily speaking. This does not reflect a sociocultural perspective; PLATO cannot explain the learning that is happening between the students in this specific context. In this sense, the student perspective is limited in Articles I and II, although some of their utterances are transcribed during conversations with the teacher.

4.7.2 Lack of the teacher perspective

The lack of input from the teacher as a primary source of data is also a concern. Since teaching is the main focus of this thesis, the teacher's perspective should have also been represented. *Teacher cognition* concerns the teacher's beliefs and knowledge regarding how students learn and how this may affect their teaching practices (Barnard & Burns, 2012, p. 1). Especially in language teaching, teaching practices can be heavily influenced by teachers' beliefs about SLA (Kubanyiova & Feryok, 2015). Observation protocols do not normally include the teacher perspective; this is in line with the belief that the end result (what one observes in the classroom) is the most important aspect of teaching (cf. Klette & Blikstad-Balas, 2018). However, PLATO cannot explain the rationale behind a teacher's actions. Although I did interview the teachers to provide background information in Articles I and II, these interviews did not become a primary data source for this article. I considered it too ambitious a scope to include three methods for one Ph.D. project. Ultimately, I prioritized the student perspective as it is less explored than teacher perspectives in CLIL, particularly in the Norwegian context (see Fokus på språk, 2010, 2012).

4.7.3 Large-scale design to small-scale study

Another challenge was fitting LISA's large-scale, quantitative design to my small-scale, mixed methods study. The LISA design prescribes four hours of video data per subject (Klette et al., 2017, p. 8). This is because it was originally tailored for a research project that studied 50 Norwegian classrooms through a random sampling approach. However, I only wished to film

three or four schools, treating one or two schools per article as a case study. Four hours did not feel sufficient to capture the context of a CLIL classroom fully. While discussing the results of Article I with the CLIL teachers of the program, the teachers commented that the low scores in text-based instruction (i.e., how much the students were reading) were not representative of their teaching. Their reading sessions were project-based a few times a year, at times when I was not collecting data. This reflects upon the reliability of the project (Johnson & Christensen, 2012, p. 138) and the magnification of the results as the teaching could have possibly been scored differently in PLATO given a longer period of time. There is, therefore, a methodological tension between LISA's quantitative research design and my desire for a more qualitative approach to this research.

4.7.4 Comparing multigroups

The comparison between schools and teachers entailed a problem of validity. Johnson and Christensen (2012) cite this as *differential selection*—a threat to the internal validity of a multigroup study when a difference between groups exists (p. 254). This involved the transition from lower to upper secondary school, as there were many variables involved. Students in lower and upper secondary school can differ in a multitude of ways, including age, ability, knowledge, and so forth. This was also true of the CLIL teachers; teachers can differ in teaching styles, knowledge, and abilities, which may explain the differences in PLATO scores more than the differences between subjects do. My initial intent was to choose a range of CLIL teachers and students and triangulate the data to see how different or similar they were. In hindsight, this was an ambitious endeavor for a Ph.D. project, and I should have delimited my focus to the upper secondary level to strengthen the internal validity of the project. Despite these limitations, comparing two groups of students with the same tool and research design has borne fruit. Section 5.1.3 discusses some striking differences between lower and upper secondary school.

4.8 Brief Summary

This concludes the methods section of the thesis. To answer my main research question, "*What characterizes CLIL teaching in three Norwegian secondary classrooms*?," I filmed and observed three CLIL classrooms. I employed a total of 24 hours of video data, filming five subjects (science, mathematics, English, geography, and social science). The research design was multiphase and emergent, with three phases. Each phase represented an article, with one phase influencing the next. Phase 1 focused on how CLIL teachers taught content and language; Phase 2 focused on how CLIL teachers implemented scaffolding; and Phase 3 focused on student perspectives. The video data from Phases 1 and 2 were analyzed with an observation protocol, PLATO. Additionally, I distributed two questionnaires to the students to elicit knowledge about their thoughts and opinions on the CLIL teaching in Phase 3. Forty-three students answered Questionnaire 1 (mine), and 42 students answered Questionnaire 2 (Tripod).

The research design has its strengths and weaknesses. If given the chance to conduct this Ph.D. project again, I would have narrowed the scope and focus of the thesis. I might have delimited my project to the PLATO instrument, treating only one or two codes per article. The use of six codes per article meant that many results had to be eliminated. Conducting two types of methods (observation protocol and questionnaires) and the use of quantitative and qualitative aspects was challenging for a novice researcher within a limited timeframe. I would have narrowed down my participants to only upper secondary school, focusing on one or two classrooms. Upper secondary school is the most relevant for CLIL in Norway, and it is challenging to keep the lower secondary school anonymous as it is the only one of its kind. I would have strayed from the LISA design and adopted a more ethnographic approach, following the classroom and filming for longer periods of time. Yet, on the whole, I consider the methods used to be useful in gaining insight as to what happens in the CLIL classroom. This will be discussed more thoroughly in the next chapter.

5 Summary and discussion

In this final chapter, I will present the summary of each individual article, discussing the findings in light of the theory and literature presented in this thesis. The chapter will furthermore present the empirical, theoretical, and methodological contributions the thesis brings to CLIL and ELT in Norway. Then, I will suggest some implications for the Norwegian CLIL context and conclude the thesis.

5.1 Summary of the articles

5.1.1 Article I

Article I was titled "Characterizing CLIL Teaching: New Insights from a Lower Secondary Classroom." It was published in *International Journal of Bilingual Education and Bilingualism*[17] in April 2018. The article was co-written with Associate Professor Lisbeth M. Brevik of English Didactics and Professor Marianne Ødegaard of Science Education; both are affiliated with the University of Oslo and the LISE project.[18] The aim of the article was to discover the primary characteristics of CLIL teaching in terms of content and language. The terms *content* and *language* have their basis in integration theory and Met's (1999) continuum of Content and Language Integration. The research question was as follows:

> *RQ. What characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language?*

I filmed at a lower secondary school (School 1) for 12 lessons (Grade 9, ages 14–15) in the subjects of science, mathematics, and English, four lessons each. Six PLATO codes were used in

[17] For Norwegian readers: Level 1 journal

[18] For a full overview of how much each author contributed, see Declaration of Authorship in appendix 1

the analysis of the video data: *Purpose*[19] (how the purpose of the lesson is made explicit by the teacher), *Representation of Content* (how the teacher explains the material), *Intellectual challenge* (the intellectual rigor of student activities), *Classroom Discourse* (the formats and amount of speaking time students are provided), *Text-Based Instruction* (how teachers approach reading and writing in the classroom), and *Accommodations for Language Learning* (strategies teachers use to make the lesson available to L2 speakers).

5.1.1.1 Characterizing language

In terms of language findings, the L2 was used extensively; the teacher and students spoke English L2 83–97% of the time. This corresponds well with findings from Sweden (Paulsrud, 2019b). The science and math CLIL teachers frequently used, defined, and prompted students to use subject-specific terminology (understood in this thesis as CALP language). Similar to other CLIL classrooms in Europe, there were no tangible language learning goals expressed in the lessons apart from key vocabulary, and no evidence of FFI was observed (e.g., Dalton-Puffer, 2007; De Graaff et al., 2007; Schuitemaker-King, 2012). This points to the belief that students will learn language through immersion. Interestingly, the English language lessons did not have any language learning goals or FFI either. This may strengthen the argument that English functions more as a second language in Norway as the students are learning *through* the language instead of the language being an object of study itself (Paulsrud, 2019a, p. 295). This also erases the boundary between CLIL lessons and English language lessons in Norway—it would appear that the English lessons are equally content-driven.

There was also evidence of language support in the CLIL lessons. The science teacher, in particular, supported the lesson with strategies for second language learners, such as visual media. This finding suggests that even though CLIL teachers are not necessarily language teachers, they can still use strategies to support second language learners. The difference in results between science and mathematics may also point to subject-specific differences, in which science has more visual-based teaching (Lemke, 1990). These findings regarding

[19] All definitions taken from Grossman (2015).

language support led to my interest in *scaffolding* (Section 3.1.3.2), which encompasses other types of support related to learning as well. I wondered how teachers supported their students in a broader sense, and this became the topic of Article II.

Finally, I noted the differences in opportunities to speak, listen, and write in English in the CLIL and EFL teaching. The CLIL teaching was highly dialogic; students were given the opportunity to speak 83–86% of the time in the CLIL lessons. However, in EFL teaching, the students were only given the opportunity to speak 60% of the time. This harmonized well with other studies that suggest that CLIL teaching is more communicative than EFL teaching (Nikula, 2015; Van Kampen et al., 2016). The high PLATO scores in Classroom Discourse would suggest that CLIL teachers prompt students to justify or explain their answers over half of the time. This differs from the findings of Llinares and Pascual Peña (2015), who only discovered this 35% of the time in Spanish classrooms. The use of questions in CLIL classes suggests that the CLIL teachers encourage students to use critical skills. However, there was very little evidence of reading and writing in either the CLIL or English language teaching. The students did not write more than a few sentences in English and mathematics but did spend one science lesson partially writing a science report. There was little evidence of reading, except for in English, in which they read excerpts and discussed literature in one lesson. Due to the nature and tradition of the content subjects of science and mathematics, I question if there are often opportunities for students to engage in lengthy, in-depth reading and writing sessions. This would be interesting to investigate in more Norwegian CLIL classrooms since earlier research has implied that Norwegian CLIL students improve their reading through CLIL (Brevik & Moe, 2012; Hellekjær & Hopfenbeck, 2012).

5.1.1.2 Characterizing content

In terms of content findings, the observed CLIL teaching was considered to be content-driven on Met's (1999) continuum of content and language integration (see Section 2.3.2). The CLIL science lessons appeared to have the same curriculum, goals, and progress as an L1 science lesson. The CLIL teachers' explanations of material were long and rich. However, the science teacher focused much more on the conceptual understanding of material (79% of

explanations), whereas the mathematics teacher only did so in brief segments (17%), focusing more on solving specific tasks. Over half of the time, both CLIL teachers provided intellectually challenging tasks. This result is positive, indicating that learning through the L2 did not mean focus was detracted from the content subject matter. Since the students were recruited based on their English abilities, Cummins' (1976) threshold level of English illustrates that students can be provided with analytical tasks in the L2.

5.1.2 Article II

Article II was titled "The Comprehending Teacher: Scaffolding in Content and Language Integrated Learning (CLIL)." It was published online in *The Language Learning Journal*[20] in January, 2020. I was the sole author of the article. Intrigued by the amount of language support found in Article I, I wished to investigate this phenomenon further. Vygotsky's (2012) Zone of Proximal Development (ZPD) was used as a conceptual understanding framework explaining the difference between how students perform alone versus with the help of a teacher (Heritage, 2013, p. 181). Using the term *scaffolding*, I searched for literature on the support that teachers use with L2 students and found limited information on scaffolding in CLIL. The aim of the article was, therefore, to systematize the studies on scaffolding in ELL contexts in North America and CLIL in Europe and identify how scaffolding is practiced in an upper secondary Norwegian CLIL school. The research question was as follows:

> *RQ. Which scaffolding strategies do three CLIL teachers use to help their L2 English students comprehend material and complete tasks?*

I filmed 12 lessons in an 11th grade (ages 16–17) CLIL program in the subjects of science, geography, and social science (four lessons per subject) at School 2. Based on Maybin et al.'s (1992) definition, I identified scaffolding as strategies that teachers use to help students

[20] For Norwegian readers: Level 1 journal

comprehend material and *complete tasks.*[21] To empirically measure how teachers help students comprehend material, I used three PLATO codes: *Connections to Prior Knowledge*[22] (how a teacher connects new material to the students' prior knowledge), *Supportive Materials* (body language, visual aids, and graphic organizers that enable L2 English students to understand a lesson), and *Academic Language*[23] (the subject-specific terminology students need to understand a lesson). To empirically measure how teachers help students complete tasks, I used three other PLATO codes: *Uptake of Student Responses*[24] (How a teacher/student elaborates or follows up on ideas), *Strategy Use and Instruction* (a teacher's use of strategies and skills that support students' learning during the task at hand), and *Modeling and Use of Models* (how a teacher visibly enacts targeted strategies, skills, and processes).

On the whole, a range of scaffolding strategies were observed in the CLIL classroom. The CLIL teachers used significantly more comprehension strategies than task-solving strategies (hence the title). However, there were considerable differences between how the natural science (science and geography) teachers and the social science teacher practiced scaffolding. This pointed to a need for disciplinary literacy and to conceptually map scaffolding in various subjects. In general, the natural science teachers scored much higher on comprehension strategies. This was a surprising result as the comprehension strategies (especially supportive materials and academic language) are typically associated with language teaching (cf. Gibbons, 2015; Walqui, 2006).

5.1.2.1 Strategies to help students comprehend material

In terms of comprehension strategies, the CLIL teachers frequently made connections between known and unknown material. Particularly, the science teacher used everyday phenomena,

[21] For more information on definitions of these terms and the theoretical framework, see Article II.

[22] Definitions taken from Grossman (2015)

[23] "Supportive Materials" and "Academic Language" are subcodes of the code "Accommodations for Language Learning" found in Article I. These codes are therefore comparable.

[24] "Uptake of Student Responses" is also a subcode of "Classroom Discourse," found in Article I.

such as cutting an apple, to illustrate scientific concepts. For supportive materials, the geography teacher scored the highest, using body language to illustrate words and showing pictures and videos of landmarks. The social science teacher scored the lowest, rarely using visual aids. The L1 was often observed as a scaffolding tool; the teachers would use L1 terminology or explanations to clarify points. This finding concurs with prior research on L1 use (Gallagher & Colohan, 2014). Much as in Article I, subject-specific terminology was present in all lessons, often introduced as the core learning objective. Here again, geography scored the highest. As argued in Article I, subject-specific terminology appears to be the intersection of content and language. The geography teacher expressed the following thought during an interview: "The language is a tool to develop precision in the subject—to learn the terms that we use."

5.1.2.2 Strategies to help students solve tasks

There were much fewer scaffolding strategies in this category, notably, in the natural sciences. The tasks in the classroom were mainly identified as discussion tasks, in which students talked about key terminology in groups, forging their own definitions. The article is critical of this as the students were not producing any tangible products, such as writing texts and so forth. This differed from the lower secondary level in Article I; there, students were regularly producing posters, creating objects, and conducting laboratory work. In other words, in the lower secondary school, the CLIL teaching was more concentrated on the tangible aspects of learning, whereas in the upper secondary, the CLIL teaching was more abstract and theory-oriented. Due to the small sample, these observations are not generalizable as the findings may have depended on when I was filming, the specific teachers.

The social science teacher frequently used scaffolding strategies to help the students discuss theoretical concepts in the subject. More open-ended questions were observed and fewer IRF sequences and teacher-led discussions. The natural sciences displayed an overweighting of closed questions, which functioned as fact-checking. This is congruent with findings from other ELL classrooms (Dalton-Puffer, 2007; Llinares & Pascual Peña, 2015; McNeil, 2011; Mortimer, 2003).

The findings reveal that scaffolding manifests itself in different forms across subjects; these differences may be explained by the traditions and nature of the subject (Nikula, Dafouz, et al., 2016, p. 7). The natural sciences have their basis in real-world phenomena and are, therefore, compatible with scaffolding strategies that support visual aspects of teaching. Conversely, the social sciences often deal with more abstract issues that require personal input and function as discussion topics that stray from the IRF pattern. The one aspect of scaffolding that all three subjects have in common is the lack of modeling and strategy use. Limited instances of the teachers providing examples and suggestions of strategies to aid students in tasks were observed.

5.1.3 Article III

Article III was titled "Something New and Different: Student Perceptions of Content and Language Integrated Learning (CLIL)." It was submitted to *ELT Journal*[25] in April 2019 and is currently under review. The article was co-written with fellow Ph.D. candidate Helga Norheim (University of South-Eastern Norway), who is an expert on quantitative methods.[26] By following a sociocultural view that emphasizes dialogic teaching (Mercer, 1996; Mercer & Littleton, 2007; Wells & Arauz, 2006), I wished to introduce more student input into the Ph.D. project. The overarching aim of the article was to explore CLIL teaching from student perspectives, drawing on the previous themes of Articles I and II. Additionally, the article included open-ended questions for the CLIL students intended to capture aspects of CLIL teaching that PLATO could not. Three research questions were posed:

RQ1. Why do students choose CLIL programs?

RQ2. What pros and cons do students experience in CLIL teaching?

RQ3. How do students assess CLIL science teaching?

[25] For Norwegian readers: Level 2 journal

[26] For a full overview of how much each author contributed, see the Declaration of Authorship in Appendix 3

During the filming of the upper secondary classrooms, I distributed two questionnaires to the students of Schools 2 and 3 (for more on the procedures, see Section 4.3.1). The first questionnaire was directed toward the CLIL program as a whole, whereas the second was delimited to the assessment of science teaching. This was because the Tripod questionnaire has only been translated for the science subject (of content subjects), and science was the only CLIL subject at School 3. Questionnaires I (mine) and II (Tripod) consisted of 43 and 42 student responses, respectively.

5.1.3.1 Questionnaire I

The results of the questionnaires were largely favorable. The attitudes of the students of both CLIL programs were more positive than negative toward CLIL teaching. This is congruent with student assessments of CLIL in a variety of other countries (see Bower, 2019; Doiz, Lasagabaster, & Sierra, 2014; Lasagabaster & Sierra, 2009b; Roiha, 2019). In line with findings by Paulsrud (2019a) in Sweden, the Norwegian CLIL students largely chose CLIL because they viewed it as prestigious and valuable for their future. Learning English was not a goal per se, but rather improving various aspects of their English, such as vocabulary and fluency. Students also believed that learning in English would be new and interesting (hence the title), possibly leading to an environment where they could find like-minded peers.

The most-cited pro was that students reported improving their level of English. This may be attributable to the output hypothesis (Swain, 1985) due to the large amount of talking time observed in the lessons (Articles I and II). The second most-cited pro was the classroom environment. Both CLIL programs were marketed as "international," and the classrooms consisted of students of various nationalities, languages, and backgrounds. The students reported that learning in this environment was interesting and educational. Some students also felt more comfortable as Norwegian was not their L1, and they did not feel the pressure of having to be perfect in Norwegian.

Most of the cons reported about the CLIL program were language-related. The largest fears were loss of the L1 and that there was little focus on subject-specific terminology in the L1. This

was problematic since, in School 3, the students would still be taking their exams in Norwegian. The students' reactions to the lack of the L1 are also congruent with the position of The Language Council of Norway (The Language Council of Norway, 2017a, 2017b). The fear of domain loss of subject-specific terminology in Norwegian is growing, even among CLIL students. This is the only study, to my knowledge, in which CLIL students request more L1 use because they are afraid the L2 will dominate. Other students reacted to the mixing of the L1 and L2 in the classroom, expressing frustration about code-switching between L1 and L2. To me, this reflects that the language policies of both CLIL schools are unclear, and the schools may need more strategic planning so that their CLIL students feel equipped for a bilingual future. Other cons included the school environment; CLIL students at School 3 perceived themselves as outsiders and not well integrated into the school environment. Furthermore, a lack of resources in English for the students was also noted; for example, the science teacher was unable to find a science book that covered all of the themes in the national curriculum. Three books were chosen to cover parts of it, leading to frustration among the students.

5.1.3.2 Questionnaire II

The Tripod questionnaire also had largely favorable results. The science teaching was perceived similarly in both CLIL programs, and many links could be drawn between the findings from PLATO, in Articles I and II, and the answers to the Tripod questionnaire. The Tripod questionnaire confirmed that the CLIL students found how the teachers explained material to be clear (content), in line with Article I. Tripod also elicited information on how students perceived their input (language). Articles I and II observed a significant amount of talking time for the students. Consistent with this, students perceived that their input was valued. This can be tied in with dialogic teaching; when the teacher creates space for students to speak, they respond more positively to the learning environment. The final aspect that Tripod examined was intellectual challenge. Much as in Article I, the students confirmed that they felt intellectually challenged during the lessons, but not overwhelmingly so. This may allay some of the fears of CLIL practitioners as the field is preoccupied with whether learning through the L2 will lead to a cognitive overload. There are mixed results regarding intellectual challenge in CLIL

(see Section 3.1.4), but it appears that this is not an issue in Schools 2 and 3. The selectivity of the programs and the threshold level of English for both programs may contribute to this. The only drawback indicated in the questionnaire was that students reported they were not given the opportunity to decide on activities. The students may not have felt as if they were co-constructors of their own knowledge, and this may affect their motivation toward the subject.

Table 4. Overview of the articles

Articles	I	II	III
Title of Article	Characterizing CLIL Teaching: New Insights from a Lower Secondary Classroom	The Comprehending Teacher: Scaffolding in Content and Language Integrated Learning (CLIL)	Something New and Different: Student Perceptions of Content and Language Integrated Learning (CLIL)
Author(s)	Karina Rose Mahan Lisbeth Brevik Marianne Ødegaard	Karina Rose Mahan	Karina Rose Mahan Helga Norheim
Research question(s)	RQ1: What characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language?	RQ2: Which scaffolding strategies do three CLIL teachers use to help their L2 English students comprehend material and complete tasks?	RQ3: Why do students choose CLIL programs? RQ4: What pros and cons do students experience in CLIL teaching? RQ5: How do students assess CLIL science teaching?
Theory	Sociocultural theory with a focus on integration and disciplinary literacy	Sociocultural theory with a focus on scaffolding in bilingual education	Sociocultural theory with a focus on the student perspective
Data	*Video data*: 4 consecutive lessons filmed in science, mathematics and English class (N=12)	*Video data*: 4 consecutive lessons filmed in science, geography and social science (N=12)	*Student survey data*: Qualitative questionnaire (N=43)

			Quantitative questionnaire (Tripod) (N=42)
Findings	The teaching was content-driven and intellectually challenging. There were clear explanations and language support. There were many opportunities to speak but not read and write.	The teachers showed many instances of strategies to help students comprehend material but few to solve tasks. There were significant differences between natural and social science teaching.	Students viewed CLIL as positive and interesting. Cons included lack of L1, code-switching school environment, and lack of resources. Students were positive toward the science teaching but noted a lack of decision input in activities.
Status	Published in the *International Journal of Bilingual Education and Bilingualism* (2018) Level 1	Published in *The Language Learning Journal* (2020) Level 1	Under review in *ELT Journal* Level 2

5.2 Research contributions

According to the guidelines for Ph.D. theses at the University of South-Eastern Norway, a Ph.D. thesis must "contribute to develop new expertise and knowledge" in its respective field(s) (The Norwegian Ministry of Education and Research, 2018, § 3-9, translation mine). Therefore, in the subsequent sections, I will highlight the empirical, theoretical, and methodological contributions my thesis contributes to the fields of CLIL and English didactics in Norway.

From a bird's-eye view, the thesis contributes by obtaining insight into the CLIL classroom. It contextualizes CLIL teaching in a country with little process-oriented research on bilingual teaching and in a milieu with heated debates about the status of English in Norway. The thesis uses a mixed methods design, which is rare in CLIL research (Nikula et al., 2013, p. 73). Furthermore, the research is interdisciplinary and uses validated tools from other fields (PLATO and Tripod), which is highly sought after in the CLIL literature (Coyle, 2007; Mearns et al., 2017). The topics chosen (in particular, teaching methods and scaffolding) have been subjected to only limited prior research, especially in the Norwegian context. It is hoped that the sum of these findings will inform a viable option for learning English outside of the English classroom in Norway.

5.2.1 Empirical contributions

The primary contributions of this thesis are empirical because this is, first and foremost, an empirical study. This means that this is a study based on observation, experiment, or experience (Johnson & Christensen, 2012, p. 584). The data inform the readers how CLIL manifests itself in Norwegian classrooms. More specifically, this thesis has investigated not only how CLIL teachers teach in terms of content, language, and scaffolding but also how students perceive this teaching. Section 5.1 provides an extensive discussion of the findings; therefore, in the following, I will provide three key findings by synthesizing findings from all three articles. The main findings of the projects are that CLIL is content-driven, language is taught implicitly, and language is shaped by the subject it is in (disciplinary literacy).

5.2.1.1 CLIL is content-driven

In line with a number of CLIL studies throughout Europe, this thesis has discovered that the three CLIL classrooms in this study can be said to be content-driven (Banegas, 2016; Llinares, 2015; Lorenzo, 2007). Following Met's (1999) definition of content-driven bilingual teaching, the observed CLIL teaching focused on content learning, with content objectives and evaluation of content mastery. The observations suggested that there was no negative impact on CLIL teaching in terms of content; although the students were learning through English L2, they were still provided with rich conceptual explanations and intellectually challenging tasks. These findings are important for future CLIL research, warranting more research with a focus on the content subject. Content was taught differently in the lower and secondary schools: the lower secondary school (Article I) focused on tangible tasks, such as posters and lab reports, whereas the upper secondary school (Article II) focused on discussion of theory.

5.2.1.2 Language is implicit in CLIL teaching

The thesis revealed that there were no language learning goals or evidence of form-focused instruction in the observed CLIL teaching. The CLIL teachers rarely corrected or discussed aspects of language outside of subject-specific terminology. This suggests that the schools in question adhere to principles of immersion, in which students learn English through exposure

and dialogue, as observed in much of the CLIL literature (Dalton-Puffer, 2007; De Graaff et al., 2007; Schuitemaker-King, 2012). The CLIL students expressed that they are positive toward CLIL teaching (in line with Bower, 2019; Pladevall-Ballester, 2013; Yoxsimer Paulsrud, 2014), but it is unclear when and how students should use the L1. Furthermore, the CLIL students commented on a lack of teaching of the L1, even though they were expected to know the presented terminology for exams in the L1. The implicitness and lack of language policies led to frustration among the students. However, students are given many opportunities to speak (over 80% of teaching time), and they use English L2 around 83–97% of the time. In line with the findings of other CLIL research, the content teachers show signs of scaffolding comprehension (Tavares, 2015; Van Kampen et al., 2016).

5.2.1.3 Integration plays a considerable role in CLIL teaching

A contribution of this thesis has been the identification and mapping of content and language integration in CLIL teaching (called for in Nikula, Dafouz, et al., 2016). In terms of integration, subject-specific terminology is identified as an intrinsic part of both content and language teaching in the CLIL literature (e.g., Lorenzo, 2007). By focusing on language in content classrooms, teachers will not only provide scaffolding for L2 students but also help them understand the nuances of content through terminology. The comparison of various CLIL subjects through one instrument (PLATO) verifies that there are indeed empirical differences between how subjects are taught. Similar findings have also been noted in teachers' self-reported practices in the Netherlands (Van Kampen et al., 2016). This points to the usefulness of disciplinary literacy (Shanahan & Shanahan, 2012) when discussing CLIL teaching as language learning will be determined by the content subject. However, we require even more information on how subjects differ from one another, particularly from the content researchers themselves. In years to come, I hope that CLIL research will develop CLIL pedagogies for each subject, e.g., CLIL science pedagogy, CLIL math pedagogy, and so forth, to tailor pedagogy to the needs of each individual subject.

5.2.2 Theoretical contributions

As stated in Section 2, the overall lens of the thesis is sociocultural, but the thesis employs a number of theories from SLA and CLIL research to explain how students acquire and learn languages. Since the thesis is empirical, the empirical findings inform the theoretical. The focus of the theoretical contribution lies in the suitability of the theories for the data and whether the findings can inform our understanding of CLIL.

5.2.2.1 A framework for scaffolding

The most significant theoretical contribution of this thesis has perhaps been the creation of a framework for scaffolding (Article II). I identified scaffolding in a language learning context to create a common language for researchers (building on the seminal works of Echevarría et al., 2017; Gibbons, 2015; Hogan & Pressley, 1997; Maybin et al., 1992; van de Pol, Volman, & Beishuizen, 2010; van Lier, 2004; Walqui, 2006). By synthesizing the relevant literature, I suggest how scaffolding can be identified and studied empirically in the CLIL classroom.

5.2.2.2 The sociocultural approach

The sociocultural approach proved useful to study CLIL. A relatively more qualitative investigation of various factors related to historical, cultural, and social aspects of the schools and classrooms captured valuable details that were of relevance to the articles (Mercer, 2004). The use of the dialogic approach (Alexander, 2014) was also of relevance because the large amount of student input serves as a meaningful springboard to discuss CLIL's contribution as a teaching method. Students reacted positively to the dialogic approach, suggesting that more student input may lead to higher levels of motivation (Article III). The sociocultural approach was, therefore, deemed a suitable lens to understand CLIL teaching. Yet, as discussed in Section 4.7.1, there were tensions between the sociocultural approach that focuses on interaction between two parties, and the use of quantitative instruments that focus largely on the teacher (see Section 4.7.1).

5.2.3 Methodological contributions

Two of the three tools used in this thesis (PLATO and Tripod) have already been piloted and validated extensively—also in the Norwegian educational context (Cohen, Schuldt, Brown, & Grossman, 2016; Grossman et al., 2014; Klette et al., 2017; Wallace et al., 2016). The methodological contributions of the thesis, therefore, are related to the suitability of the tools for studying CLIL teaching.

5.2.3.1 PLATO in bilingual education

PLATO proved useful in analyzing CLIL subjects. There were several categories related to language learning similar to the Sheltered Immersion Observation Protocol (Echevarría et al., 2017). For this reason, I found that PLATO covers many aspects of bilingual teaching and can be appropriate for CLIL. Although PLATO was originally meant for English language teaching in the United States, the codes were universal enough for a range of subjects, including mathematics, science, geography, and social sciences (also observed in Cohen, 2018). However, the codes need to be clarified and tailored to each subject. The raters and I struggled with, for instance, describing how effective reading strategies can be identified in mathematics. Additionally, the CLIL teaching consisted of many student projects. This was challenging for PLATO as it mainly measures what the teacher does (see Section 4.4.1). This illustrates a limitation of teacher-centered instruments; PLATO is only useful for teaching formats in which the teacher is at the center. Therefore, I recommend PLATO for whole-classroom teaching but not for teaching in which the students are the main actors. This underlines a major point: what the teacher does and what the students learn are not necessarily congruent (as found by Hattie, 2009, 2012).

5.2.3.2 The questionnaires

Questionnaire I (mine) was developed to quickly gather information from a number of students (Johnson & Christensen, 2012, pp. 162–163). It was comprehensive and covered many aspects of CLIL student perceptions. Nearly all of the students answered the open-ended questions (e.g., why they chose CLIL programs). However, most of the answers were short. The use of this

questionnaire is, therefore, questionable in terms of obtaining long answers to open-ended questions.

The results from Article III suggest that Questionnaire II (Tripod) is a highly compatible tool to pair with PLATO (also suggested in Klette et al., 2017). There was a high level of agreement between what was observed in Articles I and II through PLATO and what the students scored in Article III using Tripod. However, PLATO is a much more comprehensive tool than Tripod, and only a few categories were selected from each instrument. Therefore, if one were to compare the two instruments further, Tripod should be expanded on or supplemented with additional questions, as in Klette et al. (2017).

5.3 Implications for CLIL in Norway

This summation of the findings may inform stakeholders, discuss the pros and cons of CLIL in Norway, and suggest how to improve CLIL teaching. Despite a high level of English proficiency in Norway, CLIL holds an important place in the Norwegian educational system. CLIL targets CALP language and internationalizes Norwegian students, thus preparing them for the bilingual and multicultural future that they will most likely face. These goals are also expressed in the English subject curriculum, implying that they are relevant to language teaching and valued on a national level (The Norwegian Directorate for Education and Training, 2013a). The three schools reported that there were twice as many applicants as placements, illustrating the level of student demand for CLIL in Norway. By marketing CLIL as international, students from multicultural backgrounds may also surround themselves with like-minded peers. These factors make CLIL a unique learning opportunity that students largely appreciate.

5.3.1.1 Policy and planning

As useful as CLIL may be, there are a number of issues that need to be explored for CLIL to achieve its full potential in the Norwegian context. National policies and teacher training have been identified as key factors in the successful implementation of CLIL (Sylvén, 2013). The findings of this thesis, therefore, call for more policy and planning. Fears about domain loss are confirmed—although there is evidence of the L1 in the observed CLIL teaching, it is not

systematized. This may suggest a need for a CLIL curriculum or language policy that ensures proficiency in and presence of the L1 in CLIL content subjects. Lack of the L1 is also an issue in Norwegian tertiary education (see Schwach, Brandt, & Dalseng, 2012; Schwach & Dalseng, 2011; Schwach & Elken, 2018), implying that teachers in Norway generally lack the strategies needed to balance teaching through two languages. The Language Council of Norway has suggested "parallel language teaching" to counter this (The Language Council of Norway, 2017b). There is additionally a lack of resources in CLIL materials; the results of Article III call for materials in English L2 that suit the needs of the Norwegian national curricula. The lack of materials currently makes CLIL an arduous endeavor for CLIL teachers as they must individually compile their own materials (Baetens Beardsmore, 2009, p. 214).

Article III further suggests that School 3 requires more strategic planning on including CLIL students in the school. Students reported feeling isolated and different from the rest of the students at the school. Creating elite programs in Norway is a dangerous enterprise; it contradicts the Norwegian democratic notion of "fellesskolen" (comprehensive school)—that children have the right to the same level of education and equal opportunities (Store Norske Leksikon, 2009). In this context, being part of a program that recruits top students may be received poorly by the rest of the school. School 2 did not report any problems in this regard and is therefore considered to be a model school for inclusivity. This could be explored further in later research. A final suggestion concerns the potential of CLIL for foreign languages. Norway has documented CLIL in German and Spanish (The Norwegian National Center for Foreign Languages in Education, 2011a), but this has been a minor undertaking. With the growing need for more competence in foreign languages (cf. Hellekjær, 1991; Hellekjær, 2007), Norway may benefit from increasing the use of foreign languages in content subjects. Limiting the language of instruction to only Norwegian and Sámi not only affects English; it also affects a number of languages that the state has explicitly expressed wishes for students to develop more competence in (see The Norwegian Ministry of Education and Research, 2007a; The Norwegian Ministry of Education and Research, 2007b).

5.3.1.2 Improving CLIL teaching

CLIL teaching in Norway has yielded numerous favorable results. Norwegian teachers appear to be at a high enough proficiency level in English. They follow the subject curriculum, and there is no evidence of loss of content.[27] The CLIL teachers scaffold language, and students confirm that they understand the teachers. There is a high level of student input and opportunities for students to speak English. However, there are a few potential areas for improvement. One point is the lack of reading and writing in CLIL subjects; by increasing the focus on these language skills in the classroom, the students may receive more support in fostering said skills. Another point is the modeling and use of strategies to help students complete tasks. It seems implicit that students should understand what to do and how to do it; nevertheless, this could be rendered more explicit in the teaching as explicit teaching is more beneficial to learning (Hattie, 2009, p. 201). Moreover, students report that they do not feel included in the choice of topics and activities. The Norwegian national subject curricula are quite free, and students could easily be included in this process (see The Norwegian Directorate for Education and Training, 2013b). Lastly, I posit that the upper secondary teaching in School 2 may benefit from using similar teaching strategies to the lower secondary teaching in School 1: the lower secondary CLIL teachers had more variation and student involvement in tasks. The challenge lies in teaching abstract concepts and theory through more than just discussion.

5.4 Conclusion

This thesis has sought to characterize CLIL teaching in Norwegian secondary classrooms through observing content, language, scaffolding, and student perspectives. A number of traits were identified. Some traits were positive; others negative. Most importantly, we are beginning to hone in on what identifies effective bilingual teaching. By opening the black box of CLIL teaching in Norway, I hope to start a discussion surrounding these CLIL teacher "frogs" who are more than content teachers but not quite language teachers. This may move the discussion of

[27] This aspect could be explored much further, especially through comparing CLIL teaching to non-CLIL teaching, and product-oriented research on students' content knowledge

bilingual education in Europe toward *how* instead of *why* they should teach bilingual education (Navés, 2009) because, ultimately, "CLIL should not be based on guesswork, fashionable political ideas, or potential gains for a particular school; it should all be for the benefit of the student" (Sylvén, 2013, p. 316).

References

Alexander, R. (2014). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). UK: Dialogos UK Ltd.

Arribas, M. (2015). Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes. *Latin American Journal of Content and Language Integrated Learning, 9*(2), 267–292.

Ary, D., Walker, D. A., & Jacobs, L. C. (2014). *Introduction to research in education* (9th ed., international ed.). Belmont, CA: Wadsworth Cengage Learning.

Ashwin, P., & Boud, D. (2015). *Reflective teaching in higher education*. London: Bloomsbury Academic.

Baetens Beardsmore, H. (2009). Chapter 9: Language promotion by European supra-national institutions. In O. Garcia (Ed.), *Bilingual education in the 21st century: A global perspective* (pp. 197–217). Singapore: Wiley-Blackwell.

Ball, P., Kelly, K., & Clegg, J. (2016). *Putting CLIL into practice*. Oxford, UK: Oxford University Press.

Banegas, D. L. (2016). Integrating content and language in English language teaching in secondary education: Models, benefits, and challenges. *Studies in Second Language Learning and Teaching, 2*(1), 111. doi: 10.14746/ssllt.2012.2.1.6

Banse, H. W., Palacios, N. A., Merritt, E. G., & Rimm-Kaufman, S. E. (2017). Scaffolding English language learners' mathematical talk in the context of calendar math. *Journal of Educational Research, 110*(2), 199–208. doi: 10.1080/00220671.2015.1075187

Barnard, R., & Burns, A. (2012). *Researching language teacher cognition and practice: International case studies* (Vol. 27). Bristol, UK: Multilingual Matters.

Barnes, D. (1969). *Language in the secondary classroom*. Harmondsworth: Penguin.

Berg, B. L. (2009). *Qualitative research methods for the social sciences* (7th ed.). Boston: Allyn & Bacon.

Berger, A. (2016). Learning mathematics bilingually: An integrated language and mathematics model (ILMM) of word problem-solving processes in English as a foreign language. In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 73–100). UK, USA & Canada: Multilingual Matters.

Blikstad-Balas, M. (2016). Key challenges of using video when investigating social practices in education: Contextualization, magnification, and representation. *International Journal of Research & Method in Education*, 1–13. doi: 10.1080/1743727X.2016.1181162

Bower, K. (2019). 'Speaking French alive': Learner perspectives on their motivation in content and language integrated learning in England. *Innovation in Language Learning and Teaching, 13*(1), 45–60.

Brevik, L. M. (2015). *How teachers teach and readers read: Developing reading comprehension in English in Norwegian upper secondary school*. Unpublished Doctoral Thesis: University of Oslo.

Brevik, L. M., & Hellekjær, G. O. (2018). Outliers: Upper secondary school students who read better in the L2 than in L1. *International Journal of Educational Research, 89*(C), 80–91. doi: 10.1016/j.ijer.2017.10.001

Brevik, L. M., & Moe, E. (2012). Effects of CLIL teaching on language outcomes. In D. Tsagari & I. Csépes (Eds.), *Collaboration in language testing and assessment* (Vol. 26, pp. 213–227). Peter Lang AG.

Brinton, D. M., Snow, M. A., & Wesche, M. B. (1989). *Content-based second language instruction*. New York: Newbury House.

Brock-Utne, B. (2007). Language of instruction and student performance: New insights from research in Tanzania and South Africa. *International Review of Education, 53*(5), 509–530.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. system*. An International Journal of Educational Technology and Applied Linguistics, 39*(4), 523–532. doi: 10.1016/j.system.2011.08.002

Bruton, A. (2013). CLIL: Some of the reasons why … and why not. *System, 41*(3), 587–597. doi: 10.1016/j.system.2013.07.001

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford: Oxford University Press.

Bullock Report. (1975). *A language for life.* Retrieved from UK Government committee chaired by Alan Bullock.

Burner, T. (2016). Formative assessment of writing in English: A school-based study of perceptions, practices and transformations. Doctoral thesis. Retrieved from NTNU Open. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2379861

Cenoz, J. (2015). Content-based instruction and content and language integrated learning: The same or different? *Language, Culture and Curriculum, 28*(1), 8–24. doi: 10.1080/07908318.2014.1000922

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262.

Cohen, J. (2018). Practices that cross disciplines?: Revisiting explicit instruction in elementary mathematics and English language arts. *Teaching and Teacher Education, 69*, 324–335. doi: 10.1016/j.tate.2017.10.021

Cohen, J., Schuldt, L. C., Brown, L., & Grossman, P. (2016). Leveraging observation tools for instructional improvement: Exploring variability in uptake of ambitious instructional Practices. *Teachers College Record, 118*(11), 1–37.

Coyle, D. (2007). Content and language integrated learning: Toward 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. Van Deusen-Scholl & S. May (Eds.), *Encyclopedia of language and education. Second and foreign language education* (2nd ed., Vol. 4, pp. 97–111). New York: Springer Science and Business Media LLC.

Coyle, D. (2011). Post-method pedagogies: Using a second or other language as a learning tool in CLIL settings. In Y. Ruiz De Zarobe, J. M. Sierra, & F. Gallardo Del Puerto (Eds.), *Content and Foreign Language Integrated—Contributions to multilingualism in European contexts* (pp. 49–73). Bern: Peter Lang.

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. doi: 10.1080/13670050.2013.777384

Coyle, D., Hood, P., & Marsh, D. (2010). *Content and language integrated learning*. UK: Cambridge University Press.

Creese, A. (2005). Is this content-based language teaching? *Linguistics and Education, 16*(2), 188–204.

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angeles: SAGE.

Creswell, J. W. (2016). Reflections on the MMIRA *The Future of Mixed Methods* Task Force Report. *Journal of Mixed Methods Research, 10*(3), 215–219.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9*, 1–43.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121–129.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. doi: 10.13140/2.1.1334.9449.

Cummins, J. (1998). *Immersion education for the millennium: What have we learned from 30 years of research on second language immersion?* Katoh Gakuen, Japan: Second Katoh Gakuen International Symposium on Immersion and Bilingual Education.

Cummins, J. (2013). BICS and CALP: Empirical support, theoretical status, and policy implications of a controversial distinction. In M. Hawkins (Ed.), *Framing languages and literacies: Socially situated views and perspectives* (pp. 10–23). New York: Routledge.

Cummins, J., & Early, M. (2015). *Big ideas for expanding minds: Teaching English language learners across the curriculum*. Canada: Pearson.

Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms* (Vol. 20). John Benjamins Publishing Company.

Dalton-Puffer, C. (2011). Content and language integrated learning: From practice to principles? *Annual Review of Applied Linguistics, 31*, 182–204. Retrieved from http://dx.doi.org/10.1017/S0267190511000092

Dalton-Puffer, C., Hüttner, J., Schindelegger, V., & Smit, U. (2009). Technology-geeks speak out: What students think about vocational CLIL. *International CLIL Research Journal, 1*(2), 18–25.

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. doi: 10.1093/applin/amu010

Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). *Language use and language learning in CLIL classrooms*.

Daniels, H. (2001). *Vygotsky and pedagogy*. London: RoutledgeFalmer.

Davison, C. (2005). Learning your lines: Negotiating language and content in subject English. *Linguistics and Education, 16*(2), 219–237. doi: 10.1016/j.linged.2006.01.005

De Graaff, R. (2016). Foreword: Integrating content and language in education: Best of both worlds? In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education*. UK, USA & Canada: Multilingual Matters.

De Graaff, R., Koopman, G., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism, 10*(5), 603–624. doi: 10.2167/beb462.0

De Smet, A., Mettewie, L., Galand, B., Hiligsmann, P., & Van Mensel, L. (2018). Classroom anxiety and enjoyment in CLIL and non-CLIL: Does the target language matter? *Studies in Second Language Learning and Teaching, 8*(1), 47–71.

Dearden, J. (2015). *English as a medium of instruction—A growing global phenomenon*. UK: British Council.

Derry, S., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., . . . Sherin, B. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3–53. doi: 10.1080/10508400903452884

Díaz Pérez, W., Ellson, M., Harumi, I., Kärkkäinen, K., Langé, G., Lee, J., . . . Schmidt, T. (2019). *Implementing internationalization of academia: Teaching, learning and research through English*. Guadalajara: The University of Guadalajara

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). CLIL and motivation: The effect of individual and contextual variables. *The Language Learning Journal, 42*(2), 209–224. doi: 10.1080/09571736.2014.889508

Drew, I. (2013). Using content and language integrated learning (CLIL) in the 9th grade: Teaching about the Second World War from a global to local perspective in English lessons. In I. Pareliussen, B. B. Moen, A. Reinertsen, & T. Solhaug (Eds.), *FoU i Praksis 2012 conference proceedings* (pp. 69–77). Trondheim: Akademia forlag.

Echevarría, J., Vogt, M., & Short, D. (2017). *Making content comprehensible for English learners. The SIOP model* (5th ed.). USA: Pearson.

Education First. (2018). *EF English Proficiency Index*. Retrieved from https://www.ef.com/wwen/epi/

Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 35–47). Georgetown University Press.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R. (2014). Chapter 3: Principles of instructed second language learning. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 31–45). USA: National Geographic Learning.

Ellis, R., Loewen, S., Elder, C., Reinders, H., Erlam, R., & Philp, J. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Bristol, UK: Multilingual Matters.

European Commission. (1995). *White paper on education and training. Teaching and learning. Toward the learning society.* Belgium: European Commission.

Eurostat. (2018). *Foreign language learning statistics*. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics

Eurydice. (2006). *Content and language integrated learning (CLIL) at school in Europe*. Belgium: European Commission.

Evnitskaya, N., & Morton, T. (2011). Knowledge construction, meaning-making and interaction in CLIL science classroom communities of practice. *Language and Education, 25*(2), 109–127.

Fernández-Sanjurjo, J., Fernández-Costales, A., & Blanco, J. M. A. (2017). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism*, 1–14. doi: 10.1080/13670050.2017.1294142

Flamez, B., Lenz, A. S., Balkin, R. S., & Smith, R. L. (2017). *The literature review.* Alexandria, USA: American Counseling Association.

Fokus på språk. (2010). *CLIL: Engelsk og fremmedspråk i andre fag*. The Norwegian National Centre for Foreign Languages in Education.

Fokus på språk. (2012). *CLIL: Kombinert fag- og engelskopplæring i videregående skole* (Vol. 28). The Norwegian National Centre for Foreign Languages in Education.

Frankfort-Nachmias, C., & Nachmias, D. (1992). *Research methods in the social sciences* (4th ed. ed.). London: Arnold.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston, Mass: Allyn and Bacon.

Gallagher, F., & Colohan, G. (2014). T(w)o and fro: Using the L1 as a language teaching tool in the CLIL classroom. *The Language Learning Journal*, *45*(4), 1–14. doi: 10.1080/09571736.2014.947382

García Mayo, M. D. P., & Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System, 54*, 40–54. doi: 10.1016/j.system.2014.12.001

García, O. (2009). *Bilingual education in the 21st century*. Singapore: Wiley-Blackwell.

Gené Gil, M., Juan Garau, M., Salazar Noguera, J., Ministerio De Ciencia, E. I., & Generalitat De, C. (2012). A case study exploring oral language choice between the target language and the L1s in mainstream CLIL and EFL secondary education. *Revista de Lingüística y Lenguas Aplicadas, 7*(1). doi: 10.4995/rlyla.2012.1129

Georgiou, S. I. (2012). Reviewing the puzzle of CLIL. *ELT Journal, 66*(4), 495–504.

Gibbons, P. (2015). *Scaffolding language, scaffolding learning. Teaching English language learners in the mainstream classroom*. New Hampshire, USA: Heinemann.

Gibbs, G. (2003). *Analysing qualitative data (Qualitative research kit)*. Los Angeles: SAGE.

Gierlinger, E. (2015). "You can speak German, sir": On the complexity of teachers' L1 use in CLIL. *Language and Education*, *29*(4), 1–22. doi: 10.1080/09500782.2015.1023733

Gold, R. L. (1958). Roles in sociological field observations. *Social Forces, 36*(3), 217–223. doi: 10.2307/2573808

Grandinetti, M., Langellotti, M., & Ting, Y. L. T. (2013). How CLIL can provide a pragmatic means to renovate science education–Even in a sub-optimally bilingual context. *International Journal of Bilingual Education and Bilingualism, 16*(3), 354–374. doi: 10.1080/13670050.2013.777390

Grossman, P. (2015). *Protocol for Language Arts Teaching Observations (PLATO 5.0)*. Palo Alto: Stanford University.

Grossman, P., Cohen, J., & Brown, L. (2014). Understanding instructional quality in English language arts. In T. Kane, K. Kerr, & R. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 303–331). United States of America: Jossey-Bass.

Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education, 119*(3), 445–470. doi: 10.1086/669901

Grossman, P., & Morva, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184–205.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Hattie, J. (2012). Know thy impact. *Educational Leadership, 70*(1), 18–23.

Hellekjær, G. O. (1991). The use of and need for foreign language skills in 24 Norwegian export firms: A survey. *Språk og Språkundervisning, 1*, 21–44.

Hellekjær, G. O. (1996). Easy does it: Introducing pupils to bilingual instruction. *Språk og Språkundervisning, 3*, 9–14.

Hellekjær, G. O. (2005). *The acid test: Does upper secondary EFL instruction effectively prepare Norwegian students for the reading of English textbooks at colleges and universities?* (Unpublished Doctoral Thesis). University of Oslo.

Hellekjær, G. O. (2007). Fremmedspråk i norsk næringsliv: Engelsk er ikke nok! *Fokus på språk,* (3). Retrieved from: http://hdl.handle.net/11250/147925

Hellekjær, G. O., & Hopfenbeck, T. (2012). Lesing. *Fokus på språk, 28*, 84–124.

Heritage, M. (2013). Gathering evidence of student understanding. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 179–196). USA: SAGE Publications, Inc.

Heyman, R. E., Lorber, M. F., Eddy, J. M., & West, T. V. (2014). Behavioral observation and coding. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 345–372). New York: Cambridge University Press.

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning*. Canada: Brookline Books Inc.

Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education, 43*(3), 325–356. doi: 10.1080/0305764X.2013.786024

Ibsen, E. (2002). *Engelsk i Europa—2002. Norsk rapport fra en europeisk engelskundersøkelse om holdninger til og ferdigheter i engelsk ved utgangen av den obligatoriske grunnskolen* [*English in Europe—2002. Norwegian report from a European English survey about attitudes and English skills at the end of compulsory schooling*]. Oslo: University of Oslo. Retrieved from:

https://www.duo.uio.no/bitstream/handle/10852/32302/AD0402ma.pdf?sequence=1&isAllowed=y

Johnson, B., & Christensen, L. B. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4th ed.). Los Angeles, CA: SAGE.

Johnson, B., & Christensen, L. B. (2014). *Educational Research. Quantitative, Qualitative and Mixed Approaches* (5th ed.). USA: SAGE.

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. doi: 10.1080/09500780508668671

Kachru, B. (1985). Standards, codification and sociolinguistic realism. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge: Cambridge University Press.

Kinginger, C. (2002). Defining the zone of proximal development in US foreign language education. *Applied Linguistics, 23*(2), 240–261.

Klette, K., & Blikstad-Balas, M. (2018). Observation manuals as lenses to classroom teaching: Pitfalls and possibilities. *European Educational Research Journal, 17*(1), 129–146.

Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking instruction and student achievement. A research design for a new generation of classroom studies. *Acta Didactica, 11*(3), 1–19.

Kontio, J., & Sylvén, L. K. (2015). Language alternation and language norm in vocational content and language integrated learning. *Language Learning Journal, 43*(3), 271–285. doi: 10.1080/09571736.2015.1053279

Kozulin, A. (1986). Vygotsky in context. In A. Kozulin (Ed.), *Thought and language* (pp. xi–lvi). Cambridge, MA: Cambridge University Press.

Kozulin, A. (2012). Foreword. In E. Hanfmann, G. Vakar, & A. Kozulin (Eds.), *Thought and Language* (pp. ix–xxv). London: The MIT Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Kubanyiova, M., & Feryok, A. (2015). Language teacher cognition in applied linguistics research: Revisiting the territory, redrawing the boundaries, reclaiming the relevance. *Modern Language Journal, 99*(3), 435–449. doi: 10.1111/modl.12239

Kulbrandstad, L. A. (2003). Minoritetsspråk og minoritetsspråkbrukere i Norge [Minority Languages and Minority Language Users in Norway]. Retrieved from https://www.sprakradet.no/Vi-og-vart/Publikasjoner/Spraaknytt/Arkivet/2003/1_2/Minoritetsspraak/

Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). *CLIL in Spain: Implementation, results and teacher training*. Cambridge: Cambridge Scholars Publishing.

Lasagabaster, D., & Sierra, J. M. (2009a). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 94*(4), 367–375.

Lasagabaster, D., & Sierra, J. M. (2009b). Language attitudes in CLIL and traditional EFL classes. *International Journal of CLIL Research, 1*, 4–17.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, N.J: Ablex.

Lialikhova, D. (2018a). The impact of a short-term CLIL intervention project on Norwegian different ability ninth graders' oral development. *International Journal of Bilingual Education and Bilingualism*, 1–22. doi: 10.1080/13670050.2018.1509055

Lialikhova, D. (2018b). Triggers and constraints of lower secondary students' willingness to communicate orally in English in a CLIL setting in the Norwegian context. *Journal of Immersion and Content-Based Language Education, 6*(1), 27–56.

Lin, A. (2015). Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum, 28*(1), 74–89. doi: 10.1080/07908318.2014.1000926

Lin, A. (2016). *Language across the curriculum and CLIL in English as an additional language (EAL) context. Theory and Practice*. Singapore: Springer.

Llinares, A. (2015). Integration in CLIL: A proposal to inform research and successful pedagogy. *Language, Culture and Curriculum, 28*(1), 58–73. doi: 10.1080/07908318.2014.1000925

Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge: Cambridge University Press.

Llinares, A., & Pascual Peña, I. (2015). A genre approach to the effect of academic questions on CLIL students' language production. *Language and Education, 29*(1), 15–30. doi: 10.1080/09500782.2014.924964

Lo, Y. Y. (2015). How much L1 is too much? Teachers' language use in response to students' abilities and classroom interaction in content and language integrated learning. *International Journal of Bilingual Education and Bilingualism, 18*(3), 270–288. doi: 10.1080/13670050.2014.988112

Lo, Y. Y., & Macaro, E. (2015). Getting used to content and language integrated learning: What can classroom interaction reveal? *Language Learning Journal, 43*(3), 239–255. doi: 10.1080/09571736.2015.1053281

Long, M. H., & Sato, C. J. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 268–286). Rowley, MA: Newbury House.

Lorenzo, F. (2007). An analytical framework of language integration in L2 content-based courses: The European dimension. *Language and Education, 21*(6), 502–514. doi: 10.2167/le708.0

Lorenzo, F., & Dalton-Puffer, C. (2016). Historical literacy in CLIL: Telling the past in a second language. In *Conceptualising Integration in CLIL and Multilingual Education* (pp. 55–72). UK, USA & Canada: Multilingual Matters.

Mahan, K. R. (2020). The comprehending teacher: Scaffolding in Content and Language Integrated Learning (CLIL). doi: 10.1080/09571736.2019.1705879

Mahan, K. R., Brevik, L. M., & Ødegaard, M. (2018). Characterizing CLIL teaching: New insights from a lower secondary classroom. *International Journal of Bilingual Education and Bilingualism*. doi: https://doi.org/10.1080/13670050.2018.1472206

Mahan, K. R., & Norheim, H. (under review). Something new and different: Student perceptions of content and language integrated learning (CLIL).

Marsh, D. (2002). *CLIL/EMILE. The European dimension. Actions, trends, and foresight potential.* Jyväskylä: University of Jyväskylä.

Marsh, D., Maljers, A., & Hartiala, A.-K. (2001). *Profiling European CLIL classrooms. Languages Open Doors.* Jyväskylä: University of Jyväskylä and the European Platform for Dutch Education.

Maybin, J., Mercer, N., & Stierer, B. (1992). Scaffolding: Learning in the classroom. In K. Norman (Ed.), *Thinking voices. The work of the national oracy project* (pp. 186–195). London: Hodder & Stoughton.

McNeil, L. (2011). Using talk to scaffold referential questions for English language learners. *Teaching and Teacher Education, 28*(3). doi: 10.1016/j.tate.2011.11.005

Mearns, T., De Graaff, R., & Coyle, D. (2017). Motivation for or from bilingual education? A comparative study of learner views in the Netherlands. *International Journal of Bilingual Education and Bilingualism*, 1–14. doi: 10.1080/13670050.2017.1405906

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359–377. doi: 10.1016/S0959-4752(96)00021-7

Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics, 1*(2), 137–168.

Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology, 80*(1), 1–14.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. London: Routledge.

Met, M. (1999). *Content-based instruction: Defining terms, making decisions.* (T. N. F. L. Center Ed.). Washington D.C.: NFLC Reports.

Meyer, O., & Coyle, D. (2017). Pluriliteracies teaching for learning: Conceptualizing progression for deeper learning in literacies development. *European Journal of Applied Linguistics, 5*(2), 199–222.

Meyer, O., Coyle, D., Halbach, A., Schuck, K., & Ting, T. (2015). A pluriliteracies approach to content and language integrated learning–Mapping learner Progressions in knowledge construction and meaning-making. *Language, Culture and Curriculum, 28*(1), 41–57. doi: 10.1080/07908318.2014.1000924

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Los Angeles: SAGE.

Moe, E. (2010). CLIL og språkutvikling. *Fokus på språk, 24*, 38–61.

Mohan, B. (1986). *Language and content*. MA: Addison-Wesley.

Mohan, B., Leung, C., & Slater, T. (2010). Assessing language and content: A functional perspective. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 217–240). Bristol, UK: Multilingual Matters.

Moore, P. (2009). *On the emergence of L2 oracy in bilingual education. A comparative analysis of CLIL and mainstream learner talk.* (Unpublished doctoral thesis). Universidad Pablo de Olavide, Seville, Spain,

Mortimer, E. (2003). *Meaning making in secondary science classrooms*. Maidenhead: McGraw-Hill Education.

Navés, T. (2009). Effective content and language integrated learning (CLIL) programmes. In Y. Ruiz De Zarobe & R. M. Jimenéz Catalan (Eds.), *Content and language integrated learning. Evidence from research in Europe* (pp. 22–40). UK & USA: Multilingual Matters.

NESH. (2006). Guidelines for research ethics in the social sciences, law and the humanities.

Nightingale, R., & Safont, P. (2019). Conversational style and early academic language skills in CLIL and non-CLIL settings: A multilingual sociopragmatic perspective. *English Language Teaching, 12*(2), 37–56.

Nikula, T. (2005). English as an object and tool of study in classrooms: Interactional effects and pragmatic implications. *Linguistics and Education, 16*(1), 27–58. doi: 10.1016/j.linged.2005.10.001

Nikula, T. (2007). The IRF pattern and space for interaction: Comparing CLIL and EFL classrooms. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical perspectives on CLIL classroom discourse* (pp. 179–204). Frankfurt: Peter Lang

Nikula, T. (2010). On effects of CLIL on a teacher's language use. In C. Dalton-Puffer & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 105–124). Amsterdam, the Netherlands: John Benjamins.

Nikula, T. (2015). Hands-on tasks in CLIL science classrooms as sites for subject-specific language use and learning. *System, 54*, 14–27. doi: 10.1016/j.system.2015.04.003

Nikula, T., Dafouz, E., Moore, P., & Smit, U. (2016). *Conceptualising integration in CLIL and multilingual education*. UK, USA & Canada: Multilingual Matters.

Nikula, T., Dalton-Puffer, C., & Llinares, A. (2013). CLIL classroom discourse: Research from Europe. *Journal of Immersion and Content-Based Language Education, 1*(1), 70–100. doi: 10.1075/jicb.1.1.04nik

Nikula, T., Dalton-Puffer, C., Llinares, A., & Lorenzo, F. (2016). More than content and language: The complexity of integration in CLIL and bilingual education. In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 1–25). UK, USA & Canada: Multilingual Matters.

Nikula, T., & Moore, P. (2019). Exploring translanguaging in CLIL. *International Journal of Bilingual Education and Bilingualism, 22*(2), 237–249. doi: 10.1080/13670050.2016.1254151

Nikula, T., & Mård-Miettinen, K. (2014). Language Learning in immersion and CLIL classrooms. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of pragmatics*. John Benjamins Publishing Company.

Norwegian Universities and Colleges Admission Service. (2019). *Krav til norsk og engelsk* [*Requirements for Norwegian and English*]. Retrieved from https://www.samordnaopptak.no/info/utenlandsk_utdanning/romania/sprakkrav_uten_hoyere_utdanning/

Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary.* Doctoral thesis in Subject Matter Education at the Department of Education and Special Education, University of Gothenburg.

Olsson, E., & Sylvén, L. K. (2015). Extramural English and academic vocabulary. A longitudinal study of CLIL and non-CLIL students in Sweden. *Apples—Journal of Applied Language Studies, 9*(2), 77–103. doi: 10.17011/apples/urn.201512234129

Olsson, E., & Sylvén, L. K. (2019). English productive proficiency. In L. K. Sylvén (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 117–135). UK/USA: Multilingual Matters.

Ortega, L. (2009). *Understanding second language acquisition*. Routledge.

Otwinowska, A., & Forys, M. (2017). They learn the CLIL way, but do they like it? Affectivity and cognition in upper-primary CLIL classes. *International Journal of Bilingual Education and Bilingualism, 20*(5), 457–480. doi: 10.1080/13670050.2015.1051944

Paulsen, J. M. (2010a). *Hvilke faktorer har betydning for implementering og spredning av CLIL? : analyse av CLIL i videregående skole 2009–2010 : del 2*(Vol. nr. 3-2010).

Paulsen, J. M. (2010b). *Lærerperspektiver på CLIL : didaktiske valg og rammebetingelser : evaluering av CLIL i videregående skole 2009–2010 : del 1*(Vol. nr. 2-2010).

Paulsrud, B. (2019a). Just a little plus: The CLIL student perspective. In L. K. Sylvén (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 282–297). UK/USA: Multilingual Matters.

Paulsrud, B. (2019b). Mapping CLIL in Sweden. In L. K. Sylvén (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 19–34). UK/USA: Multilingual Matters.

Pavón Vázquez, V., & Ramos Ordóñez, M. d. C. (2019). Describing the use of the L1 in CLIL: An analysis of L1 communication strategies in classroom interaction. *International Journal of Bilingual Education and Bilingualism, 22*(1), 35–48. doi: 10.1080/13670050.2018.1511681

Pena Díaz, C., & Porto Requejo, M. D. (2008). Teacher beliefs in a CLIL education project. *Porta Linguarum, 10*, 151–161.

Pérez-Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. doi: 10.1080/13670050.2011.630064

Pérez-Cañado, M. L. (2016). Are teachers ready for CLIL? Evidence from a European study. *European Journal of Teacher Education, 39*(2), 202–221. doi: 10.1080/02619768.2016.1138104

Peräkylä, A. (2011). Validity in research on naturally occurring social interaction. In D. Silverman (Ed.), *Qualitative research* (pp. 365–382). London: SAGE.

Piaget, J. (1926). *The language and thought of the child*. London: Routledge & Kegan Paul.

Pladevall-Ballester, E. (2013). 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.

Rainey, E., & Moje, E. B. (2012). Building insider knowledge: Teaching students to read, write, and think within ELA and across the disciplines. *English Education, 45*(1), 71.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Italy: Cambridge University Press.

Rindal, U. E. (2013). Meaning in English: L2 attitudes, choices and pronunciation in Norway. Universitetet i Oslo.

Roiha, A. (2019). Investigating former pupils' experiences and perceptions of CLIL in Finland: A retrospective analysis. *Nordic Journal of Studies in Educational Policy*. doi: 10.1080/20020317.2019.1586514

Roiha, A., & Sommier, M. (2018). Viewing CLIL through the eyes of former pupils: Insights into foreign language and intercultural attitudes. *Language and Intercultural Communication, 18*(6), 631–647. doi: 10.1080/14708477.2018.1465069

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). Switzerland: Springer.

Ruiz de Zarobe, Y., & Jiménez Catalán, R. M. (2009). *Content and language integrated learning: Evidence from research in Europe*. UK: Multilingual Matters.

Ruiz de Zarobe, Y., & Zenotz, V. (2017). Learning strategies in CLIL classrooms: How does strategy instruction affect reading competence over time? *International Journal of Bilingual Education and Bilingualism, 21*(3), 319–331.

San Isidro, X., & Lasagabaster, D. (2018). Code-switching in a CLIL multilingual setting: A longitudinal qualitative study. *International Journal of Multilingualism*, 1–21. doi: 10.1080/14790718.2018.1477781

Sandnes, T. (2017). *Innvandrere i Norge* [*Immigrants in Norway*]. Retrieved from Norway: https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/332154?_ts=162901a1050

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, II*, 11–26.

Schuitemaker-King, J. (2012). *Teachers' strategies in providing opportunities for second language development.* (Unpublished Ph.D. thesis). University of Groningen, Groningen.

Schwach, V., Brandt, S. S., & Dalseng, C. F. (2012). "Det gikk på engelsk og norsk, engelsk og norsk". En undersøkelse av språk i pensum på grunnivå i høyere utdanning.

Schwach, V., & Dalseng, C. F. (2011). *Språk i pensumlitteratur* [*Language in the curriculum*]. Retrieved from NIFU [Nordic Institute for Studies in Innovation, Research and Education].

Schwach, V., & Elken, M. (2018). *Å snakke fag på et språk andre forstår. Norsk fagspråk i høyere utdanning og arbeidsliv* [*Talking subject in a language others understand. Norwegian terminology in higher education and work life*]. Retrieved from NIFU [Nordic Institute for Studies in Innovation, Research and Education].

Seargeant, P., & Swann, J. (2012). *English in the world: History, diversity, change*. Abingdon: Routledge.

Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education, 21*(4), 328–341.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13. doi: 10.3102/0013189X027002004

Shanahan, T., & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in Language Disorders, 32*(1), 7–18.

Shepard, L. (2005). Linking formative assessment to scaffolding. *Educational Leadership, 63*(3), 66–70.

Simensen, A. M. (2002). To fluer i en smekk: Historie på tysk, heimkunnskap på engelsk, musikk på fransk? [Two birds in one stone: History in German, Home Economics in English, Music in French?]. *Norsk pedagogisk tidsskrift, 86*(1), 18–24.

Simensen, A. M. (2014). Skolefaget engelsk ; fra britisk engelsk til mange slags "engelsker"- og veien videre. *Acta didactica Norge [elektronisk ressurs], 8*(2), 18. doi: 10.5617/adno.1138

Somers, T., & Llinares, A. (2018). Students' motivation for content and language integrated learning and the role of programme intensity. *International Journal of Bilingual Education and Bilingualism*, 1–16. doi: 10.1080/13670050.2018.1517722

Statistics Norway. (2019a). Innvandrere og norskfødte med innvandrerforeldre [Immigrants and Norwegians born with immigrant parents]. Retrieved from https://www.ssb.no/befolkning/statistikker/innvbef

Statistics Norway. (2019b). Videregående opplæring og annen videregående utdanning [High school participation]. Retrieved from https://www.ssb.no/vgu/

Store Norske Leksikon. (2009). Enhetsskole. Retrieved from https://snl.no/enhetsskole

Stray, J. H., & Wittek, L. (2014). *Pedagogikk : en grunnbok [Pedagogy - an introductory book]*. Oslo: Cappelen Damm akademisk.

Svenhard, B. W., Servant, K., Hellekjær, G. O., & Bøhn, H. (2007). Norway. In A. Maljers, D. Marsh, & D. Wolff (Eds.), *Windows on CLIL* (pp. 139–146). European Platform for Dutch Education.

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition*. Rowley, MA: Newbury House.

Swain, M., Kinnear, P., & Steinman, L. (2015). *Sociocultural theory in second language education: An introduction through narratives* (2nd ed.) S.l.: Channel View Publications.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371–391.

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. (2015). CLIL and non-CLIL students' beliefs about language. *Studies in Second Language Learning and Teaching, 5*(2), 251–272.

Sylvén, L. K. (2019a). CLIL, CLISS and the Swedish context: An overview. In L. K. Sylvén (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 3–18). UK/USA: Multilingual Matters.

Sylvén, L. K. (Ed.). (2019b). *Investigating content and language integrated learning: Insights from Swedish high schools*. UK/USA: Multilingual Matters.

Tavares, N. J. (2015). How strategic use of L1 in an L2-medium mathematics classroom facilitates L2 interaction and comprehension. *International Journal of Bilingual Education and Bilingualism, 18*(3), 319–335.

The Language Council of Norway. (2017a). *New students meet large parts of their curriculum in English*. Retrieved from http://www.sprakradet.no/Vi-og-vart/hva-skjer/Aktuelt/2017/nye-studenter-moter-store-deler-av-pensum-pa-engelsk/

The Language Council of Norway. (2017b). Parallellspråkbruk ved universiteter og høgskoler [Parallel language use at universities and colleges]. Retrieved from https://www.sprakradet.no/Sprakarbeid/Forsking-og-hogre-utdanning/parallellsprakbruk-i-universitets-og-hogskolesektoren/

The Norwegian Directorate for Education and Training. (2006). *English subject curriculum* [LK06] [Press release]. Retrieved from http://data.udir.no/kl06/ENG1-03.pdf?lang=eng

The Norwegian Directorate for Education and Training. (2013a). *English subject curriculum* [LK06]. Retrieved from http://data.udir.no/kl06/ENG1-03.pdf?lang=eng

The Norwegian Directorate for Education and Training. (2013b). *Natural science subject curriculum*. Retrieved from https://www.udir.no/kl06/NAT1-03?lplang=eng

The Norwegian Directorate for Education and Training. (2016). *Har eleven rett til særskilt språkopplæring?* [*Does the student have the right to mother tongue instruction?*].

Retrieved from https://www.udir.no/laring-og-trivsel/minoritetsspraklige/regleverk-som-gjelder-spesielt-for-minoritetsspraklige/sarskilt-sprakopplaring/Hva-betyr-retten-til-sarskilt-sprakopplaring/

The Norwegian Directorate for Education and Training. (2017). *Hvilke språk kan skoler undervise på?* [*Which languages can schools teach in?*]. Retrieved from https://www.udir.no/regelverk-og-tilsyn/skole-og-opplaring/saksbehandling/undervisningssprak/

The Norwegian Directorate for Education and Training. (2019a). *Kritisk tenkning og etisk bevissthet* [*Critical thinking and ethical consciousness*]. Retrieved from https://www.udir.no/laring-og-trivsel/lareplanverket/overordnet-del/opplaringens-verdigrunnlag/kritisk-tenkning-og-etisk-bevissthet/

The Norwegian Directorate for Education and Training. (2019b). *Læreplan i engelsk, timetall* [*English subject curriculum, number of hours*]. Retrieved from https://www.udir.no/kl06/ENG1-03/Hele/Timetall

The Norwegian Ministry of Education and Research. (2007a). *Språk bygger broer* [*Languages build bridges*]. *White Paper No. 23*.

The Norwegian Ministry of Education and Research. (2007b). *Språk åpner dører [Languages Open Doors]*.

The Norwegian Ministry of Education and Research. (2018). *Forskrift om graden ph.d. ved USN* [*The regulations for a Ph.D. degree at USN*]. Retrieved from https://lovdata.no/dokument/SF/forskrift/2017-12-14-2411

The Norwegian National Center for Foreign Languages in Education. (2011a). *CLI-prosjekter på barne- og ungdomstrinnet* [*CLIL projects at primary and lower secondary level*]. Retrieved from http://www.fremmedspraksenteret.no/nor/fremmedspraksenteret/forsking-og-utvikling/clil-content- -language-integrated-learning/clil_prosjekter-p%C3%A5-barne_-og-ungdomstrinnet

The Norwegian National Center for Foreign Languages in Education. (2011b). *CLIL-prosjekter i videregående skole 2009–2011* [*CLIL projects in upper secondary school, 2009–2011*].

Retrieved from http://www.fremmedspraksenteret.no/nor/fremmedspraksenteret/forsking-og-utvikling/clil-content-_-language-integrated-learning/clil_prosjekter-i-videreg%C3%A5ende-skole

The Norwegian National Center for Foreign Languages in Education. (2019). *Elevane sitt val av framandspråk på ungdomsskulen 2018–2019* [*Students choice of foreign languages in lower secondary school 2018–2019*]. Retrieved from: https://www.fremmedspraksenteret.no/neted/services/file/?hash=47bc6da9bcfbcc8cbe1b455ebcba416a

The Storting. (2008). *Stortingsmelding nummer 35. Mål og meining—ein heilskapleig norsk språkpolitikk [White paper no. 35. Language and meaning—A holistic Norwegian language policy]*. Retrieved from: https://www.regjeringen.no/no/dokumenter/stmeld-nr-35-2007-2008-/id519923/sec1

Tripod Education Partners. (2015). Retrieved from https://tripoded.com/survey-modules/

Turner, M. (2013). CLIL in Australia: The importance of context. *International Journal of Bilingual Education and Bilingualism, 16*(4), 395–410. doi: 10.1080/13670050.2012.691086

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271–296. doi: 10.1007/s10648-010-9127-6

Van Kampen, E., Admiraal, W., & Berry, A. (2016). Content and language integrated learning in the Netherlands: Teachers' self-reported pedagogical practices. *International Journal of Bilingual Education and Bilingualism, 1*(15).

Van Kampen, E., Mearns, T., Meirink, J., Admiraal, W., & Berry, A. (2018). How do we measure up?: A review of Dutch CLIL subject pedagogies against an international backdrop. In T. Mearns & de Graaf R. (Eds.), *CLIL and bilingual education in the Netherlands* (pp. 129–155). John Benjamins Publishing Company.

van Lier, L. (2004). *The ecology and semiotics of language learning. A sociocultural perspective*. Boston/Dordrecht/New York/London: Kluwer Academic Publishers.

Vikør, L. (2012). Fakta om norsk [Facts about Norwegian]. Retrieved from https://www.sprakradet.no/Spraka-vare/Norsk/fakta-om-norsk/

Vilbli.no. (2019). Utdanningsprogrammene [Educational programs in upper secondary school]. Retrieved from https://www.vilbli.no/nb/nb/no

Vygotsky, L. S. (2012). *Thought and Language*. Cambridge, MA: Massachusetts Institute of Technology.

Wallace, T. L., Kelcey, B., & Ruzek, E. (2016). What can student perception surveys tell us about teaching? Empirically testing the underlying structure of the Tripod student perception survey. *American Educational Research Journal, 53*(6), 1834–1868. doi: 10.3102/0002831216671864

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International Journal of Bilingual Education and Bilingualism, 9*(2), 159–180.

Watson, D. (1992). Ethnomethodology, conversation analysis and education: An overview. *International Review of Education, 38*(3), 257–274. doi: 10.1007/BF01101432

Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of the Learning Sciences, 15*(3), 379–428. doi: 10.1207/s15327809jls1503_3

Wiesemes, R. (2009). Developing theories of practices in CLIL: CLIL as post-method pedagogies? In Y. Ruiz de Zarobe & R. Jiminez Catalan (Eds.), *Content and language integrated learning: Evidence from Research in Europe* (pp. 41–62). Bristol, UK: Multilingual Matters.

Wittek, L. (2014). Kapittel 7- Sosiokulturelle tilnærminger til læring [Chapter 7- Sociocultural approaches to learning]. In J. H. Stray & L. Wittek (Eds.), *Pedagogikk - en grunnbook* (pp. 133–148). Latvia: Cappelen Damm AS.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. doi: 10.1111/j.1469-7610.1976.tb00381.

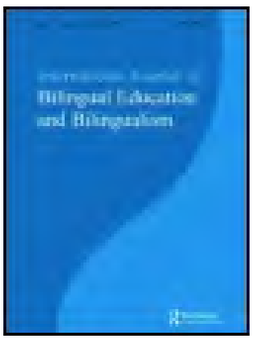
Yoxsimer Paulsrud, B. (2014). *English-medium instruction in Sweden: Perspectives and practices in two upper secondary schools.* Stockholm: Department of Language Education, Stockholm University.

Part II

The articles

Article I

Mahan, K. R., Brevik, L. M. & Ødegaard, M. (2018). Characterizing CLIL teaching: new insights from a lower secondary classroom. *International Journal of Bilingual Education and Bilingualism*, https://doi.org/10.1080/13670050.2018.1472206



International Journal of Bilingual Education and Bilingualism

ISSN: 1367-0050 (Print) 1747-7522 (Online) Journal homepage: http://www.tandfonline.com/loi/rbeb20

Characterizing CLIL teaching: new insights from a lower secondary classroom

Karina Rose Mahan, Lisbeth M. Brevik & Marianne Ødegaard

To cite this article: Karina Rose Mahan, Lisbeth M. Brevik & Marianne Ødegaard (2018): Characterizing CLIL teaching: new insights from a lower secondary classroom, International Journal of Bilingual Education and Bilingualism, DOI: 10.1080/13670050.2018.1472206

To link to this article: https://doi.org/10.1080/13670050.2018.1472206

Published online: 12 May 2018.

INTERNATIONAL JOURNAL OF BILINGUAL EDUCATION AND BILINGUALISM
https://doi.org/10.1080/13670050.2018.1472206

Characterizing CLIL teaching: new insights from a lower secondary classroom

Karina Rose Mahan [a], Lisbeth M. Brevik[b] and Marianne Ødegaard[b]

[a]Department of Language and Literature, University College of Southeast Norway, Tønsberg, Norway; [b]Department of Teacher Education and School Research, University of Oslo, Oslo, Norway

ABSTRACT

As a bilingual teaching method, Content and Language Integrated Learning (CLIL) is growing in popularity in Europe and research has primarily focused on (language) learning outcomes. Few studies have identified what characterizes teaching in the CLIL classroom in terms of content and language integration. Studying how CLIL is practiced is vital to understanding how it works and how students can benefit from it. In this study, we filmed and observed CLIL lessons in science and mathematics in a 9th grade, Norwegian CLIL class offering subjects in English. The present study uses *The Protocol for Language Arts Teaching Observation* (PLATO) to analyze video-recordings of CLIL lessons in science and mathematics, emphasizing a within-CLIL focus, and compared this with the English language teaching in the same class, as a baseline. Our findings indicate content-driven and intellectually challenging CLIL teaching with clear instructional explanations and systematic language support. English was used as frequently in the CLIL teaching as in the English teaching. Content and language were clearly integrated in the observed CLIL lessons, underscoring that the CLIL teachers successfully conveyed their subject in the target language.

ARTICLE HISTORY
Received 3 October 2017
Accepted 21 April 2018

KEYWORDS
CLIL teaching; CLIL pedagogy; content and language integrated learning; teaching practices

Introduction

A central research approach to understanding the Content and Language Integrated Learning (CLIL) methodology is identifying what happens in the CLIL classroom. Despite the growing body of CLIL research in Europe, its focus has primarily been on language outcomes in the form of tests (Brevik and Moe 2012; Coyle 2007; Georgiou 2012; Lasagabaster and Ruiz de Zarobe 2010; Llinares 2015). While recent years have shown a growing number of studies that focus on the interplay between content and language learning (e.g. Llinares, Morton, and Whittaker 2012; Nikula et al. 2016a), researchers have argued that large areas of CLIL teaching remain uninvestigated, especially content (cf. Cenoz, Genesee, and Gorter 2014; Fernández-Sanjurjo, Fernández-Costales, and Arias Blanco 2017). Studies on CLIL classrooms have mainly focused on language use (Dalton-Puffer 2007; Dalton-Puffer, Nikula, and Smit 2010; De Graaff, Koopman, Anikina and Westhoff 2007; Escobar Urmeneta 2013; Nikula 2010, 2015; Tavares 2015).

Recent research on CLIL has emphasized that integration should be a practical implementation as well as a theoretical lens (Llinares 2015; Ruiz de Zarobe and Cenoz 2015). Researchers therefore need to describe CLIL teaching practices to address 'the significant gap between CLIL theory and CLIL practice'. To close this gap, two areas in need of further research have been suggested, namely how CLIL is practiced, and how content is approached by the CLIL teachers (Admiraal, Westhoff, and De Bot 2006; Dalton-Puffer 2011). The present study addresses these issues by analyzing actions captured by video

CONTACT Karina Rose Mahan karina.mahan@usn.no

observations of naturalistic CLIL teaching in science and mathematics. The aim is *to identify what characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language.* Our approach is pedagogical and holistic, focusing on the nature of the content subjects as well.

Reviewing research on the CLIL classroom

One of the underlying justifications of CLIL is that students will benefit from it. This is referred to as the 'added value' of CLIL, understood as what CLIL contributes to the classroom that language and content subjects separately do not (Ball, Kelly, and Clegg 2016; Coyle, Hood, and Marsh 2010; Dalton-Puffer 2007; Marsh 2002). The most commonly cited advantage is the added opportunities to speak a foreign or second language (L2) in another subject. Research indicates that some skills seem to be positively affected by CLIL teaching (listening and reading) (Admiraal et al. 2006; Brevik and Moe 2012; Dalton-Puffer 2007; Lasagabaster and Ruiz De Zarobe 2010; Vollmer 2008). However, CLIL's added value has not been argued on the basis of classroom observations. Observational studies have largely used video and audio data from CLIL classrooms to describe patterns of discourse (see Dalton-Puffer 2007; Evnitskaya and Morton 2011; Llinares and Whittaker 2007; Moore 2009; Morton and Llinares 2016; Nikula 2010, 2005, 2015; Relaño Pastor 2015). We argue that observation can additionally be used to describe the 'range and practices by CLIL teachers,' which Van Kampen, Admiraal, and Berry (2016) refer to as CLIL pedagogy.

Van Kampen et al. (2016) analyzed survey data and interviews among CLIL and non-CLIL teachers, and argue that CLIL teaching is more interactive and dialogue-based than non-CLIL teaching, which corroborates findings from discourse analysis (Dalton-Puffer 2007; Dalton-Puffer, Nikula, and Smit 2010). Van Kampen et al. (2016) note that the weakest point of CLIL teachers' self-reported practices is their awareness and use of subject-specific literacies.

Another study concerns classroom observations of CLIL practices (De Graaff et al. 2007). To the best of our knowledge, this is the only CLIL study using an observation manual to determine successful practices. The study uncovered that CLIL teachers use a range of effective language teaching tools, including authentic materials and visual aids. In accordance with Van Kampen et al. (2016), neither CLIL teachers nor English teachers focused on correcting students' use of English. Instead, content subject teachers offered implicit language support (De Graaff et al. 2007, 620).

The 'disparate nature' of CLIL research combined with the many varieties of CLIL makes it difficult to characterize CLIL teaching, particularly since each content subject has its own needs and traditions (Van Kampen et al. 2016). Studies suggest that students who study science in their first language (L1) perform slightly better than their CLIL counterparts in the content subject (Fernández-Sanjurjo, Fernández-Costales, and Arias Blanco 2017), while CLIL science students largely improve their reading, writing, and grammar compared to non-CLIL science students (Pérez-Vidal and Roquet 2015). Although teachers who teach science in L1 have a wider repertoire of meaning-making, which gives nuances in instructional explanations (Nikula 2010), using L2 can facilitate science learning because CLIL teachers feel the need to plan lessons in greater detail (Grandinetti, Langellotti, and Ting 2013; Nikula 2010).

CLIL mathematics research has primarily focused on the relationship between content and language, and results are mixed. Studies have suggested that mathematics students at university level do not understand lectures or how to 'talk in math' in L2, CLIL students who are provided contextual language clues in arithmetic problems outscore students who do not (Miqdadi and Al-Jamal 2013; Van Rinsveld, Schiltz, Brunner, Landerl and Ugen 2016), and language scaffolding in CLIL mathematics unfolds similarly as in language teaching (Tavares 2015). These studies suggest that language is crucial to mathematics teaching, with a need to clarify the relationship between mathematics and language.

We conclude that CLIL teachers and students use many linguistic resources, but there needs to be a more systematic focus on how to scaffold content learning through language (Dalton-Puffer 2007;

Dalton-Puffer et al. 2010; Miqdadi and Al-Jamal 2013; Van Rinsveld et al. 2016). Furthermore, research needs to clarify how content is taught in CLIL teaching, with focus on subject-specific features (Meyer, Coyle, Halbach, Shuck and Ting 2015; Van Kampen et al. 2016). By observing how CLIL is taught in terms of content and language integration, our study aims to shed light on how CLIL science and mathematics teachers teach their respective subjects in L2. Bearing the aforementioned research gaps in mind, our study poses the following research question: *What characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language?* In order to examine this research question, we observe classroom teaching for the same students in both CLIL subjects. We have also decided to include observations of the class' English lessons for, and use their L2 teaching as a baseline to understand how English functions in their language subject, as done in other research (Nikula 2010).

Conceptualizing integration in CLIL

According to Vygotsky, learning is a social activity (Vygotsky 1978; Wertsch 1985). There is an intricate relationship between mental processes and communication, meaning learning is heavily imbued in social interaction (Hickmann 1985; Mercer 2004). In line with Vygotsky, 'the most significant moment in the course of intellectual development […] occurs when speech and practical activity, two previously completely independent lines of development, converge' (Vygotsky 1978, 24). Classroom talk, in this sense, becomes the 'chief locus of knowledge construction,' as subjects are 'talked into being' (Dalton-Puffer 2016, 29). Integration theories that operate on a local level (i.e. the classroom) are therefore often situated within a sociocultural framing (e.g. Evnitskaya and Morton 2011; Llinares and Whittaker 2010; Morton and Llinares 2016; Nikula 2010).

Following Vygotsky's thought, learning a subject means starting the process of becoming a member of a certain community (Sfard 1998, 6). In this view, 'content' and 'language' are complex processes that one cannot simply acquire, but rather participate in. Students must not only know and understand concepts in, for instance, science, but also be able to think, speak and write scientifically. Subjects are considered the result of historical processes, in which researchers and teachers are 'socialized in specific discourses and practices' (Nikula et al. 2016b, 7–8). The use of language depends on the content subject, which can vary in its structure of discourse and vocabulary (Shanahan and Shanahan 2012). Concepts which are gaining foothold in this direction include pluriliteracies (Meyer et al. 2015; Meyer and Coyle 2017) and disciplinary literacy (Airey, 2015; Shanahan and Shanahan 2012). These concepts resist language as generalized skills that can be applied across the curriculum, instead focusing on language skills necessary to understand the individual content subject.

Many CLIL scholars have attempted to clarify the integration between content and language (e.g. Berger 2016; Gajo 2007; Llinares 2015; Lorenzo 2007). Content is considered the antithesis of language, defined as 'any topic, theme, or non-language issue' (Genesee 1994, 3). Language, on the other hand, has often been perceived as developing 'skills in speaking, reading and writing, which are readily transferable to other areas of the curriculum' (Davison 2005, 221). CLIL scholars argue that what separates CLIL from other types of bilingual education is its preoccupation with integrating the two; seeing content and language as 'emergent synergies' that create a whole (Coyle, Hood, and Marsh 2010, 27). In our view, integration's goal is to draw on aspects of content and language teaching optimally in the classroom to foster learning. However, there are some problems in conceptualizing and realizing the potential of integration (Gajo and Serra 2002; Llinares 2015). For example, integration might entail 'mapping the characteristics and interplay of content and language,' De Graaff (2016, xiii). In this sense, integration is not simply adding content and language to a sum or applying language goals and methods to content subjects. Instead, it involves establishing the role and needs of content and language for each subject, as 'an integrated perspective on content and language is not the same in history as in physics teaching' (De Graaff 2016, xv).

It is acknowledged that literacy, in the fundamental and derived senses, is a crucial part of science (Norris and Phillips 2003). The fundamental sense is based on the essential role of text in science and

involves reading, writing, and being fluid in the discourse patterns and communication systems of science. The derived sense involves being knowledgeable and educated in science, and being able to take a critical stance on information (Norris and Phillips 2003; Ødegaard, Haug, Mork and Sørvik 2014). Mathematics requires a different approach. On the one hand, mathematical language has long been considered a language in its own right (Pimm 1987). Berger (2016) argues that although there are symbols and terminology unique to mathematics, 'mathematical content or understanding is inconceivable without the flexibility of everyday language' (Berger 2016, 75). Barwell (2005) concurs, emphasizing the need to understanding mathematics classroom problem genres (e.g. how to solve tasks), as well as solving them using subject-specific terminology. These understandings point to the importance of subject-specific literacies. In our study, we have used subject-specific literacies as a lens for understanding how the CLIL teachers teach the subjects within the context of its tradition. This has also aided us in understanding the use and role of language in science and mathematics.

CLIL in Norway

In Norway, CLIL is defined as teaching 30% or more of the curriculum in content subjects in a language other than L1 (Brevik and Moe 2012; Hellekjær 2005). The first CLIL initiative was sponsored by the Norwegian Ministry of Education and Research in 1993 (Svenhard et al. 2007). Since then, CLIL has been a grassroots initiative. Implementing CLIL is the responsibility of the individual county, school, or teacher. CLIL teaching varies between schools, including the number of CLIL subjects offered and the languages used. Although most CLIL initiatives use English as the teaching language, some at primary and lower secondary level have used French and German (Svenhard, Servant, Hellekjær and Bøhn 2007). The majority of CLIL classes in Norway are at the upper secondary school level, and a survey in 2004 indicated that 4%–7% of upper secondary schools offered some form of CLIL teaching (Svenhard et al. 2007). To the best of our knowledge, few lower secondary schools offer CLIL teaching in in Norway, and only one of these has done so consistently over time (since 2011). Little prior research has been conducted in these classrooms, which is why we do so in the present study.

Methods

To infer CLIL practices, we sought characteristics of CLIL teaching in mathematics and science in naturalistic video data. During the 2015–2016 school year, the Linking Instruction and Student Experiences (LISE) team collected data from one CLIL classroom. Four consecutive lessons in two CLIL subjects (science and mathematics), in addition to English L2 lessons, were filmed in the same class. This totaled 12 lessons (60 min each). The data enabled us to identify aspects of integration across CLIL subjects, and use the English lessons as a baseline concerning L2 language use.

Sample

Our sample was a 9th grade CLIL class in a Norwegian public school (ages 14–15). Students must apply for the CLIL program, taking an English reading test and an interview to prove their English level. The participants are the science, mathematics, and English teachers ($n = 3$), and the students

Table 1. Teacher background information.

Subject	Gender	L1	Teacher education	Education in the subject	Teaching experience	CLIL teacher
Science	Female	English	Yes	60 ECTS	6 years	6 years
Mathematics	Female	Norwegian	Yes	30 ECTS	2 years	1 year
English	Male	Norwegian	Yes	300 ECTS (Master's degree)	3 years	1 year

Note: ECTS = European Credit Transfer and Accumulation System.

of the CLIL class ($n = 26$). Table 1 offers background information on the teachers, including CLIL experiences.

Video recordings

Video recordings are valuable in classroom analysis due to the possibility of systematically investigating complex educational settings and deconstructing qualities of teaching (Blikstad-Balas 2016; Snell 2011). Our design relied on two cameras: one small, wall-mounted camera at the back of the classroom, facing the teacher; the other at the front, facing the whole classroom. We had two microphones; one on the teacher, the other capturing student conversations.

Data analysis

We analyzed the video data using the *Protocol for Language Arts Teaching Observation* (PLATO) (Grossman, Loeb, Cohen, Hammerness and Wyckoff 2010). PLATO is particularly relevant, as it is designed to assess content and language aspects of teaching. Despite PLATO's focus on language arts teaching, it has already been used to study mathematics teaching in the US, Finland and Norway (Cohen, Grossman, Borko, Loeb & Shavelson 2013; Cohen, Schuldt, Brown & Grossman 2016; Kane and Staiger 2012; Klette and Blikstad-Balas 2017; Klette, Blikstad-Balas, and Roe 2017; Luoto, Klette, and Blikstad-Balas under review; Stovner 2018). We applied it to CLIL teaching in mathematics and science.

PLATO consists of 13 elements considered to represent effective teaching (e.g. Klette and Blikstad-Balas 2017). Among these elements, we have chosen six that comply with the CLIL conceptualization of content and language integration. We used these elements in our data analysis to infer CLIL characteristics in the observed lessons, including integration of content and language (see Table 2). PLATO scores on a scale from 1 to 4, and assigns scores for every 15-minute segment of video data. Each recorded lesson lasted for approximately 60 min and was divided into 15-minute segments for analysis. Low-end teaching indicates there is no evidence (score 1) or little evidence (score 2) of the element in question. High-end teaching indicates evidence with some weaknesses (score 3) or strong and consistent evidence (score 4).

Purpose (PUR) examines how the purpose of a lesson is made explicit by the teacher and reflected in student activities (Grossman 2015). We used this element to investigate if the goals of the CLIL teaching were primarily content- or language-driven (Banegas 2016; Lasagabaster 2008; Met 1999).

Representation of content (ROC) denotes the teacher's accuracy in talking about their subject (Grossman 2015). PLATO differentiates between *instructional explanations* (how the teacher explains the content) and *conceptual richness* (the type of explanations offered). Accuracy means that the teacher adequately provides a sufficient level of explanation (score 3), although the explanations are not necessarily nuanced in ways that help students distinguish among different features of related ideas (Grossman 2015). Conversely, the examples, analogies, and/or explanations are not sufficiently complete to explain the concept, and only touch on surface level features of the content (score 2). As for conceptual richness, we differentiate between explanations that focus on deeper conceptual understanding (score 4), explanations that mainly focus on conceptual understanding (score 3), and superficial explanations that mostly focus on procedures, rules, or labeling terms (score 2). ROC is of relevance, as CLIL teachers often express concerns about talking accurately about their subject through L2 (Maasum, Maarof, Yamay & Zakaria 2012; Pérez-Cañado 2016; Šulista 2012).

Intellectual challenge (IC) represents the intellectual rigor of student activities, including student–teacher conversations (Grossman 2015). PLATO differentiates between low-level (rote and recall) and high-level activities (analyzing, synthesizing, and interpreting). IC enables us to observe if content and language are integrated enough for students to understand material and complete tasks (Coyle et al. 2010).

Classroom discourse (CD) examines what formats and how much speaking time the students are provided (*opportunity*), and how the teacher responds to and builds on student ideas (*uptake*)

Table 2. Content and language features of teaching, based on six PLATO elements (Grossman 2015).

	1: Almost no evidence	2: Provides limited evidence	3: Provides evidence with some weaknesses	4. Provides strong and consistent evidence
PUR	No clear learning goal or unrelated to disciplinary skills	Communicated or inferred goal, as a general disciplinary topic	Communicated, specific goal related to development of disciplinary skills. Activities align with goal	Communicated, specific goal related to development of disciplinary skills. Activities align with goal. Evidence of student awareness. Teacher refers back to goal
ROC	*Instructional explanations:* Weak or incorrect explanations of disciplinary concepts. *No conceptual richness*	*Instructional explanations*: Incomplete explanations which touch on surface-level features of subject content. *Conceptual richness*: Superficial representation, focusing on rules, labels, procedures. Little attention to deeper understanding	*Instructional explanations*: Accurate but un-nuanced explanations of disciplinary concepts. May address student misunderstandings. *Conceptual richness*: A balance of rules, labels, procedures. Attention to deeper understanding	*Instructional explanations*: Accurate and clear explanations, addressing student misunderstandings and highlighting nuances. *Conceptual richness*: Conceptual understanding of content beyond the superficial to focus on interpretation or deeper understanding
IC	Activities are rote/recall	Mostly rote/recall, some analysis/inference	Mostly analysis/inference/ idea generation/ interpretation	Mostly sophisticated or high-level analytic and inferential thinking
CD	*Opportunities*: Few or no opportunities for student talk. *Uptake*: Few or no response to students' ideas	*Opportunities*: Occasional opportunities for student talk. *Uptake*: Brief responses with no elaborative discussion or help to develop	*Opportunities*: Opportunities for student talk for at least 5 min. Only 2–3 students participate. *Uptake*: Teacher occasionally builds on student ideas (re-voices in academic language, asks for elaboration)	*Opportunities*: Opportunities for student talk for at least 5 min. The majority of students participate by speaking and/or listening. *Uptake*: A consistent engagement in high-level uptake
TBI	*Use*: No authentic text present. *Production*: No opportunities for students to engage in writing	*Use*: Refers to details in authentic text. *Production*: Brief pieces of connected text (at least 3 min)	*Use*: Active use of authentic text to gain understanding. *Production*: Sustained opportunities in a particular genre or structure (at least 7 min)	*Use*: Active use of authentic text for a sustained period of time (at least 7 min). *Production*: Sustained opportunities with attention to issues of writing, style, or genre (at least 7 min)
ALL	*Materials*: No supportive materials. *Academic language*: Teacher does not introduce, define, or prompt use of academic terms	*Materials*: Teacher provides relevant supportive materials, but are not used. *Academic language*: Teacher rarely introduces, defines, or prompts academic terms	*Materials*: Teacher provides and prompts use of relevant accessible, supportive materials. *Academic language*: Teacher introduces, defines, and highlights academic language	*Materials*: Teacher provides and prompts use of relevant materials. Evidence of use. *Academic language*: Teacher consistently introduces, defines, and highlights academic language. Students have multiple opportunities to use them

Note: PUR = Purpose. ROC = Representation of content. IC = Intellectual Challenge. CD = Classroom Discourse. TBI = Text Based Instruction. ALL = Accommodations for Language Learning.

(Grossman 2015). Opportunities for conversations about subject content is evidence toward low-end (score 2) if the conversations last less than one third of the segment, and evidence toward high-end (score 3 and 4) if the conversations last longer, stay on track, and include open-ended questions. Identifying how teachers and students talk is motivated by research suggesting that CLIL offers student-talking opportunities and high-quality conversations (Dalton-Puffer 2007; De Graaff et al. 2007; Nikula 2010).

Text-based instruction (TBI) relates to how teachers approach reading and writing in the classroom (Grossman 2015). TBI differentiates between the use of texts (reading) and production of texts (writing). This establishes the opportunities for student engagement with texts, acknowledging the subject-specific reading and writing conventions. PLATO states that 'authentic texts' comprises published material, student-generated work, pieces of music or art, graphs, tables, or film/video used for

teaching. Research indicates that writing opportunities in the CLIL classroom are limited, and often viewed as homework activity (Dalton-Puffer, 2007).

Accommodations for Language Learning (ALL) refers to strategies teachers use to make lessons available to L2 speakers through *supportive materials* and *academic language* (Grossman 2015). Supportive materials include visual aids that enable students to understand a lesson in L2. Academic language denotes subject-specific terminology related to the content of the lesson at the low end (score 2), including the teacher's strategic use of L1 to explain or prompt terminology. At the high end (scores 3–4), academic language consists of features to describe complex ideas, abstract concepts, and cognitive processes (thinking skills); including the discourse level (communicate, clarify and negotiate meaning), syntactic level (make messages, paragraphs, and sentences clear and correct), and lexical level (choose and use the best terms to convey meaning). In CLIL teaching, academic language instruction is needed for L2 learners who might struggle to understand and use the language of mathematics and science. When CLIL teachers provide language support, the amount may depend on the age or skills of the students (De Graaff et al. 2007; Harvey Tihinen, Määttä & Uusiautti 2013).

Research credibility and ethics

Several precautions were taken to ensure the trustworthiness of this study (Creswell 2009; Peräkylä 2011). First, following the ethical guidelines of the Norwegian Center for Research Data, written and informed consent was provided by parents, students, and teachers (NESH 2006). Second, using PLATO ensures the research is less prone to personal interpretations and allows for comparability (Klette and Blikstad-Balas 2017). The segments were coded by certified PLATO raters. 25% of the observations in each CLIL subject were double-scored by experts in the respective subjects to ensure high levels of ongoing interrater agreement ($\geq$ 80% exact-score agreement) (Cohen et al. 2016, 8).

A possible limitation of this study is that it will not capture all aspects of CLIL teaching, since the PLATO manual is not designed specifically for CLIL instruction. However, based on its use in the aforementioned prior studies of mathematics teaching in the US, Finland and Norway, and our own analysis, we believe that PLATO is nevertheless a useful tool in our study. The small sample does not allow for generalizability either (Johnson and Christensen 2014). In line with this, the present study is concerned with how CLIL is practiced in a specific setting, with no intention of generalization. Therefore, we believe our design provides valuable data on the characteristics of CLIL teaching.

Results

Results indicate that the CLIL teaching in science and mathematics addressed distinct characteristics relating to the integration of content and language. Both CLIL subjects were taught in English L2. Although the CLIL teachers offered language support and numerous opportunities to speak, there were few opportunities for reading and writing. The CLIL teaching was content-driven, with rich explanations, and intellectually challenging.

Language features of CLIL teaching

In the CLIL lessons, teachers and students spoke L2 83 –97– of the time, confirming a systematic L2 presence. Notably, L2 was used as much in the CLIL subjects as in the L2 subject. The CLIL teachers largely used L1 for administrative purposes, and to aid the students in understanding the content subjects.

Academic language (ALL)

While L2 seemed to be used in effective ways to scaffold content learning in the CLIL subjects, there did not appear to be any explicit focus on L2 apart from consistent use of academic language. Both

CLIL teachers used subject-specific terminology throughout their lessons (minimum score 2), but varied the extent to which they defined (score 3) or prompted students to use it (score 4). Using L2 subject as a baseline, Figure 1 shows that in L2 lessons, students only occasionally used academic language (7% high-end, scores 3–4), while they did so frequently in mathematics (58%) and science (79%), to describe complex ideas, abstract concepts, and thinking skills.

The science teacher consistently gave tasks related to subject-specific terminology, offering opportunities to negotiate scientific meaning both in L1 and L2. In Excerpt 1, she uses L1 (underlined) to aid the students in learning subject-specific L2 terminology. It is noteworthy how the teacher uses everyday language to support learning of L2 vocabulary:

Excerpt 1 (Science, Academic Language, Score 4):

Teacher: What does *etsende* mean in English? […] Anybody that would like to give the answer to that? *Etsende*, what's that? Starts with a C? What does it do?
Student: It eats your skin.
Teacher: Yeah, exactly. So, it eats up your skin … can eat your skin. […] Corrosive is the word. Corrosive.

The mathematics teacher also encouraged the students to use L1 and L2 terminology to clarify content meaning. In Excerpt 2, she asks three students to explain how one mathematical function can differ from another, prompting the students to use subject-specific terminology (underlined) and by doing so, negotiating content meaning:

Excerpt 2 (Mathematics, Academic Language, Score 4):

Teacher: Can you try to explain in mathematical terms? […]
Student 1: So, so, so, that's minus three, right?
Teacher: That's a negative. Minus negative. Slope? Ok, how can you see that?
Student 2: Because it's going that way?
Teacher: It's going the other way. Yeah. Do you want to explain further, […]?
Student 3: *Fordi stigningstallet er ved minus?* [Because the slope is at minus?]
Teacher: In English?
Student 3: Because the *stigningstall* is …
Teacher: What's the *stigningstall* in English?
Student 3: Slope.
Teacher: Yes, thank you.
Student 3: If the slope is minus, means that the Y goes downwards, not upwards, because the slope is downwards, not upwards.
Teacher: Ok. So, what does the 'B' mean here?
Student 3: When the line hits the 'Y'.
Teacher: Yeah. And what do we call that?
Student 3: The Y-intercept.

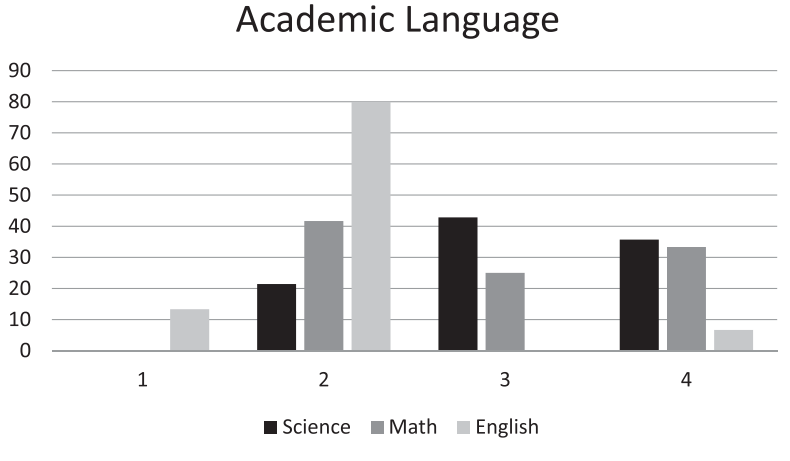


Figure 1. PLATO scores for Academic Language in science, mathematics, and English.

Table 3. Percentage of segments showing opportunities to listen, speak, read, and write across the subjects: high-end PLATO scores (3–4).

Subject	Listening and speaking	Reading	Writing
Science	86%	0%	14%
Mathematics	83%	0%	0%
English	60%	20%	0%

Note: Listening and speaking = Classroom Discourse (CD), sub-category Opportunities. Reading = Text-Based Instruction (TBI), sub-category 'Use.' Writing = Text-Based Instruction (TBI), sub-category 'Production' (See Table 3). Each segment can include any aspect of speaking, reading, and writing. Therefore, each category can score up to 100%.

Here, the mathematics teacher integrates academic language and content by prompting the students to convey subject-specific terminology, with Students 1 and 2 negotiating meaning in L2. Student 3 opts to explain in L1, with the teacher prompting him to use L2.

Supportive materials (ALL)

We also investigated the type of supportive material the teachers provided, and found several instances of visual representations to aid content learning in the CLIL subjects. Again using the L2 subject as baseline, we infrequently found materials offered as language support in the English lesson (20% high-end), quite similarly to mathematics (25%). The science teaching used extensively more supportive materials (79% high-end), but both CLIL teachers used these materials to integrate content and language. This included science props (e.g. vegetables) for the students to name, pictures with labels so students could identify certain items, and models of items they were building in the laboratory. An example of supportive materials in mathematics was a drawing of a box with terminology and descriptions, which the students actively used to construct their own boxes and negotiate content meaning during the process.

Opportunities to listen and speak (CD), read and write (TBI)

A third characteristic of language integration in the CLIL lessons concerned opportunities to use L2. We found a striking similarity in the students' opportunities to listen and speak through classroom discourse across both CLIL subjects (86% high-end in science and 83% in mathematics), even to a larger extent than the L2 subject (60%). This suggests that these opportunities are characteristic of CLIL teaching, rather than subject-dependent (Table 3). One reason is the large amount of group work in the CLIL subjects, whereas the L2 subject had more individual work.

In addition, we identified the CLIL teachers' consistent uptake of student responses and ideas in classroom discourse, within and across the CLIL lessons. In mathematics and science, the teachers prompted the students to justify their answers during most lessons (science 57% high-end; mathematics 67%). Excerpt 3 shows the mathematics teacher's engagement in high-level uptake, contributing to students' opportunities to negotiate content meaning in L2.

Excerpt 3 (Mathematics, Uptake, Score 4):

Teacher: At what values of x would the volume be zero? So, at zero, here. How could this parenthesis here be zero? Or this parenthesis here be zero.
Student 2: Ok. […]
Teacher: Yeah, but how could … could you have any other values of x that also would be zero? Other than zero. Because the question asks where it intersects the x-axis
Student 2: One hundred!
Teacher: One hundred? So, thirty minus two times one hundred. That's …
Student 1: Explain *hvor kommer* [where comes] …
Student 2: It's x minus. You can have a minus box.
Teacher: Yeah, that's true. But the volume would be less than zero.
Student 2: So, then it's before. Before one hundred x.
Teacher: Could you … , could you turn this parenthesis … could you give us an x-value that would make this parenthesis to be zero?

Student 3: Uh. Fifteen?
Teacher: Yep. Because thirty minus two times fifteen is zero. So, then the whole thing would be zero.
Student 1: Oh.
Teacher: Ok. Could you find a value of x here that would make this parenthesis, uh, to zero?

Here, the mathematics teacher asks for elaboration, addresses student ideas, and challenges students to expand on these. These aspects point to L2 integration by high uptake of student responses, and are implemented consistently throughout the mathematics teaching.

A final note concerns the (lack of) CLIL characteristics related to reading and writing. As shown in Table 3, reading and writing had largely perfunctory functions across the subjects, meaning the students read (in the L2 subject only) and wrote (in science only) primarily to solve tasks during group work. Although texts were present during most CLIL lessons, all tasks and materials provided to the students were authored by the teachers. Reading was not addressed in depth. Concerning writing, they largely took notes in the mathematics class. However, in science, the teacher discussed how to write a lab report, providing opportunities for sustained writing. Thus, while listening and speaking seem characteristic both within and across the two CLIL subjects, the opportunities to write rather seems a characteristic of science literacy.

Content-driven

We identified three content-driven features of the CLIL teaching, which entailed how the teachers expressed the purpose of the lessons (PUR), what type of instructional explanations the teachers provided (ROC), and the intellectual challenges (IC) of the tasks and activities the students were provided to fulfill said purpose.

Purpose (PUR)

While the CLIL teachers expressed no language learning goals, content-driven purposes were explicitly stated in both CLIL subjects, as here in Excerpt 4 where the science teacher expresses the goal of a science experiment:

Excerpt 4 (Science, Purpose, Score 4):

Teacher: We're going to be doing a 'red cabbage indicator.' Right, so we're going to be testing solutions for whether or not they are acids or bases. We're going to see what kind of … what kind of effect the, um, the indicator has. Ok? [Writes: Goals for lab -> finding the pH of different substances using universal indicators] Ok, I'm going to give you your lab [assignment]. Please read it now, for five minutes. Then I want you to figure out what are the goals for the lab. Ok? [Students read assignment] What is the goal, what is the purpose of the lab? Goals? Or goals? [Student], you had an idea?
Student: To figure out the pH values of different things by using different indicators.
Teacher: Yeah? So. Finding the pH of different substances. Right? Using … [student], yeah?
Student: But also, do you think we will be able to find the properties of acids and bases?
Teacher: Hmm, good question. I don't know if you can do that. If you think about it, what are you going to see? What are you going to observe?
Student: We are going to observe whether the substance or object or whatever we are testing is acidic or …
Teacher: Yeah, but what are the properties? We're not going to exactly be looking at the properties. No. So, this is all about seeing … learning about the pH scale, right? Um. The pH of different … and then, and you said, the pH scale, so we need to know about what this means.

This example indicates how the goal relates to the development of students' science skills. The teacher prompts students to state the purpose of the lesson, and there is evidence of student awareness. Throughout the segment, the activities aligned with this goal, and toward the end of the segment, the teacher referred back to it. This example is representative of the high-end, content-driven purposes observed in 64% of the science segments. In mathematics, the goal was explicitly communicated in only 17% of the segments, and also content-driven.

Rich explanations (ROC)

Another characteristic of the content-driven teaching we observed concerned instructional explanations and conceptual richness. The explanations of content in both CLIL subjects were rich, lengthy, and accurate (science 71% high-end, mathematics 58%), suggesting integration of content and language as the CLIL teachers conveyed the content of their subjects in L2. A difference though, was the science teacher's focus on conceptual understanding in most of the lessons (79% high-end), whereas the mathematics teacher did so in a few segments only (17% high-end), primarily focusing on explaining rules and procedures. This points to more conceptual richness in the science teaching.

The science teacher focused on representing scientific phenomena and providing clear examples, analogies, and explanations, as in this example where she explains the theory behind the pH-experiment they are going to conduct:

Excerpt 5 (Science, Instructional Explanations and Conceptual Richness, Score 4):

Teacher: What is a pH scale? What is the range on it? What do you think, [student]? The range of the pH scale.
Student: From zero to … fourteen?
Teacher: Yeah. Do you think in this world that we have just that, the … that it's just between … hmm … zero and fourteen? The pH? H stands for? What do you think? So, it's between zero and fourteen, ok, you say that. This one is between zero and fourteen. What I was going to say was that the … there is more. Beyond fourteen and beyond zero. There are substances that are minus twenty-five. But in general, this is our scale. Just know that it is not limited. What does H stand for?
Student: Hydrogen.
Teacher: Good. So, this is hydrogen. What do you think 'p' stands for? They don't really know, but they assume that it's called the 'power of hydrogen.' How cool is that? It's like a super power of hydrogen. And then you had these things … what is this power of hydrogen? What does it do? Well, it creates different environments. Acidic or basic. Right? So, the hydrogen in it is … it's the hydrogen's fault! The hydrogen is at fault for creating acidity or … basic … basidity as well, I've seen. Ok, so, finding the pH of different substances using a universal indicator, and also testing different indicators.

The science teacher's L2 explanation of the pH scale is long, accurate, and clear, which she ties to an earlier lesson where they talked about hydrogen, and explains how to use the theory during the experiment. The primary focus concerns a conceptual understanding of the pH scale. Although the mathematics teacher rarely explained concepts in depth, the L2 explanations were mostly accurate and addressed student misunderstandings, as in Excerpt 6:

Excerpt 6 (Mathematics, Instructional Explanations, Score 3, Conceptual Richness, Score 2):

Teacher: So, these two [graphs] show the difference for mobile phone subscription. So, this one is more expensive, right?
Student: Yeah?
Teacher: And this one is less expensive, unless you use many megabytes per month. So, if you use any megabyte, this one increases. But this won't increase that much. Since the slope is different. So here, this mobile phone subscription will be more expensive. If you use a lot of megabytes. So, in this point, this one will start to be more expensive when you use more than how many megabytes? Fifty. So, if you use more than fifty megabytes, this subscription will be more expensive. But if you use less than fifty megabytes, this one will be more expensive. That's the answer to this one. So, when it says cheapest to use the subscription, it's f of x. F of x is the first one, right? That's this one.

The mathematics teacher is explaining how to interpret and compare two graphs. The example is accurate and clear to help the students solve the task at hand, although there is no focus on conceptual understanding of the graphs. Both CLIL teachers' representations of content in L2 were not only rich, but also consistently focused on subject-specific content.

Intellectually challenging

The third content-driven characteristic of the CLIL teaching was that it provided high intellectual challenges in terms of analytic/inferential tasks in more than half of both the science segments (57% high-end), and the mathematics segments (58% high-end). Figure 2 gives an overview.

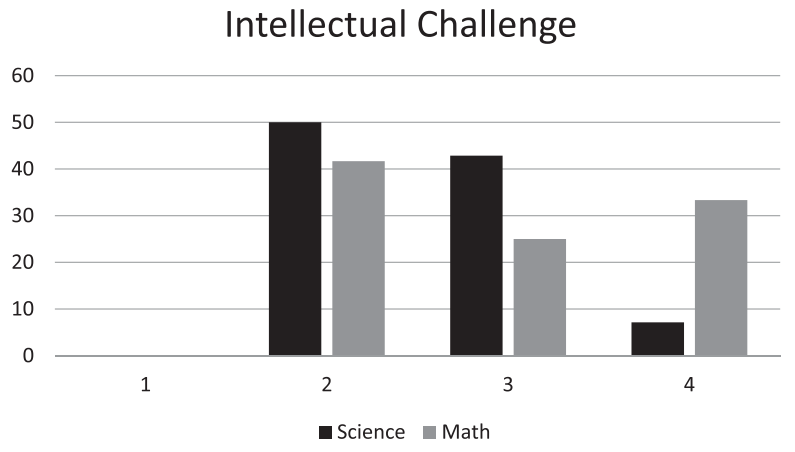


Figure 2. PLATO scores for Intellectual Challenge in science, mathematics, and English.

Both CLIL subjects showed a mixture of rote/recall (score 2) and analytic/inference activities (score 3–4). In both subjects, the teachers encouraged the students to think like scientists and mathematicians; formulating hypotheses, observing, testing, and justifying conclusions in science, and interpreting visual representations of numbers in mathematics. Excerpt 7 offers an example:

Excerpt 7 (Mathematics, Intellectual Challenge, Score 3):

Teacher: So, when the graph intersects the x-axis, what will the value of y be then? When it intersects here? For example, if it intersects about here? What is the value of y?
Student 1: Zero.
Teacher: Zero! If it intersects here? What will the value of y be?
Student 1: Fourteen point five?
Student 2: What? No.
Teacher: Zero. Because it's always zero. On this axis. Where it intersects the x-axis, the y will always be zero. Because here, the y is always zero. So, at this point, the coordinate will be twelve point zero. Eleven point zero, ten point zero, nine point zero, eight point zero, seven point zero, zero point zero. And what happen … what do you think the volume is? Here?
Student 1: Zero.
Teacher: Zero.

The first question was inferential, asking the students in L2 to infer based on the axes. When Student 1 answered correctly, the teacher checked by asking a related question. When Student 1 then answered incorrectly, the teacher reformulated it to prompt the students to approach the question differently, instead of reducing the intellectual challenge of the task. However, a lack of student responses to analytic/inference questions reduced the level of intellectual challenge in some instances.

Discussion

In this section, we return to our research question: *What characterizes CLIL teaching in science and mathematics in terms of content and language when taught in English as a second language?* We will discuss our findings in light of previous CLIL research and our understanding of integration. One challenge is to discuss potential benefits or pitfalls when using L2 teaching as a baseline for comparison with the use of L2 in CLIL teaching. Another is whether the content-driven features of the science and mathematics teaching are due to CLIL teaching or are common content features for the subjects.

Although the differences between CLIL subjects are ultimately more interesting and significant for understanding CLIL than the differences between them and English L2 teaching, using the latter as a

baseline contributes to our understanding of the integration of language and content. Specifically, it helps identify CLIL characteristics concerning content learning through L2. Characterizing the oral use of languages of the CLIL classroom, this class spoke primarily in English L2, both in science and mathematics. Using L2 for 83 –97– of the time is extremely high, particularly compared to prior research in Norway with 30% or more L2 use (Brevik and Moe 2012). Language and content integration seemed to be a CLIL characteristic, as realized for example through the use of L1 to aid content understanding (especially translating subject-specific terminology). This concurs with previous research that suggests L1 is often used strategically by CLIL teachers (Gallagher and Colohan 2014; Gierlinger 2015; Martínez Adrián and Gutiérrez Mangado 2015; Tavares 2015).

Both CLIL subjects scored high on use of academic language, illustrating how the teachers used and prompted subject-specific terminology consistently throughout the lessons. This points to process-oriented integration, where the students are becoming members of the subject communities (Sfard 1998). Through prompting them to use terminology, the CLIL teachers push their students toward being able to think and speak scientifically/mathematically in L2. We argue that this is the embodiment of the added value of CLIL – students are not only learning to express themselves in L2, but through integration express themselves in specific disciplinary ways (Berger 2016; Llinares et al. 2012; Nikula et al. 2016a, 2016b; Norris and Phillips 2003; Ødegaard et al. 2014).

We furthermore revealed that another characteristic of the CLIL teaching was how it offered the students many opportunities to speak. This corroborates with theories of learning as social interaction (Mercer 2004; Vygotsky 1978) and previous research that aligns CLIL with a sociocultural framing (Dalton-Puffer 2007; Nikula 2010; Van Kampen et al. 2016). Moreover, this refers to speaking in ways that may also scaffold socialization into becoming a member of the field, in other words, not just a matter of speaking more for sake of speaking. Unsurprisingly, the opportunities to speak are also reflected in research on language outcomes that suggests CLIL students primarily improve their oral proficiency (Admiraal, Westhoff, and De Bot 2006; Lasagabaster 2008). However, the CLIL students were provided limited opportunities to read and write. Interestingly, this is echoed in CLIL literature commenting that writing was perceived as a homework activity (Dalton-Puffer 2007). We therefore question if the overt focus on oral communication may draw attention away from reading and writing aspects of the content subjects, particularly as these are considered essential features of scientific/mathematical literacy (Berger 2016; Ødegaard et al. 2014). Other studies of science classrooms show between 17% and 30% writing (Ødegaard et al. 2014; Ødegaard and Arnesen 2010), where Ødegaard et al. (2014) was an intervention study with a focus on literacy. In mathematics, it is usually around 50% (Bergem 2016).

The within-CLIL analysis further probed the depths of content-driven features of the CLIL teaching in science and mathematics, demonstrating more traces of the integration of content and language. Since both CLIL subjects purely focused on content goals and no tangible language goals, this raises another question of integration: Can CLIL be CLIL without explicit language goals, or is it sufficient to say that CLIL has a dual focus on content and language if there is an implicit focus on language learning? Several researchers have commented on this dichotomy (cf. Coyle et al. 2010; Dalton-Puffer 2007; Georgiou 2012; Marsh 2002). Nikula et al. (2016b) propose that our understanding of how language functions in content subjects is underdeveloped and needs to be further explored before we can begin to discuss how CLIL should be taught.

Delving into issues of content, there is also the question of whether content-driven features of the science and mathematics teaching reported above are due to teaching through an L2, or if these are subject-specific. Although Nikula (2010) argues that teaching through L2 may cause CLIL teachers to lose some nuances in their instructional explanations, we found the instruction to be rich, lengthy and accurate. In line with disciplinary literacy (Airey 2015; Ødegaard et al. 2014; Shanahan and Shanahan 2012), the science teacher focused on conceptual understanding of scientific phenomena, while the mathematics teacher primarily focused on mathematical rules and procedures. This is an interesting finding, which may relate to the culture of mathematics teaching in Norway (Stigler and Hiebert 1999). Another aspect of language support that might be considered a content rather

than CLIL characteristic, was visual aids. We discovered that the science lessons provided the students with more visual aids than the mathematics lessons; including models, and pictures with labels. We attribute this to the nature of the content subjects (Nikula et al. 2016b), since science traditionally use many visual representations (Tytler et al. 2013).

Much CLIL literature is preoccupied with intellectual challenge, questioning if learning a content subject through L2 will render the students less capable of completing tasks (Coyle et al. 2010; Gibbons 2015; Lin 2016). Examining intellectual challenge, we found that approximately half of the time in science, students were given analytical/high inference tasks, while in mathematics, they were given slightly more rote and recall tasks. The observation of tasks and dialogues between the teachers and students suggest that the students' levels of language and the type of challenges were successfully integrated. Briefly put, the students are still provided complex instructional explanations and intellectually challenging tasks through L2.

In conclusion, the strength with our within-CLIL comparison approach is that the participants are students in the same class. Studies that have compared CLIL students with non-CLIL students have been problematized, since CLIL students are often handpicked from disproportionately higher socio-economic backgrounds, have higher grade averages and L2 proficiency (see Aro and Mikkilä-Erdmann 2015; Bruton 2011, 2013). This makes for difficult comparisons, as CLIL teachers may teach their subjects differently to high-achieving CLIL students as opposed to non-CLIL students. However, a within-CLIL comparison means that we cannot say for certain if our findings are subject-specific (e.g. typical of science/mathematics) or CLIL-specific (teaching through L2). We acknowledge that our study only provides insight into the workings of one CLIL classroom. However, our design allowed for a systematic and detailed description of CLIL teaching across subjects and adding to the body of much-needed detailed studies of CLIL in practice. We hope these observations can serve as a starting point for further research, particularly into issues of how teachers support their students through scaffolding not only language, but also content.

Acknowledgements

The authors would like to thank Professor Kirsti Klette for inviting this study into the Linking Instruction and Student Experiences (LISE) project at the Department of Teacher Education and School Research, at the University of Oslo. We would also like to thank Roar Bakken Stovner (PLATO coder in mathematics), and Bjørn Sverre Gulheim (technician) for their invaluable input in the data analysis process.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID


Karina Rose Mahan http://orcid.org/0000-0002-7083-1171

References

Admiraal, Wilfried, Gerard Westhoff, and Kees De Bot. 2006. "Evaluation of Bilingual Secondary Education in the Netherlands: Students' Language Proficiency in English." *Educational Research and Evaluation* 12 (1): 75–93. doi:10.1080/13803610500392160.

Airey, John. 2015. "From Stimulated Recall to Disciplinary Literacy: Summarizing Ten Years of Research into Teaching and Learning in English." In *English-Medium Instruction in European Higher Education*, edited by Dimova Slobodanka, Anna Kristina Hultgren, and Christiansen Jensen, 157–176. Boston/Berlin: De Gruyter Mouton.

Airey, John. 2015. "From Stimulated Recall to Disciplinary Literacy: Summarizing Ten Years of Research Into Teaching and Learning in English." In *English-Medium Instruction in European Higher Education*, edited by Slobodanka Dimova, Anna Kristina Hultgren, and Christian Jensen, 157–176. De Gruyter Mouton.

Aro, Sophie, and Mirjamaija Mikkilä-Erdmann. 2015. "School-External Factors in Finnish Content and Language Integrated Learning (CLIL) Programs." *Scandinavian Journal of Educational Research* 59 (2): 127–142. doi:10.1080/00313831.2014.894937.

Ball, Phil, Keith Kelly, and John Clegg. 2016. *Putting CLIL Into Practice*. Oxford: Oxford University Press.

Banegas, Darío Luis. 2016. "Integrating Content and Language in English Language Teaching in Secondary Education: Models, Benefits, and Challenges." *Studies in Second Language Learning and Teaching* 2 (1): 111. doi:10.14746/ssllt.2012.2.1.6.

Barwell, Richard. 2005. "Integrating Language and Content: Issues from the Mathematics Classroom." *Linguistics and Education* 16 (2): 205–218.

Bergem, O. K. 2016. "Usually We Are Not Where the Teacher Is." In *Teaching and Learning in Lower Secondary Schools in the Era of PISA and TIMMS*, 47–61. Cham: Springer.

Berger, Angela. 2016. "Learning Mathematics Bilingually: An Integrated Language and Mathematics Model (ILMM) of Word Problem-Solving Processes in English as a Foreign Language." In *Conceptualising Integration in CLIL and Multilingual Education*, edited by Tarja Nikula, Emma Dafouz, Pat Moore, and Ute Smit, 73–100. UK, USA & Canada: Multilingual Matters.

Blikstad-Balas, Marte. 2016. "Key Challenges of Using Video When Investigating Social Practices in Education: Contextualization, Magnification, and Representation." *International Journal of Research & Method in Education* 40 (5): 511–523.

Brevik, Lisbeth M., and Eli Moe. 2012. "Effects of CLIL Teaching on Language Outcomes." In *Collaboration in Language Testing and Assessment*, edited by Dina Tsagari and Ildikó Csépes, 213–227. Frankfurt, Berlin, Bern, Brussels, New York, Oxford & Wien: Peter Lang AG.

Bruton, Anthony. 2011. "Is CLIL so Beneficial, or Just Selective? Re-Evaluating Some of the Research." *System: An International Journal of Educational Technology and Applied Linguistics* 39 (4): 523–532. doi:10.1016/j.system.2011.08.002.

Bruton, Anthony. 2013. "CLIL: Some of the Reasons Why … And Why Not." *System* 41 (3): 587–597. doi:10.1016/j.system.2013.07.001.

Cenoz, Jasone, Fred Genesee, and Durk Gorter. 2014. "Critical Analysis of CLIL: Taking Stock and Looking Forward." *Applied Linguistics* 35 (3): 243–262.

Cohen, Julia Jackson, Pamela Lynn Grossman, Hilda Borko, Susanna Loeb, and Richard Shavelson. 2013. *Practices That Cross Disciplines?: A Closer Look at Instruction in Elementary Math and English Language Arts*. Stanford: Stanford University.

Cohen, J., L. C. Schuldt, L. Brown, and P. Grossman. 2016. "Leveraging Observation Tools for Instructional Improvement: Exploring Variability in Uptake of Ambitious Instructional Practices." *Teachers College Record* 118 (110306): 1–37.

Coyle, Do. 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, Do, Philip Hood, and David Marsh. 2010. *Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Creswell, John W. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd ed. Los Angeles: Sage.

Dalton-Puffer, Christiane. 2007. *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: John Benjamins Publishing.

Dalton-Puffer, Christiane. 2011. "Content-and-Language Integrated Learning: From Practice to Principles?" *Annual Review of Applied Linguistics* 31: 182–204.

Dalton-Puffer, Christiane. 2016. "Cognitive Discourse Functions: Specifying an Integrative Interdisciplinary Construct." In *Conceptualising Integration in CLIL and Multilingual Education*, edited by Tarja Nikula, Emma Dafouz, Pat Moore, and Ute Smit, 29–54. UK, USA & Canada: Multilingual Matters.

Dalton-Puffer, Christiane, Tarja Nikula, and Ute Smit. 2010. *Language Use and Language Learning in CLIL Classrooms*. Amsterdam: John Benjamins Publishing.

Davison, Chris. 2005. "Learning Your Lines: Negotiating Language and Content in Subject English." *Linguistics and Education* 16 (2): 219–237. doi:10.1016/j.linged.2006.01.005.

De Graaff, Rick. 2016. "Foreword: Integrating Content and Language in Education: Best of Both Worlds?" In *Conceptualising Integration in CLIL and Multilingual Education*, edited by Tarja Nikula, Emma Dafouz, Pat Moore, and Ute Smit, xiii–xvi. UK, USA & Canada: Multilingual Matters.

De Graaff, Rick, Gerrit Koopman, Yulia Anikina, and Gerard Westhoff. 2007. "An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL)." *International Journal of Bilingual Education and Bilingualism* 10 (5): 603–624. doi:10.2167/beb462.0.

Escobar Urmeneta, Cristina. 2013. "Learning to Become a CLIL Teacher: Teaching, Reflection and Professional Development." *International Journal of Bilingual Education and Bilingualism* 16 (3): 334–353. doi:10.1080/13670050.2013.777389.

Evnitskaya, Natalia, and Tom Morton. 2011. "Knowledge Construction, Meaning-Making and Interaction in CLIL Science Classroom Communities of Practice." *Language and Education* 25 (2): 109–127.

Fernández-Sanjurjo, J., A. Fernández-Costales, and J. M. Arias Blanco. 2017. "Analysing Students' Content-Learning in Science in CLIL vs. Non-CLIL Programmes: Empirical Evidence from Spain." *International Journal of Bilingual Education and Bilingualism* 1: 1–14. doi:10.1080/13670050.2017.1294142.

Gajo, Laurent. 2007. "Linguistic Knowledge and Subject Knowledge: How Does Bilingualism Contribute to Subject Development?" *International Journal of Bilingual Education and Bilingualism* 10 (5): 563–581. doi:10.2167/beb460.0.

Gajo, Laurent, and Cecilia Serra. 2002. "Bilingual Teaching: Connecting Language and Concepts in Mathematics." In *Education and Society in Plurilingual Context*, edited by D. W. C. So and G. M. Jones, 563–581. Brussels: Brussels University Press.

Gallagher, Fiona, and Gerry Colohan. 2014. "T(w)o and Fro: Using the L1 as a Language Teaching Tool in the CLIL Classroom." *The Language Learning Journal*: 1–14. doi:10.1080/09571736.2014.947382.

Genesee, Fred. 1994. *Integrating Language and Content: Lessons from Immersion*. Santa Cruz, CA: National Center for Research on Cultural Diversity and Second Language Learning.

Georgiou, Sophie Ioannou. 2012. "Reviewing the Puzzle of CLIL." *ELT Journal* 66 (4): 495–504.

Gibbons, Pauline. 2015. *Scaffolding Language, Scaffolding Learning. Teaching English Language Learners in the Mainstream Classroom*. New Hampshire: Heinemann.

Gierlinger, Erwin. 2015. "'You Can Speak German, Sir': On the Complexity of Teachers' L1 Use in CLIL." *Language and Education*: 1–22. doi:10.1080/09500782.2015.1023733.

Grandinetti, Maria, Margherita Langellotti, and Y. L. Teresa Ting. 2013. "How CLIL can Provide a Pragmatic Means to Renovate Science Education – Even in a Sub-Optimally Bilingual Context." *International Journal of Bilingual Education and Bilingualism* 16 (3): 354–374. doi:10.1080/13670050.2013.777390.

Grossman, P. L. 2015. *Protocol for Language Arts Teaching Observations (PLATO 5.0), Center to Support Excellence in Teaching (CSET)*. Palo Alto: Stanford University.

Grossman, Pamela Lynn, Susanna Loeb, Julia Jackson Cohen, Karen Hammerness, and James Wyckoff. 2010. "Measure for Measure: The Relationship Between Measures of Instructional Practice in Middle School English Language Arts and Teachers' Value-Added." *American Journal of Education* 119 (3): 445–470.

Harvey, Tiina, Tiina Tihinen, Kaarina Määttä, and Satu Uusiautti. 2013. "Content and Language Integrated Learning in Practice: Comparison of Three Cases." *International Journal of Academic Research in Progressive Education and Development* 2 (2): 86–98.

Hellekjær, Glenn Ole. 2005. "The Acid Test: Does Upper Secondary EFL Instruction Effectively Prepare Norwegian Students for the Reading of English Textbooks at Colleges and Universities?" *Unpublished Doctoral Thesis*., University of Oslo.

Hickmann, M. E. 1985. "The Implications of Discourse Skills in Vygotsky's Developmental Theories." In *Culture, Communication and Cognition*, edited by James V. Wertsch, 236–257. Cambridge: Cambridge University Press.

Johnson, Burke, and Larry B. Christensen. 2014. *Educational Research. Quantitative, Qualitative and Mixed Approaches*. 5th ed. USA: Sage.

Kane, T. J., and D. O. Staiger. 2012. *Gathering Feedback for Teaching*. Seattle, WA: Mill and Melinda Gates Foundation.

Klette, Kirsti, and Marte Blikstad-Balas. 2017. "Observation Manuals as Lenses to Classroom Teaching: Pitfalls and Possibilities." *European Educational Research Journal* 17: 1–18.

Klette, Kirsti, Marte Blikstad-Balas, and Astrid Roe. 2017. "LInking Instruction and Student Achievement. A Research Design for a New Generation of Classroom Studies." *Acta Didactica* 11 (3): 1–19.

Lasagabaster, David. 2008. "Foreign Language Competence in Content and Language Integrated Learning." *Open Applied Linguistics Journal* 1: 4–17.

Lasagabaster, David, and Yolanda Ruiz De Zarobe. 2010. *CLIL in Spain: Implementation, Results and Teacher Training*. Cambridge: Cambridge Scholars Publishing.

Lin, A. 2016. *Language Across the Curriculum and CLIL in English as an Additional Language (EAL) Context. Theory and Practice*. Singapore: Springer.

Llinares, Ana. 2015. "Integration in CLIL: A Proposal to Inform Research and Successful Pedagogy." *Language, Culture and Curriculum* 28 (1): 58–73. doi:10.1080/07908318.2014.1000925.

Llinares, Ana, Tom Morton, and Rachel Whittaker. 2012. *The Roles of Language in CLIL*. Cambridge: Cambridge University Press.

Llinares, Ana, and Rachel Whittaker. 2007. "Talking and Writing in the Social Sciences in a Foreign Language: A Linguistic Analysis of Secondary School Learners of Geography and History." *Revista española de lingüística aplicada* 1: 83–94.

Llinares, Ana, and Rachel Whittaker. 2010. "Writing and Speaking the History Class: Data from CLIL and First Language Contexts." In *Language Use and Language Learning in CLIL Classrooms*, edited by Christiane Dalton-Puffer, Tarja Nikula, and Ute Smit, 125–144. Amsterdam: John Benjamins Publishing.

Lorenzo, Francisco. 2007. "An Analytical Framework of Language Integration in L2 Content-Based Courses: The European Dimension." *Language and Education* 21 (6): 502–514. doi:10.2167/le708.0.

Luoto, Jennifer, Kirsti Klette, and Marte Blikstad-Balas. Under Review. "Activity Formats and Instructional Patterns in Mathematics Classrooms: A Comparative Study of Norwegian and Finnish-Swedish Lower Secondary Classrooms."

Maasum, Tengku Nor Rizan Tengku Mohamad, Nooreiny Maarof, Hamidah Yamat, and Effandi Zakaria. 2012. "An Investigation of Teachers' Pedagogical Skills and Content Knowledge in a Content-Based Instruction Context." *Indonesian Journal of Applied Linguistics* 1 (2): 75–90.

Marsh, David. 2002. *CLIL/EMILE. The European Dimension. Actions, Trends, and Foresight Potential*. Jyväskylä: University of Jyväskylä.
Martínez Adrián, María, and María Juncal Gutiérrez Mangado. 2015. "L1 Use, Lexical Richness, Accuracy and Syntactic Complexity in the Oral Production of CLIL and NON-CLIL Learners of English." *Atlantis: Revista de la Asociación Española de Estudios Anglo-Norteamericanos* 37 (2): 175–197.
Mercer, Neil. 2004. "Sociocultural Discourse Analysis: Analysing Classroom Talk as a Social Mode of Thinking." *Journal of Applied Linguistics* 1 (2): 137–168.
Met, Myriam. 1999. *Content-Based Instruction: Defining Terms, Making Decisions*. Washington, DC: NFLC Reports.
Meyer, Oliver, and Do Coyle. 2017. "Pluriliteracies Teaching for Learning: Conceptualizing Progression for Deeper Learning in Literacies Development." *European Journal of Applied Linguistics* 5 (2): 199–222.
Meyer, Oliver, Do Coyle, Ana Halbach, Kevin Schuck, and Teresa Ting. 2015. "A Pluriliteracies Approach to Content and Language Integrated Learning – Mapping Learner Progressions in Knowledge Construction and Meaning-Making." *Language, Culture and Curriculum* 28 (1): 41–57. doi:10.1080/07908318.2014.1000924.
Miqdadi, Ruba, and Dina Al-Jamal. 2013. "Difficulties in Content and Language Integrated Learning: The Case of Math." *Jordan Journal of Educational Science* 9 (4): 449–459.
Moore, P. 2009. "On the Emergence of L2 Oracy in Bilingual Education. A Comparative Analysis of CLIL and Mainstream Learner Talk." *Unpublished Doctoral Thesis*., Universidad Pablo de Olavide.
Morton, Tom, and Ana Llinares. 2016. "Students' Use of Evaluative Language in L2 English to Talk and Write About History in a Bilingual Education Programme." *International Journal of Bilingual Education and Bilingualism*: 1–13. doi:10.1080/13670050.2016.1192101.
NESH. 2006. *Guidelines for Research Ethics in the Social Sciences, Law and the Humanities*. National Committees for Research Ethics in Norway. Accessed July 7, 2017.
Nikula, Tarja. 2005. "English as an Object and Tool of Study in Classrooms: Interactional Effects and Pragmatic Implications." *Linguistics and Education* 16 (1): 27–58. doi:10.1016/j.linged.2005.10.001.
Nikula, Tarja. 2010. "On Effects of CLIL on a Teacher's Language Use." In *Language Use and Language Learning in CLIL Classrooms*, edited by Christiane Dalton-Puffer and U. Smit, 105–124. Amsterdam: John Benjamins.
Nikula, Tarja. 2015. "Hands-On Tasks in CLIL Science Classrooms as Sites for Subject-Specific Language Use and Learning." *System* 54: 14–27. doi:10.1016/j.system.2015.04.003.
Nikula, Tarja, Emma Dafouz, Pat Moore, and Ute Smit. 2016a. *Conceptualising Integration in CLIL and Multilingual Education*. UK, USA & Canada: Multilingual Matters.
Nikula, Tarja, Christiane Dalton-Puffer, Ana Llinares, and Francisco Lorenzo. 2016b. "More Than Content and Language: The Complexity of Integration in CLIL and Bilingual Education." In *Conceptualising Integration in CLIL and Multilingual Education*, edited by Tarja Nikula, Emma Dafouz, Pat Moore, and Ute Smit, 1–25. UK, USA & Canada: Multilingual Matters.
Norris, S. P., and L. M. Phillips. 2003. "How Literacy in its Fundamental Sense is Central to Scientific Literacy." *Science Education* 87 (2): 224–240.
Ødegaard, Marianne, and Nina Elisabeth Arnesen. 2010. "Hva skjer i naturfagsklasserommet? Resultater fra en videobasert klasseromsstudie, PISA +." *NorDiNa: Nordic Studies in Science Education* 6 (1): 16–32.
Ødegaard, Marianne, B. S. Haug, S. M. Mork, and G. O. Sørvik. 2014. "Challenges and Support When Teaching Science Through an Integrated Inquiry and Literacy Approach." *International Journal of Science Education* 36 (18): 2997–3020.
Peräkylä, Anssi. 2011. "Validity in Research on Naturally Occurring Social Interaction." In *Qualitative Research*, edited by David Silverman, 365–382. London: Sage.
Pérez-Cañado, María Luisa. 2016. "Are Teachers Ready for CLIL? Evidence from a European Study." *European Journal of Teacher Education* 39 (2): 202–221. doi:10.1080/02619768.2016.1138104.
Pérez-Vidal, Carmen, and Helena Roquet. 2015. "The Linguistic Impact of a CLIL Science Programme: An Analysis Measuring Relative Gains." *System* 54: 80–90. doi:10.1016/j.system.2015.05.004.
Pimm, D. 1987. *Speaking Mathematically: Communication in Mathematics Classrooms*. London: Routledge & Keegan Paul.
Relaño Pastor, Ana. 2015. "The Commodification of English in 'Madrid, Comunidad Bilingüe': Insights from the CLIL Classroom." *Language Policy* 14 (2): 131–152. doi:10.1007/s10993-014-9338-7.
Ruiz de Zarobe, Yolanda, and Jasone Cenoz. 2015. "Way Forward in the Twenty-First Century in Content-Based Instruction: Moving Towards Integration." *Language, Culture and Curriculum* 28 (1): 90–96. doi:10.1080/07908318.2014.1000927.
Sfard, Anna. 1998. "On Two Metaphors for Learning and the Dangers of Choosing Just One." *Educational Researcher* 27 (2): 4–13. doi:10.3102/0013189X027002004.
Shanahan, Timothy, and Cynthia Shanahan. 2012. "What Is Disciplinary Literacy and Why Does It Matter?" *Topics in Language Disorders* 32 (1): 7–18.
Snell, Julia. 2011. "Interrogating Video Data: Systematic Quantitative Analysis Versus Micro-Ethnographic Analysis." *International Journal of Social Research Methdology* 14: 253–258.
Stigler, J., and J. Hiebert. 1999. *The Teaching Gap*. New York: The Free Press.
Stovner, Roar Bakken. 2018. "Do Mathematics-Specific and General Observation Instruments Capture Different Aspects of Instruction?" *Conference proceedings of the American education research association (AERA) annual meeting*, New York.

Šulista, Marek. 2012. *CLIL Implementation in Mathematics Lessons*. Prague: Lambert Academic Publishing.

Svenhard, Britt Wenche, Kim Servant, Glenn Ole Hellekjær, and Henrik Bøhn. 2007. "Norway." In *Windows on CLIL*, edited by Anne Maljers, David Marsh, and Dieter Wolff, 139–146. The Hague: European Platform for Dutch Education.

Tavares, Nicole Judith. 2015. "How Strategic Use of L1 in an L2-Medium Mathematics Classroom Facilitates L2 Interaction and Comprehension." *International Journal of Bilingual Education and Bilingualism* 18 (3): 319–335.

Tytler, Russell, Vaughan Prain, Peter Hubber, and Bruce Waldrop. 2013. *Constructing Representations to Learn in Science*. Amsterdam: Springer Science and Business Media.

Van Kampen, Evelyn, Wilfried Admiraal, and Amanda Berry. 2016. "Content and Language Integrated Learning in the Netherlands: Teachers' Self-Reported Pedagogical Practices." *International Journal of Bilingual Education and Bilingualism* 1 (15): 222–236.

Van Rinsveld, Amandine, Christine Schiltz, Martin Brunner, Karin Landerl, and Sonja Ugen. 2016. "Solving Arithmetic Problems in First and Second Language: Does the Language Context Matter?" *Learning and Instruction* 42: 72–82. doi:10.1016/j.learninstruc.2016.01.003.

Vollmer, H. J. 2008. "Constructing Tasks for Content and Language Integrated Assessment." In *Research on Task-Based Language Learning and Teaching. Theoretical, Methodological and Pedagogical Perspectives*, edited by J. Eckhert and S. Siekmann, 227–290. Franfurt: Peter Lang.

Vygotsky, Lev Semenovič. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wertsch, James V. 1985. *Culture, Communication, and Cognition: Vygotskian Perspectives*. Cambridge: Cambridge University Press.

Article II

Mahan, K. R. (2020) The comprehending teacher: scaffolding in content and language integrated learning (CLIL), *The Language Learning Journal*, https://doi.org/10.1080/09571736.2019.1705879

The Comprehending Teacher: Scaffolding in Content and Language Integrated Learning (CLIL)

Journal:	*The Language Learning Journal*
Manuscript ID	Draft
Manuscript Type:	Original Paper
Keywords:	Content And Language Integrated Learning (CLIL), English as a Foreign Language, Scaffolding, Classroom-based research, Teacher Education

The Comprehending Teacher: Scaffolding in Content and Language Integrated Learning (CLIL)

Abstract

Teaching through a second language (L2) poses many challenges, as second language learners (SLLs) have fewer linguistic resources in the language of instruction. This study investigates how three Norwegian Content and Language Integrated Learning (CLIL) teachers support learning for second language learners (SLL) through scaffolding. Twelve lessons (science, geography and social science) were filmed in one 11th grade CLIL class. A coding manual (PLATO) was used to identify the scaffolding strategies the teachers used. The findings indicate that CLIL teachers scaffold their students to comprehend material. However, they provide few metacognitive strategies to help students solve tasks. CLIL teachers scaffold differently in the natural and social sciences; the natural science teaching has more visual aids, whereas the social science teaching allows for more student talk. The results imply that natural and social science teacher complement each other. However, CLIL teachers need to create more specific learning activities to provide their students with more support.

Keywords: Content and Language Integrated Learning (CLIL), English as a Foreign Language (EFL), Scaffolding, Classroom-Based Research, Teacher Education

Introduction

This study investigates how teachers use scaffolding strategies to support students learning English L2 in a content and language integrated learning (CLIL) classroom. CLIL is a bilingual teaching approach defined as an additional language integrated into a non-language subject (Coyle, Hood, and Marsh 2010, 1). CLIL students have greater difficulties learning material than L1 students because they learn material at the same level as L1 students but with larger language deficits in the language of instruction (Cummins and Early 2015). CLIL teachers are generally untrained in teaching second language learners (SLLs), and they express concerns about how to teach them (Pérez-Cañado 2016). SLL researchers claim that scaffolding is a promising way to help SLLs (Gibbons 2015; van de Pol, Volman, and Beishuizen 2010). By using scaffolding strategies, CLIL teachers can integrate language learning into content subjects (Pawan 2008), thus exploring meaning negotiation and linguistic assistance in the classroom. This is crucial to the language development of SLLs (Kayi-Aydar 2013). However, even though many SLL researchers note the potential benefits of scaffolding to SLLs, the research on CLIL is disparate and limited (Author 1, Colleague 1, and Colleague 2 2018). There is a need for empirically-grounded studies on naturally occurring CLIL teaching in order to map out how content teachers scaffold. The current study uses a coding manual to identify scaffolding in video-recorded classroom interaction in a CLIL classroom in which science, geography, and social science is taught. The main unit of analysis is the interaction between the teacher and the students. The study contributes to mapping what the teachers do and do not do to scaffold students, and the results may be used to further discuss how CLIL teachers may more effectively support their students in their learning processes. This study is guided by the following research question: **Which scaffolding strategies do three CLIL teachers use to help their L2 English students comprehend material and complete tasks?**

Scaffolding in theory and practice

Bruner introduced the term *scaffolding* in an educational sense in the 1970s. It refers to the 'interactional instructional relationship' between adults and learners that 'enables a child or novice to solve a problem [. . .] beyond his unassisted efforts' (Wood, Bruner, and Ross 1976, 90). Scaffolding has its roots in psychology but has since expanded into the educational sciences. Due to its flexible nature, scaffolding is a broad concept. Some researchers understand scaffolding as a new metaphor for Vygotsky's zone of proximal development, placing it firmly in sociocultural theory (Bliss, Askew, and Macrae 1996; Smagorinsky 2018; Verenikina 2004). Others insist on further developing it as a tool to use in the classroom, leaning toward more constructivist approaches (Hogan and Pressley 1997).

Researchers generally agree that the goal of scaffolding is student autonomy (van de Pol, Volman, and Beishuizen 2010), which is realized through tailored support from a teacher or more capable peer and involves the responsibility of learning slowly transferring from the teacher to the student (Lin et al. 2012). This study uses Maybin, Mercer, and Stierer (1992)'s definition of scaffolding: a type of teacher assistance that helps students learn new skills, concepts, or levels of understanding (hereafter **comprehension of material**) that leads to the student successfully **completing a task** ('a specific learning activity with finite goals') (Maybin, Mercer, and Stierer 1992, 188). These two aspects of scaffolding, as shown in Figure 1, are a basis for the framework for SLL scaffolding in the present study. They are realized through specific tools (scaffolding strategies) that are employed by the teacher (van de Pol, Volman, and Beishuizen 2010).

The above understandings of scaffolding have theoretical approaches. However, the field of SLL, which is the focus of this study, largely takes a practical approach to scaffolding by identifying what teachers do or should do (Echevarría, Vogt, and Short 2017; Gibbons 2015; Masako and Hiroko 2008). Scaffolding strategies operate from a macro level (e.g., curriculum planning that integrates language systematically) to a micro level (i.e., interactional scaffolding). Interactional scaffolding is

the minute-to-minute support teachers give their students in the classroom (van Lier 2004, 148). Interactional scaffolding poses a challenge to teachers because they must support students with unpredicted problems on the spur of the moment (Many et al. 2009; Walqui 2006). The present study focuses on scaffolding strategies CLIL teachers use during interactional scaffolding. What follows is a review of SLL studies that explicitly regard scaffolding strategies within the framework of scaffolding strategies that this study uses (see Figure 1 below).

The majority of SLL scaffolding research is qualitative and descriptive and takes place in naturally occurring teaching (Lin et al. 2012). SLL researchers typically create their own frameworks in a bottom-up approach to identify scaffolding practices in the classroom (Gibbons 2003; Kayi-Aydar 2013; Li 2012). The main unit of analysis is the dialogue between teachers and students, although some studies include non-verbal behavior and gestures (Miller 2005). Most SLL studies use video observation and create coding schemes (e.g., Ajayi 2014; van de Pol and Elbers 2013). Researchers use vastly divergent conceptualizations, approaches, and terms—in other words, they measure disparate items. As van de Pol, Volman, and Beishuizen (2010) put it, 'the measurement and analysis of scaffolding still appears to be in its infancy' (287). To move forward, they suggest agreeing on a clear conceptualization of scaffolding and how to empirically operationalize and measure it.

Since there are many rich descriptions of scaffolding, the current study aims to research scaffolding in a top-down manner by building on existing literature to work toward a more unified understanding of scaffolding. The following section synthesizes SLL research in five emerging themes that researchers have used to describe how SLL teachers practice **interactional scaffolding** (Figure 1). The framework builds on literature primarily from **English language learner (ELL) contexts** (immigrant students in an English-speaking country) and **CLIL contexts** (students from the majority language learning English L2). ELLs and CLIL students represent two of the largest SLL groups and have the most research on scaffolding, although as discussed later, ELLs and CLIL students come from very different educational contexts. The literature review focuses on five scaffolding themes related to comprehending material and solving tasks, following Maybin, Mercer, and Stierer (1992)'s

classification of scaffolding (see Figure 1). This classification of scaffolding was used because it provides clear goals for scaffolding. The five emerging themes also correspond to the coding manual (Protocol for Language Arts Teaching Observation [PLATO]) used in the methods section. In what follows, each emerging theme will be discussed. The methods section will explain how PLATO empirically measures these emerging themes in the present study.

[FIGURE 1 NEAR HERE]

Comprehension strategies in SLL scaffolding

Scaffolding that aids comprehension emphasizes how to help students understand new material (Maybin, Mercer, and Stierer 1992). Pawan (2008) found that content teachers generally focus little on comprehension scaffolding strategies (as little as 28%). The first emerging theme to support comprehension draws on the **previous knowledge** of SLLs to introduce new material (Walqui 2006). In PLATO, this concept is known as 'connecting to prior knowledge' (Grossman 2015). It stems from the idea that SLLs are not 'empty vessels' but that they bring with them 'a collection of prior knowledge and skills acquired in their native language' (Dong 2017, 145). Linking known knowledge to unknown knowledge is pivotal, as prior knowledge is one of the most important factors in student learning (Tomlinson and Moon 2013, 421). Examples of comprehension strategies include assessing what students already know, referring to prior lessons, or using relatable real-world examples. Gallagher and Colohan (2014) argue that L1 can be a powerful scaffolding strategy in CLIL contexts (in which students and teachers have a common L1 and cultural background). Author 1, Colleague 1, and Colleague 2 (2018) and Dalton-Puffer (2007) have found that CLIL teachers frequently use L1 as a resource for helping students comprehend, drawing connections between concepts in L1 and L2.

The second emerging theme concerning comprehension is the role of **supportive materials** (Gibbons 2015; Walqui 2006). Supportive materials comprise visual aids, graphic organizers, use of body language, and other items to help students understand language in context (Grossman 2015). Academic language can be more difficult to acquire because one often cannot infer the meaning of

an academic word from context (Cummins 2013). Walqui (2006) asserts that SLLs therefore require rich extralinguistic contexts and supportive materials to 'construct their understanding on the basis of multiple clues and perspectives' (169). Boche and Henning (2015) describe how a teacher of eleventh- and twelfth-grade history used supportive materials to scaffold. By contextualizing texts with visual aids, sounds, and other ways of organizing information, the teacher helped students understand content. Likewise, Author 1, Colleague 1, and Colleague 2 (2018) investigated supportive materials in CLIL teaching. They found that CLIL science teaching involved visual aids, graphic organizers, and film clips that help students understand abstract concepts.

The third emerging theme is how to support SLL's **academic language development** so students can use correct terminology (cf. Meyer et al. 2015; Meyer and Coyle 2017; Morton 2015). Gibbons (2015) suggests that although academic language is also new to L1 students, they have a clear advantage because they have a solid foundation upon which to build. Scaffolding strategies include allowing students to use their own words to describe terminology; bilingual translations, and so forth (Barr, Eslami, and Joshi 2012). Ajayi (2014) found that Mexican-American ELLs learned vocabulary more efficiently when their English teacher employed explicit scaffolding strategies that targeted academic language. Researchers have found that vocabulary teaching can be implicit in CLIL contexts because those teachers are often not language teachers (Dalton-Puffer 2007; De Graaff et al. 2007). However, one Norwegian study revealed a ninth-grade English L2 CLIL math and science class in which the CLIL teachers used several scaffolding strategies to support academic language development (Author 1, Colleague 1, and Colleague 2 2018).

Task-solving strategies in SLL scaffolding

Task-solving strategies comprise scaffolding strategies aimed at helping students complete a specific learning activity (Maybin, Mercer, and Stierer 1992). Pawan (2008) has found that 70% of scaffolding (as self-reported by SLL content teachers) focuses on completing content-related tasks. The fourth

theme is how teachers use **discourse** as a supportive tool to help students with tasks (cf. Gibbons 2003; Kayi-Aydar 2013). According to McNeil (2011), key scaffolding strategies include *revoicing* (repeating the student's answer in academic language), *repetition* (echoing a student's answer in class), and *elaboration* (prompting the student to justify or lengthen their answer) (398). Author 1, Colleague 1, and Colleague 2 (2018) provide evidence of these three scaffolding strategies in CLIL teaching, but they found more strategies in mathematics than in science. The science discourse included more patterns of Initiation-Response-Evaluation (IRE). McNeil (2011), Dalton-Puffer (2007), and Banse et al. (2017) have investigated types of teacher questions in ELL/CLIL classrooms. They differentiate between *referential questions* (in which the teacher does not know the answer) and *display questions* (in which the teacher knows the answer) (definitions taken from Long and Sato 1983). All three studies conclude an overabundance of display questions. Referential questions are more relevant for language learning because they prompt students to form longer and more complex sentences (Farooq 2007). The overabundance of display questions, particularly in the natural sciences, indicates that students do not have many opportunities in which to speak or use L2 creatively (Banse et al. 2017; Lemke 1990; McNeil 2011).

The fifth and final emerging theme is **metacognition**, or 'learning to learn' (Coyle, Hood, and Marsh 2010, 29). This theme focuses on how teachers support students in completing tasks by making students aware of their own learning processes (Gaskins et al. 1997). Research suggests that one of the most effective ways of creating independent students is by showing them how to solve tasks (Gritter, Beers, and Knaus 2013; van de Pol, Volman, and Beishuizen 2010). In science teaching, metacognition has been emphasized in 72% of scaffolding frameworks (Pawan 2008). Scaffolding strategies that target metacognition include providing examples of tasks and discussing them (e.g., modeling) and suggesting meta-strategies to help students complete tasks. In CLIL contexts, only two studies have focused on metacognition. These studies were conducted in Basque Country on fifth- and sixth-grade English L2 science students (Ruiz de Zarobe and Cenoz 2015; Ruiz de Zarobe and

Zenotz 2017). The studies conclude that reading strategies have a moderate impact on reading comprehension and that they encourage the use of strategies in completing tasks.

In conclusion, many aspects of scaffolding have been examined in SLL classrooms, but very few studies have measured them similarly. Scaffolding is a more comprehensive field in ELL literature than in CLIL (Gibbons 2015; Walqui 2006). CLIL is just beginning to scrape the surface of scaffolding, and there is a need for more systematic, empirical research to describe how CLIL teachers scaffold their students (Author 1, Colleague 1, and Colleague 2 2018; Ruiz de Zarobe and Zenotz 2015). However, not all aspects of ELL scaffolding transfer to CLIL contexts. There are important differences between ELLs and CLIL students, such as socioeconomic differences and the status and use of L1 (Lasagabaster and Sierra 2009; Nikula and Mård-Miettinen 2014). The present study seeks to contribute to CLIL literature by mapping how content teachers scaffold when they and their students have a common L1. CLIL students represent a significant number of English SLLs in Europe (Cenoz, Genesee, and Gorter 2014). Examining how secondary CLIL teachers teach their content subjects in English can shed light on how they prepare students for English at the university level.

Methods

The present study is an analysis of twelve video-recorded lessons from one CLIL classroom in which three CLIL subjects (science, geography, and social science) were taught. It was filmed over the span of one month during the 2015–2016 school year. The video data were transcribed and coded with an observation manual (PLATO). The research design was developed and validated by the [research team, deleted for anonymity] at [university] (Colleague 3, Colleague 4, and 5 2017).

Sample

The sample was an eleventh-grade CLIL class at an upper secondary Norwegian school (ages 15–16). It was a convenience sample, as only 4–7% of upper secondary schools in Norway offer some form of CLIL teaching (Svenhard et al. 2007). The school offered an English CLIL program for science,

geography, and social science. Students apply for the CLIL program and are accepted based on their grades. The participants in this study were the science, geography, and social science teachers (n = 3) and the CLIL students (n = 25). All the teachers and students were female, and most had Norwegian as their L1. The CLIL teachers had 1–2 years of experience teaching CLIL, and two had attended CLIL courses. Three CLIL subjects were chosen for cross-comparison to see if the CLIL teachers scaffolded similarly to the same class regardless of subject (see Author 1, Colleague 1, and Colleague 2 2018).

Data collection and analysis

Video recordings were used for this study, as they allow researchers to systematically investigate complex educational settings (Snell 2011) and because they are useful in studying interactional scaffolding (van de Pol, Volman, and Beishuizen 2010). The CLIL classroom was filmed using two cameras: one small camera mounted in the back of the classroom, and one above the blackboard. The teacher wore a microphone; another was placed in the middle of the classroom to capture student utterances.

The video data were analyzed with PLATO, which is a teacher-centered observation manual that describes twelve aspects (here called 'elements') of teaching (Grossman et al. 2013). PLATO classifies elements on a scale from 1–4. Raters assign scores for every fifteen-minute segment of video data (approximately ten segments per subject in this study). A score of 1 or 2 signifies low-end teaching, and a score of 3 or 4 signifies high-end teaching. Low-end teaching indicates no evidence (score 1) to little evidence (score 2) of an element, whereas high-end teaching indicates limited evidence (score 3) to strong and consistent evidence (score 4). This study uses the percentage of segments that score within high-end teaching. For example, a score of 80% means that eight of the ten segments scored a 3 or 4. PLATO was chosen because it is a useful tool with which to identify, label, and measure teaching practices across subjects, and the scaffolding field calls for reliable and valid instruments of measurement (van de Pol, Volman, and Beishuizen 2010). Six of the PLATO elements correspond well with the emerging scaffolding themes (see Figure 1), allowing the

researcher to accurately score them. PLATO was originally created for language arts teaching but has been used to study science and mathematics teaching, and it takes SLLs into account (Author 1, Colleague 1, and Colleague 2 2018; Cohen 2018; Klette, Blikstad-Balas, and Roe 2017).

Six PLATO elements were selected to identify various scaffolding strategies in CLIL teaching. Three elements captured **comprehension scaffolding strategies** (*connections to prior knowledge, supportive materials,* and *academic language*). Three others captured **task-solving strategies** (*uptake of student responses, strategy use and instruction,* and *modeling and use of models*). Each element in the video data was identified, scored, and described. Table 1 defines each element and what constitutes each score. All definitions are taken from Grossman (2015).

[TABLE 1 NEAR HERE]

Research credibility and ethics

In accordance with the ethical guidelines of the Norwegian Center for Research Data, the teachers and students were informed orally and in writing about the project, and they each provided written consent (NESH 2006). A certified PLATO rater coded the video data. A second certified PLATO rater double-scored 25% of the video data to ensure reliability (interrater reliability = 90%). PLATO provided a useful lens with which to observe and interpret the aspects of scaffolding (Klette and Blikstad-Balas 2018). PLATO is supported by years of research on effective teaching, and using it will allow for comparison with other research that uses the same tool (Author 1, Colleague 1, and Colleague 2 2018; Klette, Blikstad-Balas, and Roe 2017; Klette and Blikstad-Balas 2018). However, using a manual with pre-determined codes may have limited the researcher's perception of scaffolding, and cannot measure the effect of the scaffolding strategies on the students' learning processes. The limited sample does not allow for generalizability.

Findings

The findings indicate that CLIL teachers employ an array of scaffolding strategies to help students comprehend material, but they employ limited strategies to help students complete tasks. The CLIL teachers frequently make connections between known and unknown material, provide the students with supportive materials, and consistently use, define, and prompt subject-specific terminology. The teachers consistently engage in dialogue that helps students solve tasks. However, there is limited evidence of strategies and models (metacognition).

Comprehension scaffolding strategies

Connections to prior knowledge (CPK)

The CLIL teachers consistently create connections between known and unknown material (high-end teaching, score 3–4, science 80%, geography 50%, social science 46%). Prominent scaffolding strategies include asking students if they are familiar with concepts, making explicit links to previous lessons, and using real life or personal examples.

The science teacher in particular refers to observable, scientific phenomena, e.g., she asks the students what happens when the students cut an apple in half. **The geography teacher** uses geographical land formations with which the students are familiar. **The social science teacher** prompts students to draw on everyday experiences to understand sociological phenomena, such as discussing how the students have resocialized from lower to upper secondary school.

In the following excerpt, the science teacher illustrates a redox reaction (new topic) by dropping a sink nail into copper chloride. She draws on the students' prior knowledge to guess what will happen, and why redox reactions are relevant for Norwegians:

Excerpt 1 (Science, Connections to Prior Knowledge, score 4):

Interaction	Action
Science teacher: What color does copper have when it's solid? Do you know? I didn't bring any copper out, but . . . is there any copper here? At least there is copper inside the cords, but I don't see any copper here. What kind of color is copper, Student 1?	Teacher elicits prior knowledge in students.
Student 1: It's like red-brown.	Student answers.
Science teacher: Yes. So, when it's solid, you know the different states, don't you? Solid, liquid, and gas. So, when copper is solid, it's a copper . . . sort of, you have a small chunk of copper, it would be some brownish red. Brownish red. When it's liquid, it's blue. Or light blue. So, it's mixed with chloride. I can't smell any . . . it doesn't smell like [local swimming pool]. But you will notice that later when we do an experiment. And this nail that I got. It's zinc. And now I'm going to put it in here. And do you have any idea if something will happen. Do you have any suggestions? Hypothesis? If anything will happen at all? It looks like sort of just blue-ish water. So, if I put zinc in here, do I expect . . . what do you expect? Student 2?	Teacher refers to the different states of elements, which she had covered earlier in class. Teacher refers to an element students are familiar with, relating it to a local place. Teacher prompts students to guess what will happen to the nail, presumably based on what they know about nails.
Student 2: Maybe it will start to rust?	Student provides suggestion based on prior experience.
Science teacher: Maybe it will start to rust. Why do you suggest that? Because that's very interesting. Because she suggests that corrosion will happen. That it will start to rust. And why did you suggest that?	Teacher asks why.
Student 2: Because I've seen it before?	Student confirms her belief based on prior knowledge.
Science teacher: You've seen it before! So, any other suggestions? Or do you stick with corrosion? So, let's see, then. So, I have an extra here, so you won't forget what it look like. It's a nail made out of zinc. [Drops it in a test tube with chloride]. Ok, so it turned black. [. . .] So, what is happening? It's a redox reaction. [. . .] So, this is, maybe not actually this, but this is sort of an introduction to a process that we use a lot	Teacher loops back to new material (redox reactions).

in Norway to create metals. Because how do you make metals? Well, you find some sort of chunk of the earth that you know contains a metal. But you don't want a chunk, you want only the metal. And then you can do something with electrons. You can sort of add or take away electrons to make the metal pure. And they do that a lot in Norway in Haugesund area, Karmøy, Vestlandet. And do you know why we do it there?	Teacher elicits prior knowledge of Norwegian industry and geography.
Student 3: Because there's a lot of water?	Student replies.

Here, the science teacher uses several tools to elicit and refer to prior knowledge. She creates a clear link between known material (what they know about copper) and how this is relevant to the unknown material (redox reactions). The segment scores a 4 because the new material builds explicitly on prior knowledge (see Table 1 for more information).

Supportive materials

There is a large difference among the CLIL teachers' use of supportive materials (science 60%, geography 90%, social science 18%). The most striking difference is the role of video clips and animations to show phenomena in the natural sciences. The **science** and **geography teacher** consistently use body language to illustrate the meanings of words. The science teacher uses Bohr models and the periodic table as aids for helping students understand the compositions of atoms. She shows a webpage that allows users to build atoms by adding and subtracting electrons and protons. The **geography teacher** uses instructional videos and pictures to illustrate geographical phenomena. The use of instructional videos allows students to see how land formations occur over time. She introduces a video clip with a song about erosion. She uses her hands and fingers to physically demonstrate the meaning of words, such as 'vertical' and 'horizontal.' Finally, the **social science teacher** uses a graphic organizer to help students categorize terminology, but she does not use other supportive materials.

Academic language

Academic language is present in all lessons, and the teachers employ many scaffolding strategies to support academic language development (science 60%, geography 90%, social science 54%). Geography concentrates the most on the meanings of many terms, and it provides the students with the most opportunities to discuss terminology. All the teachers appear highly aware of academic language, and most lessons center around terminology. Throughout the lessons, students must identify, define, and explain subject-specific terminology. The teachers strategically use L1 to provide bilingual translations. The **science** and **geography teacher** frequently ask for definitions of subject-specific terminology, whereas the **social science teacher** asks how students personally interpret abstract concepts (see excerpt 4).

In the next excerpt, the geography teacher began the lesson by moving from one topic (weathering) to a new topic (erosion). The students were given two minutes to discuss the difference between these topics in groups, and now they have a classroom discussion:

Excerpt 2 (Geography, Academic Language, score 4):

Interaction	Action
Geography teacher: There's a difference, isn't it? Between weathering and erosion. And the main difference being?	Teacher introduces vocabulary of the day. She starts by prompting students to discuss the difference between two terms.
Student 1: Weathering is breaking?	
Geography teacher: Yeah.	Student 1 provides a definition of 'weathering' and 'erosion.'
Student 1: And erosion is like carrying it. . .	
Geography teacher: Carrying it away [gesticulates]. Yeah, that's it. [. . .] Alright, then. You talked for a couple of minutes, right? One minute. One minute. So, what is it? Weathering is, you know, in situ. Right there. Breaking it down, right there. Right? What about erosion? Student 2.	Teacher repeats and asks for clarification of 'erosion.'
Student 2: It's the transfer of sediments. Like wind and the sea.	Another student gives a more accurate definition.
Geography teacher: That's it. You know . . . moving. . . ?	Teacher is still asking for a different definition
Student 3: Rock?	
Geography teacher: It away. Transportation. Right? Transportation. Alright? What else? Student 4?	Teacher highlights key word: transportation.
Student 4: Um.	

Geography teacher: Erosion.	Teacher asks for more information about erosion.
Student 4: Well, she mentions kind of like taking away the residue?	Student provides answer, reformulating Student 1's answer.
Geography teacher: Yeah, mm-hmm?	
Student 4: Already broken down through weathering?	
Geography teacher: Wonderful. The sediments, right? Yeah. Mm-hmm. That's it. Mm-hmm. Alright. What else? Student 5?	Teacher confirms and asks for more information.
Student 5: Well, that was kind of what we talked about.	
Geography teacher: What you talked about, yeah?	
Student 5: You had the erosion, right? It's only the transport of rocks …	Student repeats information.
Geography teacher: Yeah? Mm-hmm, mm-hmm.	
Student 5: By water.	
Geography teacher: Mm, ok	
Student 6: Oh, uh, erosion combined together with weathering is what breaks the mountains apart, and if you were to only say that weathering is a power that breaks everything and erosion is what picks everything up and moves it.	Student understands what teacher is prompting and reformulates erosion and weathering in her own words.
Geography teacher: Yeah. It moves it around. That's it.	

There is a high use of terminology in the excerpt. An interesting observation is the tension between everyday explanations of scientific terminology (e.g., 'weathering is breaking'). The segment scores a 4 because the teacher consistently introduces, defines, and prompts terminology and because the students have many opportunities to use their own definitions.

Task-solving scaffolding strategies

Uptake of student responses (UP)

The students have many opportunities to speak, and the teachers often expand on their ideas (science 50%, geography 60%, social science 91%). The teachers revoice student answers into academic language, prompt students to elaborate, and use student examples to further build on

ideas and concepts. However, there is a noticeable difference between the natural sciences (science and geography) and social science. **Science and geography** are characterized by display questions with yes/no answers half of the time, which leads to briefer student responses. This in turn leads to several IRE sequences. **The social science teacher** poses more referential questions and allots more time to open classroom discussions. The next excerpt is from social science. In this excerpt, the students are working in groups to discuss the difference between the terms 'rule' and 'law.' The teacher stops by a group for a desk talk:

Excerpt 3 (Social Science, Uptake of Student Responses, score 4)

Interaction	Action
Student 1: Ok, so rules are like . . . for smaller places, like schools and, like, organizations, and stuff, but laws are like for all of society.	Students are discussing a referential question in a group of three students: what is the difference between rules and laws? Student 1 is trying to explain to the other students.
Student 2: [Unintelligible]	
Student 1: Yeah, but he's written down, like . . . They're kind of the same, but not the same. Like, it depends on, like . . . uh, the school has kind of laws, but they're like rules, because. . .	
Social science teacher: Yeah, you should listen to what she has to say.	Teacher is encouraging student to continue with her train of thought.
Student 1: Society. . .	
Social science teacher: Mm.	Teacher prompts student to continue.
Student 1: And society has rules, but they're called laws because they apply to everyone.	
Social science teacher: So, rules are more limited, for example, like you said, school regulations are an example of rules. And there might be rules, sort of, anywhere. You could have rules for your class, you know, you probably did that when you were in … yeah? Or might be, even, you know, in public buildings, or if you go to a gym, there might be rules, how to use the locker room, what to do or what not to do, you know. So, rules are more limited, like I said, legislation generally is, you know, nation-wide.	Teacher expands on student idea (rules are more limited). She introduces more subject-specific terminology (school regulations, legislation).
Student 1: Isn't that kind of like … you can say that if you break the rules, you can have some sort of punishment.	Student builds on her own idea of rules.
Social science teacher: Some kind of sanction, yeah.	Teacher revoices in academic language.
Student 1: But, if you break the law, it's quite the hardest punishment you. . .	Student continues building on the differences between rules and laws.

Social science teacher: Mm. . .	
Student 1: And it's more like . . . yeah, it's more serious.	
Social science teacher: Yeah.	Teacher confirms idea.
Student 1: And. . .	
Social science teacher: Generally. And depending on what law you break. Of course, if you killed somebody, it's extremely serious. If you drive too fast, if it's not TOO fast, you just have to pay if you're caught.	Teacher nuances student idea.
Student 1: Yeah.	
Social science teacher: If you're not caught, of course, there are no consequences. Other than actually, maybe, causing more danger on the roads, in a way.	Teacher continues building on student idea.
Student 1: And there are different kinds of laws. And rules.	

The student responses are long and not teacher-directed. The teacher responds by building on student ideas and revoicing ideas in academic language. The teacher does not pose any questions, but the task allows students to explain how they understand terminology. The segment scores a 4, as the teacher is consistently referring to and building on student ideas.

Strategy use and instruction (SUI)

There is little evidence of strategy instruction except in science (science 40%, geography 0%, social science 18%). This means that, overall, CLIL students are provided little explicit and detailed instruction on strategies to help them complete tasks. However, the students are mainly working with discussion tasks. They do not make any tangible products, such as texts or posters. Therefore, a central question to pose could be whether strategy instruction becomes more prominent if students are working on tangible products.

Modeling and use of models (MOD)

No models were found, and there are limited instances of modeling (science 30%, geography 30%, social science 27%). Modeling consists of walkthroughs in which the teacher asks students to define

terminology and later models an answer. The teachers do not decompose features of modeling to explicitly illustrate what they are doing.

Discussion

This study has sought to shed light on how CLIL teachers scaffold their students by identifying their scaffolding strategies during interaction. The findings indicate that CLIL teachers provide many scaffolding strategies with which to comprehend material. This is realized through linking concepts in L1 and L2, defining and prompting students to use subject-specific terminology, and the use of visual aids. Some of these strategies have been identified in previous CLIL literature (Author 1, Colleague 1, and Colleague 2 2018; Dalton-Puffer 2007). They stand in contrast to Pawan (2008)'s study, which suggests that content teachers in ELL classrooms only use scaffolding strategies for comprehending material 28% of the time. Nineteen percent of ELL teachers expressed that aiding ELLs in comprehending material was not their responsibility. This may suggest a contrast between CLIL and ELL teaching: CLIL teachers are more preoccupied with students understanding the material, perhaps because all their students are SLLs. In ELL contexts (i.e., immigrant students placed in classrooms with L1 students), the needs of ELLs may be overshadowed by the needs of L1 students. This may suggest that the more homogenous a group of language learners is, the more advantages teachers will have in helping their students comprehend material. The findings of this study further suggest that content teachers without formal training in language teaching can use scaffolding strategies to support the comprehension of material.

On the other hand, the findings show that CLIL teachers use limited strategies to help students solve tasks (metacognition). It is worth noting that the students do not create any tangible products (texts, posters, presentations) in the course of the twelve hours. They are largely discussing and trying to comprehend material. This may lead to a lack of strategies and modeling for students to

complete tasks. These forms of metacognition are an important step toward becoming an independent learner, and it is problematic that they are absent in teaching (Gaskins et al. 1997).

Subject-specificity in CLIL teaching

An important finding is the divide in the use of scaffolding strategies between natural sciences (science and geography) and social science subjects. This divide may be explained by the historicity and nature of the subjects—the way they have been developed, practiced, and taught over the years (Nikula et al. 2016). The natural sciences provide multiple supportive materials, whereas social science provides limited supportive materials, which is in line with Author 1, Colleague 1, and Colleague 2 (2018). This difference incidentally makes natural sciences more understandable, as they provides students with contextual clues. The social science teaching, in turn, has more in-depth conversations. Discussions are student-led, have fewer IRE patterns, and provide more referential questions. This leads to longer stretches of student speech and allows students to expand more on their ideas. Several studies have found an overabundance of display questions and IRE patterns in the natural sciences (Lemke 1990; McNeil 2011; Mortimer 2003). The IRE pattern is incompatible with tenets of scaffolding, as it may stifle student autonomy and shorten student answers (Kinginger 2002; Walqui 2006).

Although some of the scaffolding elements in PLATO score similarly, the teachers may still use different strategies. Science uses the most real-world examples to connect to prior knowledge, reflecting that it is a subject that expresses how the world works (Mortimer 2003). Geography connects to national and local knowledge, showing that it is a subject that builds national identity (Sætre 2013). Lastly, social science relates to more personal examples, relating to its promotion of civic competence (Torrez and Claunch-Lebsack 2013). These findings highlight the importance of

subject-specificity in teaching. Natural science and social science subjects provide different types of support for SLLs, and these differences appear to complement each other.

Conclusion

This study has used existing literature to create a framework with which to study scaffolding. In this study, twelve hours of CLIL teaching were observed, and scaffolding strategies were identified with a coding manual to determine which scaffolding strategies three CLIL teachers use to help their students comprehend material and solve tasks. The findings indicate that CLIL teachers use a variety of scaffolding strategies in science, geography, and social science. Many of the scaffolding strategies pertain to comprehension, in which the teachers show connections between known and unknown knowledge, use supportive materials, and define and prompt academic language. The teachers build on student ideas to help students solve tasks, but they show little evidence of metacognition. There are further differences between how teachers scaffold in the natural sciences and social sciences. One implication from the findings is that context is important: there are clear differences between how CLIL and ELL teachers scaffold. The homogeneity of CLIL teachers and students allows them to better scaffold the comprehension of material. However, these teachers show less evidence of scaffolding the solving of tasks. Lastly, this study suggests that content teachers support their L2 students even when they do not have a background in language teaching.

The strength of this study is that it unifies understandings of scaffolding in SLL literature. It cross-compares three subjects and teachers in one classroom (see Author 1, Colleague 1, and Colleague 2 2018). The design is systematic and detailed and uses a validated and reliable tool (PLATO) to measure scaffolding. However, the limitations of this study are that it provides insight into only one CLIL classroom and that it does not consider student perspectives. The next step in

scaffolding research is to discuss how we can empirically measure how students perceive scaffolding strategies and how they become more independent learners. Teacher-centered approaches like PLATO do not fully cover these dimensions of scaffolding. Further research could delve into student-centered approaches and how students may experience scaffolding (Koole and Elbers 2014; Maybin, Mercer, and Stierer 1992).

References

Ajayi, Lasisi. 2014. "Vocabulary instruction and Mexican–American bilingual students: how two high school teachers integrate multiple strategies to build word consciousness in English language arts classrooms." *International Journal of Bilingual Education and Bilingualism* 18 (4):1-22. doi: 10.1080/13670050.2014.924475.

Author 1, Colleague 1, and Colleague 2. 2018.

Banse, Holland W., Natalia A. Palacios, Eileen G. Merritt, and Sara E. Rimm-Kaufman. 2017. "Scaffolding English Language Learners' Mathematical Talk in the Context of Calendar Math." *Journal of Educational Research* 110 (2):199-208. doi: 10.1080/00220671.2015.1075187.

Barr, Sheldon, Zohreh R. Eslami, and R. Malatesha Joshi. 2012. "Core Strategies to Support English Language Learners." *Educational Forum* 76 (1):105-17. doi: 10.1080/00131725.2011.628196.

Bliss, Joan, Mike Askew, and Sheila Macrae. 1996. "Effective Teaching and Learning: scaffolding revisited." *Oxford Review of Education* 22 (1):37-61. doi: 10.1080/0305498960220103.

Boche, Benjamin, and Megan Henning. 2015. "Multimodal Scaffolding in the Secondary English Classroom Curriculum." *Journal of Adolescent and Adult Literacy* 58 (7):579-90.

Cenoz, Jasone, Fred Genesee, and Durk Gorter. 2014. "Critical Analysis of CLIL: Taking Stock and Looking Forward." *Applied Linguistics* 35 (3):243-62.

Cohen, Julie. 2018. "Practices that cross disciplines?: Revisiting explicit instruction in elementary mathematics and English language arts." *Teaching and Teacher Education* 69:324-35. doi: 10.1016/j.tate.2017.10.021.

Colleague 3, Colleague 4, and Colleague 5. 2017.

Coyle, Do, Philip Hood, and David Marsh. 2010. *Content and Language Integrated Learning*. UK: Cambridge University Press.

Cummins, Jim. 2013. "BICS and CALP: Empirical Support, theoretical status, and policy implications of a controversial distinction." In *Framing languages and literacies: Socially situated views and perspectives*, edited by Margaret Hawkins, 10-23. New York: Routledge.

Cummins, Jim, and Margaret Early. 2015. *Big Ideas for Expanding Minds: Teaching English Language Learners Across the Curriculum*. Canada: Pearson.

Dalton-Puffer, Christiane. 2007. *Discourse in content and language integrated learning (CLIL) classrooms*. Amsterdam: John Benjamins Publishing.

De Graaff, Rick, Gerrit Koopman, Yulia Anikina, and Gerard Westhoff. 2007. "An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL)." *International Journal of Bilingual Education and Bilingualism* 10 (5):603-24. doi: 10.2167/beb462.0.

Dong, Yu Ren. 2017. "Tapping into English Language Learners' (ELLs') Prior Knowledge in Social Studies Instruction." *The Social Studies* 108 (4):143-51. doi: 10.1080/00377996.2017.1342161.

Echevarría, Jana, MaryEllen Vogt, and Deborah Short. 2017. *Making Content Comprehensible for English Learners. The SIOP Model*. 5th ed. USA: Pearson.

Farooq, M. 2007. "Exploring the effects of spoken English classes of Japanese EFL learners." *The Journal of Liberal Arts* 2:35-57.

Gallagher, Fiona, and Gerry Colohan. 2014. "T(w)o and fro: using the L1 as a language teaching tool in the CLIL classroom." *The Language Learning Journal*:1-14. doi: 10.1080/09571736.2014.947382.

Gaskins, Irene W., Sharon Rauch, Eleanor Gensemer, Elizabeth Cunicelli, Colleen O'Hara, Linda Six, and Theresa Scott. 1997. "Scaffolding the Development of Intelligence among Children who are Delayed in Learning to Read." In *Scaffolding Student Learning*, edited by Kathleen Hogan and Michael Pressley, 43-73. Canada: Brookline Books, Inc.

Gibbons, Pauline. 2003. "Mediating Language Learning: Teacher Interactions with ESL Students in a Content-Based Classroom." *TESOL Quarterly* 37 (2):247-73. doi: 10.5054/tj.2010.215611.

———. 2015. *Scaffolding Language, Scaffolding Learning. Teaching English Language Learners in the Mainstream Classroom*. New Hampshire, USA: Heinemann.

Gritter, Kristine, Scott Beers, and Robert W. Knaus. 2013. "Teacher Scaffolding of Academic Language in an Advanced Placement U.S. History Class." *Journal of Adolescent & Adult Literacy* 56 (5):409-18. doi: 10.1002/JAAL.158.

Grossman, Pamela. 2015. *Protocol for Language Arts Teaching Observations (PLATO 5.0)*, *Center to Support Excellence in Teaching (CSET)*. Palo Alto: Stanford University.

Grossman, Pamela, Susanna Loeb, Julie Cohen, and James Wyckoff. 2013. "Measure for Measure: The Relationship between Measures of Instructional Practice in Middle School English Language Arts and Teachers' Value-Added Scores." *American Journal of Education* 119 (3):445-70. doi: 10.1086/669901.

Hogan, Kathleen, and Michael Pressley. 1997. *Scaffolding Student Learning*. Canada: Brookline Books Inc.

Kayi-Aydar, Hayriye. 2013. "Scaffolding language learning in an academic ESL classroom." *ELT journal* 67 (3):324-35. doi: 10.1093/elt/cct016.

Kinginger, Celeste. 2002. "Defining the Zone of Proximal Development in US Foreign Language Education." *Applied Linguistics* 23 (2):240-61.

Klette, Kirsti, and Marte Blikstad-Balas. 2018. "Observation Manuals as Lenses to Classroom Teaching: Pitfalls and Possibilities." *European Educational Research Journal* 17 (1):129-46.

Klette, Kirsti, Marte Blikstad-Balas, and Astrid Roe. 2017. "Linking Instruction and Student Achievement. A research design for a new generation of classroom studies." *Acta Didactica* 11 (3):1-19.

Koole, Tom, and Ed Elbers. 2014. "Responsiveness in teacher explanations: A conversation analytical perspective on scaffolding." *Linguistics and Education* 26:57.

Lasagabaster, David, and Juan Manuel Sierra. 2009. "Immersion and CLIL in English: more differences than similarities." *ELT journal* 94 (4):367-75.

Lemke, Jay L. 1990. *Talking science : language, learning, and values*, *Language and educational processes*. Norwood, N.J: Ablex.

Li, Danli. 2012. "Scaffolding adult learners of English in learning target form in a Hong Kong EFL university classroom." *Innovation in Language Learning and Teaching* 6 (2):127-44. doi: 10.1080/17501229.2011.626858.

Lin, Tzu-Chiang, Ying-Shao Hsu, Shu-Sheng Lin, Maio-Li Changlai, Kun-Yuan Yang, and Ting-Ling Lai. 2012. "A review of empirical evidence on scaffolding for science education " *International Journal of Science and Mathematics Education* 10 (2):437-55. doi: 10.1007/s10763-011-9322-z.

Long, M. H., and C. J. Sato. 1983. "Classroom foreigner talk discourse: forms and functions of teachers' questions." In *Classroom oriented research in second language acquisition*, edited by H. W. Seliger and M. H. Long, 268-86. Rowley, MA: Newbury House.

Many, Joyce E., Deborah Dewberry, Donna Lester Taylor, and Kim Coady. 2009. "Profiles of Three Preservice ESOL Teachers' Development of Instructional Scaffolding." *Reading Psychology* 30 (2):148-74. doi: 10.1080/02702710802275256.

Masako, Douglas O, and Kataoka C Hiroko. 2008. "Scaffolding in Content Based Instruction of Japanese." *Japanese Language and Literature* 42 (2):337-59.

Maybin, J, Neil Mercer, and B Stierer. 1992. "Scaffolding: Learning in the classroom." In *Thinking voices. The work of the national oracy project*, edited by Kate Norman, 186-95. London: Hodder & Stoughton.

McNeil, Levi. 2011. "Using talk to scaffold referential questions for English language learners." *Teaching and Teacher Education* 28 (3). doi: 10.1016/j.tate.2011.11.005.

Meyer, Oliver, and Do Coyle. 2017. "Pluriliteracies Teaching for Learning: conceptualizing progression for deeper learning in literacies development." *European Journal of Applied Linguistics* 5 (2):199-222.

Meyer, Oliver, Do Coyle, Ana Halbach, Kevin Schuck, and Teresa Ting. 2015. "A Pluriliteracies Approach to Content and Language Integrated Learning--Mapping Learner Progressions in Knowledge Construction and Meaning-Making." *Language, Culture and Curriculum* 28 (1):41-57. doi: 10.1080/07908318.2014.1000924.

Miller, Patricia H. 2005. "Commentary on: Scaffolding: Constructing and deconstructing development." *New Ideas in Psychology* 23 (3):207-11. doi: 10.1016/j.newideapsych.2006.07.001.

Mortimer, Eduardo. 2003. *Meaning Making in Secondary Science Classrooms*. Maidenhead: McGraw-Hill Education.

Morton, Tom. 2015. "Vocabulary Explanations in CLIL Classrooms: A Conversation Analysis Perspective." *Language Learning Journal* 43 (3):256-70. doi: 10.1080/09571736.2015.1053283.

NESH. "Guidelines for Research Ethics in the Social Sciences, Law and the Humanities." National Committees for Research Ethics in Norway, Accessed 07.07.2017.

Nikula, Tarja, Emma Dafouz, Pat Moore, and Ute Smit. 2016. *Conceptualising Integration in CLIL and Multilingual Education*. UK, USA & Canada: Multilingual Matters.

Nikula, Tarja, and Karita Mård-Miettinen. 2014. "Language Learning in Immersion and CLIL Classrooms." In *Handbook of Pragmatics*, edited by Jan-Ola Östman and Jef Verschueren. John Benjamins Publishing Company.

Pawan, Faridah. 2008. "Content-Area Teachers and Scaffolded Instruction for English Language Learners." *Teaching and Teacher Education: An International Journal of Research and Studies* 24 (6):1450-62. doi: 10.1016/j.tate.2008.02.003.

Pérez-Cañado, María Luisa. 2016. "Are teachers ready for CLIL? Evidence from a European study." *European Journal of Teacher Education* 39 (2):202-21. doi: 10.1080/02619768.2016.1138104.

Ruiz de Zarobe, Yolanda, and Jasone Cenoz. 2015. "Way Forward in the Twenty-First Century in Content-Based Instruction: Moving towards Integration." *Language, Culture and Curriculum* 28 (1):90-6. doi: 10.1080/07908318.2014.1000927.

Ruiz de Zarobe, Yolanda, and Victoria Zenotz. 2015. "Reading strategies and CLIL: the effect of training in formal instruction." *The Language Learning Journal*:1-15. doi: 10.1080/09571736.2015.1053284.

———. 2017. "Learning strategies in CLIL classrooms: how does strategy instruction affect reading competence over time?" *International Journal of Bilingual Education and Bilingualism* 21 (3):319-31.

Smagorinsky, Peter. 2018. "Deconflating the ZPD and instructional scaffolding: Retranslating and reconceiving the zone of proximal development as the zone of next development." *Learning, Culture and Social Interaction* 16:70-5. doi: 10.1016/j.lcsi.2017.10.009.

Snell, Julia. 2011. "Interrogating video data: Systematic quantitative analysis versus micro-ethnographic analysis." *International Journal of Social Research Methdology* 14 (3):253-8.

Svenhard, Britt Wenche, Kim Servant, Glenn Ole Hellekjær, and Henrik Bøhn. 2007. "Norway." In *Windows on CLIL*, edited by Anne Maljers, David Marsh and Dieter Wolff, 139-46. European Platform for Dutch Education.

Sætre, Per Jarle. 2013. "The beginning of geography didactics in Norway." *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* 67 (3):120-7. doi: 10.1080/00291951.2013.796568.

Tomlinson, Carol Ann, and Tonya R. Moon. 2013. "Differentiation and Classroom Assessment." In *SAGE Handbook of Research on Classroom Assessment*, edited by James H. McMillan, 415-30. USA: SAGE Publications, Inc.

Torrez, Cheryl A., and Elizabeth Ann Claunch-Lebsack. 2013. "Research on Assessment in the Social Studies Classroom." In *SAGE Handbook of Research on Classroom Assessment*, edited by James H. McMillan, 461-72. USA: SAGE Publications Inc.

van de Pol, Janneke, and E. Elbers. 2013. "Scaffolding student learning: A micro-analysis of teacher-student interaction." *Learn. Cult. Soc. Interact.* 2 (1):32-41. doi: 10.1016/j.lcsi.2012.12.001.

van de Pol, Janneke, Monique Volman, and Jos Beishuizen. 2010. "Scaffolding in Teacher-Student Interaction: A Decade of Research." *Educational Psychology Review* 22 (3):271-96. doi: 10.1007/s10648-010-9127-6.

van Lier, Leo. 2004. *The Ecology and Semiotics of Language Learning. A Sociocultural Perspective*. Boston/Dordrecht/New York/London: Kluwer Academic Publishers.

Verenikina, Irina. 2004. "From Theory to Practice: What does the Metaphor of Scaffolding Mean to Educators Today?" *Outlines: Critical Practice Studies* 6 (2):5-16.

Walqui, Aida. 2006. "Scaffolding Instruction for English Language Learners: A Conceptual Framework." *International Journal of Bilingual Education and Bilingualism* 9 (2):159-80.

Wood, David, Jerome S. Bruner, and Gail Ross. 1976. "The Role of Tutoring in Problem Solving." *Journal of Child Psychology and Psychiatry* 17 (2):89-100. doi: 10.1111/j.1469-7610.1976.tb00381.x.

Tables and figures

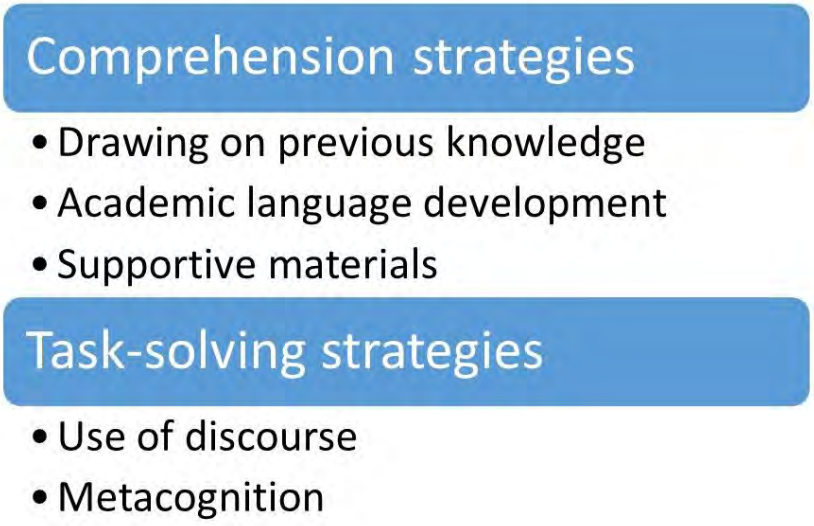


Figure 1. Author's classification of SLL scaffolding strategies, modified from Maybin, Mercer, and Stierer (1992)

Table 1. Classification of scaffolding strategies and how they were scored in PLATO

Type of scaffolding	PLATO element	Definition and coding process
Comprehension of material	Connections to prior knowledge (CPK)	*Connections to prior knowledge (CPK)* refers to what degree a teacher connects new material to the students' prior knowledge (Grossman 2015). This can be done through references (e.g., 'Last week, we talked about [. . .]'), and connections (e.g., 'We all know about weathering, so now we will talk about erosion. The main difference being?'). At the low end, the teacher does not refer to (score 1) or superficially refers to prior knowledge (score 2). At the high end, the teacher refers to prior knowledge multiple times (score 3), and new material builds on prior knowledge in such a way that students can understand it (score 4).
	Supportive materials (SUP)	*Supportive materials (SUP)* include body language and different sorts of differentiated materials, visual aids, and graphic organizers that enable L2 English students to understand a lesson (see Author 1, Colleague 1, and Colleague 2 2018, for examples). At the low end, the teacher does not provide supportive materials (score 1) or the provided materials are not used by the teacher or students (score 2). At the high end, the teacher provides supportive materials, prompts the students to use them (score 3), and there is evidence that students use them (score 4).
	Academic language	*Academic language (AL)* refers to the subject-specific terminology students will need to understand a lesson (see Author 1, Colleague 1, and Colleague 2 2018). At the low end, the teacher does not use academic language (score 1) or uses it without explaining it to the students (score 2). At the high end, the teacher introduces, defines, and prompts students to use academic language (score 3) and provides opportunities for students to use the terminology (score 4).

Completion of task	Uptake of student responses (UP)	*Uptake of student responses (UP)* refers to the degree to which a teacher/student elaborates or follows up on ideas (Grossman 2015). This can include revoicing an idea in academic language or commenting, elaborating, clarifying, or expanding on an idea (see Author 1, Colleague 1, and Colleague 2 2018). At the low end, the teacher/student provides no or few responses (score 1) or responds briefly to ideas without pushing for elaboration or expanding upon them (score 2). At the high end, the teacher revoices ideas in academic language, asks for elaboration or evidence, responds in a way that expands on student ideas, or enables students to explain, clarify, or explain their thinking briefly (score 3) or consistently (score 4).
	Strategy use and instruction (SUI)	*Strategy use and instruction (SUI)* describes a teacher's use of strategies and skills that support students' learning during the task at hand. A strategy is a general/flexible method or 'how to' that a teacher suggests to solve a task (Grossman 2015). At the low end, the teacher does not provide any strategy instruction (score 1) or briefly introduces a strategy but does not provide explicit instruction on how to use it (score 2). At the high end, the teacher provides explicit instruction (score 3) and specifies how, why, and when to use it (score 4).
	Modeling and use of models (MOD)	*Modeling and use of models (MOD)* refers to the degree to which a teacher visibly enacts targeted strategies, skills, and processes in a lesson (Grossman 2015). PLATO differentiates between physical models (e.g., an example text or a model of an item to be built) and modeling (e.g., when a teacher orally 'walks through' the process of how to solve a task). At the low end, the teacher does not provide a model (score 1) or only partially provides a model (2). At the high end, the model is complete (score 3) and the teacher decomposes specific features of it, explaining how and why to use it (score 4).

Article III

Mahan, K. R., Norheim, H. (Under review). Something new and different: Student perceptions of Content and Language Integrated Learning (CLIL). Submitted to *ELT Journal.*

Under review, not included in online publication

Appendix 1: Declaration of authorship for Articles I and III

Declaration of authorship

Title: *Characterizing CLIL teaching: new insights from a lower secondary classroom*

The International Journal of Bilingual Education and Bilingualism

Mahan, K. R., Brevik, L.M., Ødegaard, M, 2018.

In correspondence with the Vancouver convention of defining authorship, we declare the following:

Karina Rose Mahan was the first and corresponding author of this article. She made substantial contribution to the conception of the article, the acquisition and interpretation of data. She wrote the majority of the article and carried out the majority of revisions. She has approved the final version and has agreed to be accountable for all aspects of the work's integrity.

Lisbeth M. Brevik was the second author. She was responsible[1] for the design and interrater agreement procedures, and involved in the interpretation of data. She wrote the majority of the methods section and the context section (CLIL in Norway), contributed to the structuring of the article and authored parts of the discussion. She carried out the revisions with the first author, approved the final version and has agreed to be accountable for all aspects of the work's integrity.

Marianne Ødegaard was the third author. She was involved in the conception, design and interpretation of data. She wrote the sections on science and science education in the theory section (Conceptualizing integration in CLIL) and parts of the discussion. She revised the article critically for important intellectual content, approved the final version and has agreed to be accountable for all aspects of the work's integrity.

Signature of candidate

Signature of co-authors

Date: 21.05.2018

[1] Lisbeth M Brevik is the project coordinator of the *Linking Instruction and Student Experiences* (LISE) project, which this article builds on, including conception and design. Professor Kirsti Klette, University of Oslo, is the project leader of LISE.

Declaration of authorship

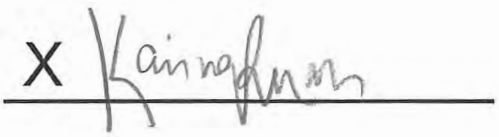
Title: *Something New and Different. Student Perceptions of CLIL*

Submitted to *ELT Journal*, 21.10.2019

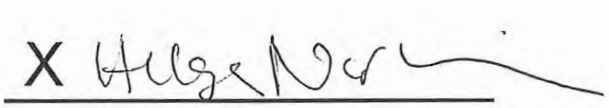
In correspondence with the Vancouver convention of defining authorship, we declare the following:

Karina Rose Mahan was the first and corresponding author of this article. She made substantial contributions to the conception of the article, the acquisition and interpretation of data. She wrote the majority of the article and carried out the majority of revisions. She has approved the final version and has agreed to accountable for all aspects of the work's integrity.

Helga Norheim was the second author. She was responsible for the majority of the data analysis for the Tripod questionnaire. She wrote parts of the methods and results. She carried out the revisions with the first author, revised the article critically for important intellectual content, approved the final version and has agreed to accountable for all aspects of the work's integrity.

X 

Signature of candidate

X 

Signature of co-author

Date: 22.01.2020

Appendix 2: Questionnaire I, used in Article III

Questionnaire I

Spørreundersøkelse i engelsk

Hei og takk for at du vil delta i mitt ph.d. prosjekt. Her er noen spørsmål angående hvordan du oppfatter og bruker engelsk i ditt daglige liv. Svar så ærlig og presist som du kan.

Jeg er *

Mitt fulle navn er: *

Bakgrunn

Hva er ditt førstespråk? *

Har du bodd i et annet land? *

Hvilke(t) land og hvor lenge bodde du der?

Gikk du på engelsktalende skole? Hvis ja, hvor lenge?

Engelskvaner

Snakker du engelsk utenfor skolen? *

I hvilke sammenheng snakker du engelsk? *

Hvor mye engelsk snakker du? Oppgi hvor mange timer daglig ca.

Hva slags engelske tekster leser du utenom skole? *

Hvor mye tid bruker du på å lese engelsk utenom skolen? Oppgi hvor mange timer daglig ca. *

Hva slags engelske tekster skriver du? *

Hvor mye tid bruker du på å skrive engelsk utenom skolen? Oppgi hvor mange timer daglig. *

Hva lytter du til på engelsk utenom skolen? *

Hvor mye tid bruker du på å lytte på engelsk? Oppgi hvor mange timer daglig ca. *

Jeg foretrekker følgende på engelsk utenom skolen

Kan krysse av mer enn én boks

Lytte

Skrive

Snakke

Skrive

Hva mener du om engelsk som språk? *

Vurder engelsknivået ditt *

Hvilken karakter hadde du i engelsk skriftlig til sommeren på 10. trinn? *

Hvilken karakter hadde du i engelsk muntlig til sommeren på 10. trinn? *

Hvorfor valgte du å begynne i en klasse som har undervisning på engelsk i andre fag?

Ser du noen fordeler eller ulemper ved å gå i denne klassen?

Takk for at du deltok i spørreundersøkelsen min!

Appendix 3: Information slip and consent form for project

Høgskolen i Sørøst-Norge

Til elever og foresatte

Institutt for språkfag
Raveien 215, 3184 Borre
Campus Vestfold

Telefon: 93 23 13 28
Mail: karina.mahan@hbv.no

To use or not to use CLIL? A Bird's-Eye View of Content and Integrated Learning in Norway: Teacher Practices and Student Achievement in Secondary Schools.

Jeg er en doktorgradsstipendiat i engelsk fagdidaktikk ved Høgskolen i Sørøst-Norge og undersøker tospråklig undervisning i Norge – Content and Language Integrated Learning (CLIL).

I forbindelse med dette ønsker jeg å gjøre videoopptak tilknyttet den tospråklige undervisningen hos dere i fire timer i et eller flere av følgende fag høsten 2016: engelsk, samfunnsfag, geografi og naturfag (vg1), internasjonal engelsk, sosiologi og sosialantropologi og historie (vg2), samfunnsfaglig engelsk, historie og religion (vg3). Jeg vil bruke en to-kameraløsning der ett kamera filmer hele klassen og ett kamera filmer læreren.

Jeg ber med dette om tillatelse til å foreta lyd- og videoopptak av klassen som helhet. Jeg ber derfor om at elevene skriver under på den vedlagte avtalen.

Prosjektet vil bli gjennomført i henhold til gjeldende lovverk for personvern og forskningsetiske retningslinjer. Studien er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD), og alle personopplysninger vil bli behandlet konfidensielt. Deltakerne vil ikke kunne bli gjenkjent i publikasjoner fra prosjektet.

Deltagelse i prosjektet er frivillig og det vil ikke ha konsekvenser for undervisning eller skole om den enkelte elev deltar eller ikke.

Med vennlig hilsen

Karina Rose Mahan
Stipendiat/ prosjektansvarlig

Lisbeth M Brevik
Førsteamanuensis/ Veileder

Høgskolen i Sørøst-Norge

Deltakelse i tilknytning til forskningsprosjektet «To use or not to use CLIL? A Bird's-Eye View of Content and Integrated Learning in Norway: Teacher Practices and Student Achievement in Secondary Schools."

Elevens navn

- ☐ Ja, jeg godtar å være med i prosjektet
- ☐ Undertegnede godtar at kopi av elevarbeider samles inn

Dato Sted

Elevenes underskrift

All deltakelse i prosjektet er frivillig, og du kan når som helst trekke ditt samtykke uten noen grunn. Dersom du trekker deg, vil alle opplysninger bli anonymisert.

Dersom du har noen spørsmål til studien, vennligst ta kontakt med stipendiat Karina Rose Mahan (karina.mahan@hbv.no), tlf. 93 23 13 28.

Vennligst returner svarslippen til læreren så fort som mulig.

Kort om CLIL-prosjektet (Content and Language Integrated Learning)

CLIL-prosjektet er et doktorgradsprosjekt ved Høgskolen i Sørøst-Norge som ønsker å gjennomføre en studie i CLIL-undervisning på ungdomstrinnet og videregående. Prosjektet vil innhente opplysninger fra elevarbeider, spørreskjemaer, samt lyd- og videoopptak av lærere og elever på ungdomstrinnet og videregående. Målet er å etablere hva en CLIL-time består av, elevprestasjoner i engelsk og faget som undervises, synspunkter om undervisning og engelsk som språk.

Prosjektet vil bli gjennomført i henhold til gjeldende lovverk for personvern og forskningsetiske retningslinjer og er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD). Alle personopplysninger vil bli behandlet konfidensielt og deltakerne vil ikke kunne bli gjenkjent i publikasjoner fra prosjektet. Prosjektet har finansiering fram til 2019.

Jeg samarbeider med LISA-prosjektet (Linking Instruction and Student Achievement), ledet av Professor Kirsti Klette ved Institutt for lærerutdanning og skoleforskning (ILS), Universitetet i Oslo, ved å bruke tilsvarende analyseverktøy og forskningsdesign. Datamaterialet vil bli lagret i sikker forskningsserver ved UiO/USIT til 2022. Alle deltakere vil bli anonymiserte, som betyr at navn fjernes fra elevarbeider straks de blir samlet inn, og videodata slettes i 2022. Det kan være aktuelt å komme tilbake til enkelte klasser for en mulig oppfølgingsstudie.

Deltagelse i prosjektet er frivillig og det ikke vil ha konsekvenser for undervisning eller skole om den enkelte elev deltar eller ikke.

Takk for hjelpen!

Karina Rose Mahan
Doktorgradsstipendiat
Høgskolen i Sørøst-Norge

Doctoral thesis no. 71
2020

—

Teaching Content and Language Integrated Learning (CLIL): Classroom practices and student perspectives in three Norwegian classrooms

Dissertation for the degree of Ph.D

—

Karina Rose Mahan

—

ISBN: 978-82-7860-416-8 (print)
ISBN: 978-82-7860-417-5 (online)

usn.no