

CLIL in Context: Profiling Language Abilities

Carmen Pérez-Vidal and Helena Roquet

1 Introduction

The contrast of different learning contexts and their impact on learners' linguistic development in a second or foreign language is one of the new areas of interest in SLA research. Collentine and Freed (2004: 158) have argued for the relevance of analysing context dependent effects: “[...] The study of SLA within and across various contexts of learning forces a broadening of our perspective of the most important variables that affect and impede acquisition in general”. Their study dealt with an unconventional learning context, Study Abroad (SA), and its impact on learners' linguistic and attitudinal development. The CLIL (content and language integrated learning) approach to education represents another unconventional learning context in that respect. It is also one that holds enormous promise for the field of SLA research, in spite of being still in its infancy, in terms of both accumulated experiences and assessment of results. Indeed, in France, Italy, Spain, Finland, The Netherlands, to take a few reported cases, programmes, either in foreign or second languages, really only got off the ground at the beginning of and throughout the 1990s, as described in Grenfell (2002). However, there are a few exceptions such as Germany, where the first CLIL programmes date from the mid 1960s (Wolff 2002; Zydattiss 2012), and Belgium (Van de Craen et al. 2007). They were visibly dovetailing the European policies of the time (European Commission 1995). In sum, both CLIL experiences as such and CLIL studies are young, particularly in comparison with nearly 50 years of Canadian immersion experiences.

CLIL has been defined as an European approach to education in which a language different from the domestic language is used as the medium of instruction for

C. Pérez-Vidal (✉) • H. Roquet

Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra (UPF),
Barcelona, Spain

e-mail: carmen.perez@upf.edu; helena.roquet@upf.edu

© Springer International Publishing Switzerland 2015

M. Juan-Garau, J. Salazar-Noguera (eds.), *Content-based Language Learning in Multilingual Educational Environments*, Educational Linguistics 23,
DOI 10.1007/978-3-319-11496-5_14

237

curricular subjects at primary and secondary stages of education.¹ The origins of the CLIL proposal lie in bilingual approaches to education in Europe and around the world either in particular schools or school regions (see Lasagabaster 2015). CLIL rapidly showed the capacity to take on board lessons drawn from them and in particular from the Canadian and US immersion programmes and content based instruction (CBI).² The ultimate objectives of CLIL programmes have been summarised by Zydattiss, who underscores the ‘bifocal’ nature of the approach encompassing both content and language, as has been strongly advocated by specialists, in the following highly realistic terms:

[...] the overriding purpose of the CLIL approach in our multilingual highly mobile societies would seem to be empowerment of school learners (through the performance of scholastic tasks) to acquire subject knowledge, study skills and cognitive operations (based on verbal thought) via a foreign language, almost regardless of which particular school subject or topic may be chosen in a specific instructional setting. (Zydattiss 2012: 23)

It is indeed true that CLIL can be described as an idiosyncratic development in modern European educational policies vis-à-vis languages, as Pérez-Vidal (2015) and Lasagabaster (2015) present in detail. It reflects multilingual policies and the promotion of mobility and internationalisation as the ultimate goal across the educational systems of the 27 member states in the Union, with its total 23 languages and populations “exhibiting mostly a monolingual habitus”, as Dalton-Puffer (2011: 185) very rightly points out (for an update on European multilingual educational policies, see Cenoz and Genesee 1998; Cenoz and Jessner 2000; Wilkinson 2004; Dalton-Puffer 2007, 2008; Dafouz and Guerini 2009).³

If we now turn to CLIL studies, research on the linguistic, content and attitudinal effects of CLIL has gained momentum throughout the past decade producing the first interesting findings both regarding its pedagogical dimension, with an emphasis on the pragmatic and discourse features of classroom language, content learning attainment and linguistic progress, the latter being the focus of the study presented in this chapter. The psycholinguistic dimension of CLIL has been described in terms of the quality of its input (meaning oriented) and interaction (focused on subject matter) and the cognitive/learning abilities which it fosters (see Muñoz 2015). Focusing on such language learning outcomes, we are beginning to identify the areas of second/foreign language competence, which are most likely to benefit from CLIL instruction and those which seem to do less so, and the variables which seem to affect progress or lack thereof (see in this volume: Ruiz de Zarobe 2015; Prieto-Arranz et al. 2015; Gené-Gil et al. 2015; Rallo Fabra and Jacob 2015; Juan-Garau et al. 2015;

¹ ‘Integrating content and language’ (in ‘higher education’), that is ILC/ILCHE are the corresponding terms used to refer to the same or similar approaches, respectively, at tertiary level, or the specific ones ‘English-medium instruction’ (EMI) and ‘English taught programmes’ (ETP) (see Pérez-Vidal 2015 for further details).

² See Pérez-Vidal (2009, 2013) for the European perspective; Genesee (1987, 2013) and Wesche (2002) on the Canadian experiences; and Brinton et al. (1989) for an update of CBI in the US.

³ Dalton-Puffer (2011: 184) also notes that an analogous process seems to have taken place in other continents, Latin America and China being cases in point.

Amengual-Pizarro and Prieto-Arranz 2015; Menezes and Juan-Garau 2015). This chapter focuses on such different linguistic gains when analysing secondary school CLIL learners enrolled on a programme offered at a school in a large Catalan city in Spain (see Juan-Garau and Salazar-Noguera 2015 for a general description of secondary school contexts in our setting). The programme was carefully designed by the school board and language specialists, both internal and external, well in advance of its implementation, in order to ensure maximum efficacy and stability over the years. The study features intergroup contrasts with a non-CLIL group. The chapter first summarises the current state of thinking regarding CLIL effects on linguistic progress, then presents the study and its results and finally discusses them and draws some conclusions.

2 CLIL Under Scrutiny

The interest in the investigation of CLIL programmes is undeniable, and only paralleled by the undiminishing interest in the analysis of the educational experiences on which CLIL has undoubtedly been modelled. These are the immersion programmes set up in the U.S., mostly for Spanish, and those in Canada, both in the mid 1960s, the latter, at the start, with the explicit goal of “additive bilingualism” for English-speaking students in a language other than that of their home and wider community, namely French. It is thus not surprising that research on the effects of European CLIL contexts of learning has often sought to substantiate North American findings with regard to such immersion programmes and probed in the same areas of learner development.⁴

The CLIL research literature has thus focused on three main areas of enquiry in which such an educational approach was expected to have an impact and become a “catalyst for change”, as Dalton-Puffer calls it (2011: 186): the learners’ L1 and how it may be affected by the use of an L2 as the medium of instruction; content learning attainment through an L2 and target language progress (see Ruiz de Zarobe 2015 as an example of a chapter encompassing those three research strands).

With respect to the first issue, Canadian research not only shows the absence of negative effects of immersion on the development of learners’ L1s but it actually posits cognitive benefits (see for example Cummins 1976) and advantages on content learning (Genesee 1994). Turning to the linguistic effects of immersion, “meticulous research has put it under the microscope in its various forms for the past 35

⁴In Canada, early French immersion started with a well-researched experiment in St. Lambert, Quebec, in a kindergarten in 1965. Thirty-five years later, the last count stood at over 325,000 students enrolled on the programmes, that is 10 % of the children in English language schools studying French (that represents 55 % of elementary school students and 47 % of secondary school students) (Wesche 2002: 358). Immersion exists in indigenous and non-official heritage languages. In the U.S., immersion started in the Coral Way school programme in Miami in the 1960s to serve school populations with either English or Spanish in both languages. It is highly successful in around 248 two-way programmes in 23 states and Washington D.C. in more than ten languages.

[now 45] years, and documented it in several thousands of reports to school boards, articles, book chapters, master and doctoral theses, and books” (Wesche 2002: 357).⁵ What emerges from those studies is that, in comparison to non-immersion students, immersion students develop (a) almost nativelike comprehension skills as measured by tests of listening and reading comprehension; and (b) high levels of fluency and confidence in using the second language and a more open attitude towards French culture “helping to close the gap between Canada’s English and French solitudes” (Swain 2000: 208), while (c) production skills seem to be non-nativelike in terms of grammatical accuracy, lexical variety and sociolinguistic appropriateness. Consequently, immersion students in Canada have been found to be second language speakers who are relatively fluent and effective communicators, but non-targetlike in terms of grammatical structure and non-idiomatic in their lexical choices and pragmatic expression in comparison to native speakers of the same age. In contrast with students learning French in traditional core French language arts courses, Wesche (2002) summarises results stating that all types of French immersion programmes, that is early immersion—starting between 4 and 6 years of age—, middle or delayed immersion—starting at age 9— and late immersion—starting at either 11 or 12—consistently lead to far stronger French proficiency in all skills than does traditional language instruction (forty minutes per day) and prepare students for bilingual secondary school programmes with approximately a third of the course work taught through French.

Against such a backdrop, if we now turn to CLIL effects concerning language achievement, general statements regarding the CLIL impact on students’ language learning outcomes are by and large very positive. This is the case of the Netherlands (Admiraal et al. 2006); Spain (Lasagabaster 2008; Alejo and MacArthur 2009; Lasagabaster and Ruiz de Zarobe 2010; Lorenzo et al. 2010; Pérez-Vidal and Juan-Garau 2010); Austria (Ackerl 2007; Dalton-Puffer 2007, 2008, 2011); Norway (Hellekjaer 2010); Sweden (Sylvén 2004); Finland (Nikula 2007), Belgium (Van de Craen et al. 2007) or Germany (Zydatiss 2007, 2012), to name but a few.

More specifically, recent updated replications by Dalton-Puffer (2011) and Ruiz de Zarobe (2008) to initial well-known reviews of findings (Dalton-Puffer 2007, 2008) emphasize the potential contrasting CLIL versus non-CLIL effects on receptive versus productive abilities as follows: (a) reading clearly improves in CLIL groups but results are mixed with respect to listening (Hellekjaer 2010); (b) CLIL groups’ receptive vocabulary clearly improves: it is larger, including words from lower frequency bands used more appropriately and with a wider stylistic range than in non-CLIL groups (Zydatiss 2007; Jexenflicker and Dalton-Puffer 2010); (c) only some morphological phenomena such as sentence complexity and affixal inflection improve with CLIL (Dalton-Puffer 2007), but not the use of null subjects, negation and suppletive forms (Villarreal Olaizola and García Mayo 2009); (d) spoken fluency rates and risk-taking rise most noticeably (Escobar Urmeneta 2004; Zydatiss 2007; Lasagabaster 2008; Ruiz de Zarobe 2008; Moore 2009); (e) written fluency

⁵ See, for example, Genesee (1987), Allen et al. (1990), Harley et al. (1990), Johnson and Swain (1997), and Genesee (2013).

and lexical and syntactic complexity improve (see previous references); and (f) so do emotive-affective factors. On the other hand, those aspects which are either unaffected by CLIL or for which research is inexistent or inconclusive are: (a) syntax; (b) productive vocabulary; (c) written accuracy and discourse skills such as cohesion and coherence, discourse structuring, paragraphing, register awareness, genre and style and pragmatic efficiency (see Whittaker and Llinares 2009; Llinares et al. 2013 for comparisons of L2 and L1 subject writing); (d) informal/non-technical language; (e) pronunciation (degree of foreign accent).

Of particular interest are studies which triangulate findings in an attempt to model patterns of learning. Zydatiss' (2007) empirical study relating language proficiency scores in the L2 and academic development is a case in point. He suggests a double language threshold (a lower one and an upper one), which would act as "intervening variables that either impede or support subject-matter learning in German CLIL classrooms" (Zydatiss 2007: 27).

However, critical voices are beginning to make themselves heard both in relation to the CLIL programmes themselves and to the research measuring outcomes. One shared general observation with data from Austria seems to be reduced active student participation in the classroom, which, as stated by Dalton-Puffer et al. (2008), may lead to less learning. Another is the finding that content teaching is conducted almost entirely without writing activities, as reflected by research findings. Criticism has been strong at the methodological level and indeed CLIL research is still at an early stage: due to the continuous growth in the number of CLIL programmes, often those under scrutiny are either in a pilot phase (see Eurydice 2006, 2008) or are purely experimental, with the array of methodological consequences that entail in terms of the reliability and validity of findings (as discussed in Moore 2009). In addition to that most samples analysed can only be compared to the same age groups of learners exposed to foreign language instruction, without the time advantage of the CLIL lessons unless ages are matched, thus representing yet another obstacle for the generalizability of results. In that vein, Bruton (2011) re-evaluates some of the existing research on CLIL particularly in terms of sampling, pretesting and observation data and questions both quantitative and qualitative results and the conclusions drawn thus far.

3 The Study: A CLIL Programme in Practice

In the study presented in this chapter, the effect of a CLIL programme on English as foreign language (EFL) linguistic progress is examined. Data were collected at a well-established school located in the city of Barcelona, Catalonia. The whole process which the school went through to launch the programme can be seen to be an example of good practice with regard to the implementation of a robust long-term programme (see Escobar Urmeneta and Pérez-Vidal 2004 for a full description of the planning phase). Eventually, assessment of results was allowed to take place and afforded the data presented in this chapter.

3.1 The CLIL Programme in Context

The school council of the Catalan educational institution in which data for this study were collected had decided to adopt a CLIL approach after taking into consideration and evaluating different innovative initiatives for their language department. Their aim was to guarantee adequate exit levels in English as a foreign language and a good preparation for a university degree where knowledge of languages was seen to be an asset.

A team of university experts, one of them being the first author of this chapter, was contacted to act as school consultants for the preparation and subsequent follow-up of the programme. They were to: (a) provide the school managers with advice on the decisions to be taken in relation with the design of the programme; (b) provide advice on how best to communicate the novelties and the rationale behind them to parents; (c) train the teachers and advise them in the design and selection of appropriate activities and materials, and in their choice of suitable teaching techniques; and (d) counsel and monitor the teachers during the first year of the programme.

Throughout that year and prior to the implementation of the CLIL programme the school undertook a preparatory period which, stage by stage, involved the three distinct parties with a major role in the programme:

- Stage I was devoted to the School Board: the Head teacher and the Language Coordinator. At this stage decisions were made as to the design of the implementation programme. It was to affect Grades 3 and 5 (8- and 11-year-olds respectively) with CLIL lessons in Science.
- Stage II included the families: a lecture was given and a leaflet was issued with answers to the most frequently asked questions in relation with CLIL.
- Stage III was addressed to the teachers and it involved an extensive 30-h Teacher Education Programme over 1 year. Twelve primary class teachers, four primary English teachers, and four secondary teachers (two specialised in EFL and two in Science) took part. The course was centred on developing strategies for fostering learners' listening and speaking abilities, unit design and lesson planning, and, finally, assessment.

Thus, the model adopted by the school constitutes an unusual case of fruitful collaboration between research experts and school administrators and practitioners taking place in Europe (Escobar Urmeneta and Pérez-Vidal 2004).

3.2 The Linguistic Impact of the CLIL Programme

In order to analyse the linguistic impact of the CLIL programme described above, the present study collected data from the first cohort of learners on the Science CLIL programme in Grade 7 and compared them with learners who had not been involved

in the programme. They both followed the conventional official curriculum in which EFL is taught as formal instruction (FI). That is, the CLIL group (Group A), follows FI and in parallel CLIL instruction (FI+CLIL), and hence it receives some ‘extra’ hours which are CLIL hours. The non-CLIL group (Group B) follows an FI only programme. The combination of FI and CLIL in parallel is the current arrangement in most CLIL programmes in Barcelona.

The study addresses the following research question: When contrasting a group experiencing FI in combination with CLIL, and a group experiencing FI only, which programme results in linguistic benefits, if at all, and which skills benefit the most, if any? On the basis of the review of the literature presented above, we establish the hypothesis that the group in the FI+CLIL programme will improve significantly more than the FI group, and the receptive skills to a larger degree than the productive skills.

3.3 Participants

Participants were 2 groups of Catalan/Spanish bilingual EFL learners for which English was their L3. Group A ($N=50$) was the experimental group experiencing the FI plus CLIL, so they are the FI+CLIL group (from now on GA: FI+CLIL group). Group B is the control group ($N=50$) experiencing only FI, so they are the FI group (from now on GB: FI group). There are 50 % of males and females in each group.

Having been together in the same school since nursery, both groups had started learning English at the age of 5/6 (Nursery), hence shared the age of onset of instruction (AoI). Data collection started when at the end of their first year of secondary education (Grade 7) at the age of 13. They had both therefore had 8 years of FI. However, GA: FI+CLIL had received 3 years of the extra CLIL hours from the age of 10 years (Grade 5). In order to make comparisons possible, GA: FI+CLIL was not matched for age with GB: FI, which would have created a disadvantage in terms of time of exposure to English, but for total number of hours of exposure. Consequently, this entailed that the latter group included learners who were a year older than the former, as Table 1 displays.

Table 1 Participants ($N=50$)

AoI in English	Data collection	
	T1	T2
GA: FI+CLIL	Grade 7 (12/13 years)	Grade 8 (13/14 years)
FI: Nursery (5/6 years)		
CLIL: Grade 5 (10/11 years)		
GB: FI	Grade 8 (13/14 years)	Grade 9 (14/15 years)
FI: Nursery (5/6 years)		

3.4 Design and Rationale of the Study

The study has a longitudinal pretest-posttest design as Table 2 below shows. Both groups of learners were measured respectively before and after one academic year in order to tap into gains obtained over the course of that year. Then, as their respective accumulated hours of exposure to English were very similar at the first data collection time (T1), although for GA: FI+CLIL some of the hours were CLIL hours, the difference in gains obtained by each group over that year was calculated. The quantity of hours being similar and the quality being different, any contrasts in the gains obtained by each group over a year treatment was expected to reveal whether or not CLIL hours have a significantly higher positive effect on learners' linguistic progress than non-CLIL hours of FI.

GA: FI+CLIL learners were measured in secondary when they were 13 (pretest) and 14 (posttest) years old at the end of Grades 7 and 8 respectively. They had had altogether 8 years of FI and 3 years of CLIL when data were collected for the first time (T1), and 9 and 4 respectively when data were collected a second time (T2). GB: FI learners were measured in Grade 8 and 9 when they were 14 (pretest) and 15 (posttest) years old respectively, also at the end of each academic year. They had had altogether 9 years of FI when at the first data collection time (T1), and 10 years of FI when data were collected second time (T2).

Table 2 below displays the accumulated number of hours of English at T1 and T2 for each group. In the case of GA: FI+CLIL, at T1 data collection, in addition to 1,120 h of FI (approximately 140/year since Nursery) they had had 3 years of CLIL, hence a total of 210 CLIL hours (70/year). Their total exposure to English was 1,330 h. One year later, at T2, GA had had 1,260 h of FI and 280 h of CLIL, that is 1,540 h in total. GB: FI, at T1 data collection, had had 1,260 h of FI (approximately 140/year since Nursery), and at T2, 1,400 h.

In order to assess the differential degree of gain between both groups, GA: FI+CLIL gains between T1 and T2 are compared with gains by GB: FI, the control group. The design allows for a between-groups comparison of the effect of a relatively similar amount of hours of instruction: 210 h (140 FI + 70 CLIL) in GA versus 140 h (FI) in GB.

Table 2 Design

	T1	T2
GA:	Grade 7 (12/13 years)	Grade 8 (13/14 years)
FI+CLIL	FI: 1,120 h + CLIL: 210 h = 1,330 h	FI: 1,260 h + CLIL: 280 h = 1,540 h
GB: FI	Grade 8 (13/14 years)	Grade 9 (14/15 years)
	FI: 1,260 h	FI: 1,400 h

3.5 *Instruments and Data Collection Procedures*

Data were elicited from intact class groups in an exam-like situation, both for productive and receptive skills. Production was elicited in writing, and reception in writing and orally. In addition, lexico-grammatical abilities were also tapped into. The instruments used to obtain the data were: (i) a composition on a given topic measuring written production; (ii) a reading task (cloze) and a dictation measuring written and oral comprehension; (iii) a sentence transformation test and a grammaticality judgement test with progressive degrees of difficulty in multiple choice format measuring lexico-grammatical ability. Data collection in two 1-h sessions was handled by the class teachers due to institutional conventions. It took place in an exam-like situation.

3.6 *Analysis and Measures*

Different procedures were used for the analysis of the data gathered. The reading task, the dictation, the grammar and the grammaticality judgement tests were straightforwardly corrected using objective criteria with a correcting profile. The data obtained from the writing test were transcribed using the CLAN programme. They were then analysed quantitatively for lexical and syntactic complexity, fluency and accuracy features, as Table 3 shows (Wolfe-Quintero et al. 1998). The data were also analysed qualitatively following a rating scale (Friedl and Auer 2007) whereby task fulfilment, organisation, grammar and vocabulary features were measured. Results were introduced to a Stats Graphic matrix, and the formulae for each ratio were calculated. Finally, mean results for all measures per group were drawn and compared with an ANOVA statistical analysis, the significance level set at <0.05 .

Finally, the frequency figure counting for correct/incorrect items was calculated per task. A final figure representing a general score was thus obtained for each task in order to calculate linguistic progress for each specific competence dimension analysed.

Table 3 Measures used to analyse written development

Quantitative measures	Syntactic complexity	Lexical complexity	Accuracy	Fluency
Qualitative measures	Task fulfilment	Organisation	Grammar	Vocabulary

4 Results and Discussion

The results of this study for the measures presented above are displayed in Table 4. The degree of significance has been set at .05. The T2 column for each group displays amount of gains (+) or losses (-). Results reveal that both groups improve over the course of the year under study, between T1 and T2, not surprisingly, in all measures but written fluency, in which they both even lose. GB also suffers losses at T2 in written syntactic complexity, that is, amount of subordination used. However, when the degree of gains achieved by each group is compared, it is only GA, and not GB, which exhibits levels of improvement. These gains are significantly higher than GB's in three out of the four domains of competence gauged, affecting the abilities of writing, reading and lexico-grammatical competence.

Concerning the results skill by skill, GA's performance in the written composition task yields significantly larger gains for accuracy than GB's ($F[1,196]=4.41$, $p=0.037$), and a tendency towards higher use of subordination ($F[1,196]=0.25$, $p=0.6201$), whereas GB tends to show larger gains in vocabulary ($F[1,196]=0.69$, $p=0.406$). They both show losses rather than gains in written fluency, higher in the case of GB ($F[1,196]=0.08$, $p=0.7801$). As for the qualitative measures of written production, GA also outperformed GB, even in vocabulary, in contrast to the already mentioned quantitative measures ($F[1,196]=2.37$, $p=0.1256$). However, no qualitative results reached statistical significance. GB seems to only make a larger improvement than GA as far as listening is concerned, and in one written measure, lexical complexity, albeit not significantly either.

Turning to results related to reading comprehension, as tested by means of a cloze test, they reveal that GA gained significantly more than GB over the course of a year ($F[1,98]=5.14$, $p=0.0255$). When focusing on listening comprehension, as tested by

Table 4 Mean values

	GA: FI+CLIL		GB: FI	
	T1	T2	T1	T2
Writing				
Syntactic complexity	0.40	0.39 (+0.01)	0.47	0.49 (-0.02)
Lexical complexity	6.50	6.71 (+0.21)	6.31	6.73 (+0.42)
Accuracy	0.120	0.078 (+0.042*)	0.092	0.086 (+0.006)
Fluency	146.2	145.1 (-1.1)	149.1	144.7 (-4.4)
Task Fulfilment	2.92	3.29 (+0.37)	2.63	2.87 (+0.24)
Organisation	2.84	3.24 (+0.4)	2.49	2.76 (+0.27)
Grammar	2.40	3.06 (+0.66)	2.34	2.70 (+0.36)
Vocabulary	2.52	3.18 (+0.66)	2.53	2.74 (+0.21)
Reading	14.3	16.1 (+1.69*)	14.6	14.8 (+0.22)
Listening	109.4	112.2 (+2.8)	109.7	112.7 (+3.1)
Grammar (Lexico-grammatical ability)	37.1	39.8 (+2.72*)	38.5	38.8 (+0.3)

Note: The higher the value for syntactic complexity and accuracy, the lower the competence level
T1 first data collection time, *T2* second data collection time

means of a dictation given by the teacher, our results yield no significant differences between both groups ($F[1,198]=0.01, p=0.924$). In fact, they both showed improvement at T2, however, and contrary to what we had hypothesized, GB also presented a tendency towards higher results than GA. Finally, when turning to the last linguistic domain scrutinized, grammar results, as tested through a fill-in-the-gaps task and an error correction task, results related to lexico-grammatical ability again indicated that GA's performance was significantly better than GB's ($F[1,98]=7.39, p=0.0078$).

In sum, GA consistently produced significantly more accurate texts and grammar manipulation tasks, and read significantly better. Their written texts tended to use more subordination, be better organized, lexically richer and more purposeful. However, they were less fluent. GB showed a tendency towards better listening abilities and use of lexis when measured quantitatively. They improved in the rest of the measures, but less than GA, lost ground in fluency, like GA, and, contrary to GA, also lost in the use of subordination. GA's significant leap forward in accuracy and general lexico-grammatical ability is relevant since, as already mentioned, CLIL courses are thought to focus on meaning rather than on form. In this respect, only the extra amount of practice or transfer of skills can explain these results, as is discussed further below.

Looking at these differences in greater detail, in written accuracy, GA shows a 0.042 progress over one academic year. This is significantly higher than GB's results, which only improved 0.006 from T1 to T2 ($F[1,196]=4.41, p=0.037$). As for reading, GA obtained a 1.69 figure, which is significantly higher than GB's 0.22 improvement ($F[1,98]=5.14, p=0.0255$). In the case of listening, GA's progress reached a value of 2.8 whereas GB's progress reached 3.1, but the difference is not statistically significant. This, together with lexical complexity in writing ($F[1,196]=0.69, p=0.406$), is the only areas in which GB shows a tendency to outperform GA, as already noted.

In the light of these findings, we can address the hypothesis in the study. Our results show that the CLIL programme seems to lead learners to improve significantly more than non-CLIL learners in their abilities to write more accurate and syntactically complex texts, and to generally improve in the whole set of qualitative measures (task fulfilment, organisation, grammar and vocabulary). Significantly higher improvement also accrues in their reading comprehension and lexico-grammatical competence. It is only in the domain of listening comprehension that GB tends to perform better than GA. These findings allow us to state that the second part of our hypothesis concerning the greater progress in receptive skills for the CLIL group is only partially confirmed. Indeed, whereas reading improves significantly, listening does not. Furthermore, our findings show a significant improvement in productive skills, whereas we had hypothesized they would lag behind those for receptive skills, as writing, and particularly accuracy, significantly progress. The same occurs with lexico-grammatical abilities. This is in contrast with findings published in previous studies and will be discussed further below.

Furthermore, although it is true that significant benefits do not accrue in all skills and measurements, it is also true that tendencies in the differential progress between both groups can allow us to establish the benefits of the school's CLIL programme. Where no benefits are found it can be argued that an academic year might not have been sufficient for learners to register more substantial benefits, and that a longer course of study

might eventually show that tendencies become significant differences. Hence we would posit that our results confirm the effectiveness of a CLIL programme.

Several general considerations concerning such general progress made by the CLIL group should be made here. First, when we review the research conducted in such settings, and more specifically in other bilingual contexts, such as Catalonia and the Basque Country, studies with detailed results for each skill seem to report similar findings to ours regarding productive skills (Muñoz and Navés 2007; Lasagabaster 2008; Ruiz de Zarobe 2008; Villareal and Gacía-Mayo 2009; Pérez-Vidal and Juan-Garau 2010). This is in contrast to other studies from Europe, as CLIL students in Spain tend to show an improvement not only in receptive but also in productive skills. It is interesting to highlight that in Lasagabaster's study (2008), as in ours, younger CLIL groups also scored lower than 1 year older non-CLIL groups in the listening tests (in the present study CLIL learners scored lower than FI learners for the listening ability not only when they were younger but also when both groups shared the same age).

Second, it has to be noted that the CLIL group has more hours of exposure. However, it is of utmost importance to realise that in spite of GA having a few more hours (70) than GB, when measured at T1 GA did not always outperform GB. For example, while it is true that, as far as written competence goes, the former started at a higher onset level in the domain of lexical complexity, task fulfilment organisation and, in contrast, grammar, they had a lower onset level in the domain of accuracy, vocabulary and fluency, just as in the domains of reading, listening comprehension and lexico-grammatical ability. Hence, it could be argued that in those domains in which GA is lower at T1, quantity of hours is not what matters but other factors such as quality, readiness to learn, and motivation, among others. Another possible explanation would be the maturational constraints of GA, as the group is a year younger than GB, an issue which is beyond the scope of this chapter (but see Muñoz 2015 on this issue). What is interesting is that even in some of those domains in which GA had lower onset levels, such as reading and lexico-grammatical ability, they still outperformed GB at T2, after 70 extra CLIL hours plus 140 FI hours.

We now turn to a different set of considerations concerning the specific language skills analysed. We will first address the issue of accuracy and lexico-grammatical ability. The significant improvement found in the area of accuracy in the writing skill and in lexico-grammatical abilities is a rather surprising finding. Opposite results were obtained by the empirical studies carried out in Canada and Europe. In Canada, this led to a concern for fostering accuracy, as proposed by Harley et al. (1990) and more recently Lyster (2007). More specifically, these authors have proposed balancing the experiential and analytical approaches, that is, introducing approaches that focus on form in order to compensate for the low level in accuracy. Therefore, the fact that accuracy in the writing skill and lexico-grammatical abilities in general showed significant improvement in the case of our CLIL participants might be explained by transfer of knowledge and skills from a FI context to a CLIL context, since they are "often" and "very often" practised in the FI context. This idea is further developed below (see Table 5).

We would now like to suggest an interpretation of our results in the light of the theories related to the role of practice and skill transfer models. As regards the issue

of transfer of knowledge and skills, there are two main differences between the context on focus in our study, CLIL, and the contrasting context, FI, CLIL being nearer a natural context than FI (see Pérez-Vidal 2015 for a detailed discussion). One is the type and amount of input learners are exposed to in one and the other, the second is the type of skills practice that learners engage in. Regarding the former, we must remember that our setting is one where little input exposure can be expected to be available outside the classroom walls, hence the additional CLIL hours are quantitatively important. Additionally, CLIL's qualitative differences with FI concerning meaningfulness of interaction and authenticity of topics and materials are also key for language development. Regarding the latter, the study of practice in SLA literature has been recently retackled, especially with DeKeyser's (2007a) monographic book on practice, claiming that not only the amount of practice but also the type is crucial to language learning. Previous studies on practice had assumed a dual division between input practice and output practice. Two confronted positions have developed over the years on this issue: VanPatten and colleagues, defending the position within the input processing studies that comprehension practice alone is enough to bring about significant development, not only in comprehension but also in production (vanPatten and Cadierno 1993), and the skill-specificity theory approach, represented by DeKeyser and Sokalski (1996) and DeKeyser (2007b), which replicated vanPatten and Cadierno's (1993) study and reached the conclusion that "input practice is better for comprehension and output practice for production" (DeKeyser and Sokalski 1996: 635). Thus, adopting the latter view, we can expect that in learning contexts where sufficient input practice is provided, comprehension skills (both reading and listening) will improve after a certain period of time. What seems not so straightforward is whether production skills (speaking and writing) will also improve in learning contexts where only comprehension practice is provided (with limited production practice) such as CLIL contexts. Hence we have to resort to a different explanation for our results, that provided by transferability of practice further below.

In our research study, reportedly each of both contexts allows different patterns of language skills practice. As Table 5 below displays, in FI writing and reading skills are often practiced, at least once a week, just as lexico-grammatical abilities, practiced often in every single class session. Listening is seldom practiced, particularly

Table 5 Skill practice

Skill	CLIL context	FI context
Writing	<i>Seldom practised</i> (short exercises)	<i>Often practised</i> (at least once a week)
Reading	<i>Highly practised</i> (every class session)	<i>Often practised</i> (at least once a week)
Listening	<i>Seldom practised</i> (teacher talk)	<i>Seldom practised</i> (teacher talk)
Grammar (Lexico-grammatical ability)	<i>Very seldom practised</i> (once a month)	<i>Very often practised</i> (at least once a week)

bidirectional listening, only through teacher talk. Oral production practice is limited. In the CLIL context, whereas reading is practiced in every class session with a considerable amount of authentic texts unusual in FI, practice in listening and writing abilities is limited to teacher talk and very short exercises. Furthermore, lexicogrammatical abilities are hardly ever practiced.⁶

In addition to the impact of practice within contexts and in order to interpret our results for written production, we should take into account the possibility of transferability of practice occurring in a particular context onto another. As GA in our study experiences a CLIL context together with a FI setting, their ability to transfer linguistic skills and competences learnt in the FI classes to the communication situations encountered in CLIL sessions might have been at play and foster improvement. This might explain why, although writing skills and lexicogrammatical abilities are hardly practiced in the CLIL sessions, GA participants obtain significantly better results than GB in these domains of competence. It could be argued that the amounts of writing and grammar practice typical of FI are used in the CLIL context and what students proceduralise in a FI context is automatized while in the CLIL setting (DeKeyser 2007b). That is, the accumulated experience of FI is what may play a major role in the relative benefits of an innovative or relatively unconventional and more naturalistic CLIL learning context such as the one enjoyed by the learners in this study.

Conclusions

Results obtained to answer the research question in this study confirmed the effectiveness of the CLIL programme, something which previous research had already shown. However, significant benefits did not accrue in all skills and measurements. Therefore, our hypothesis, which predicted that when contrasting the differential effects on learners' linguistic progress of the two programmes, the group in the FI+CLIL would improve significantly more than the FI group, especially in receptive skills, can be only partially confirmed. Reading, but not listening, improves significantly. Furthermore, our findings show significant improvement in productive skills on behalf of the FI+CLIL group. This is something we had not hypothesised, as writing, and particularly accuracy, significantly progresses. A similar situation occurs with lexicogrammatical abilities. This is in contrast with findings published in previous studies. Therefore, with the present study we have contributed to showing how, under CLIL conditions, certain aspects of language competence which did not seem to register clear gains in previous studies can also be developed. This would be the case for productive skills (writing), and formal aspects such as accuracy (also in writing) or lexicogrammatical abilities.

⁶Oral production is not contemplated in our study due to unavailability of data.

These results are in line with those from the COLE project reported in Chapters “[Testing Progress on Receptive Skills in CLIL and Non-CLIL Contexts](#)”, “[Writing Development Under CLIL Provision](#)”, and “[Does CLIL Enhance Oral Skills? Fluency and Pronunciation Errors by Spanish-Catalan Learners of English](#)” in Part II of this volume. Indeed, in COLE, receptive skills improve the most in the case of CLIL learners in contrast with results from non-CLIL learners, similarly to what the data presented in this chapter reveal. More specifically, reading comprehension improves to a greater extent than listening comprehension, and particularly with texts of a more specific kind. As for productive skills, COLE data reveal the greater effectiveness of the CLIL approach, in combination with formal instruction, with regard to overall written production, compared to formal instruction on its own, where significant progress is attained by the non-CLIL group only in accuracy. Although this is not in line with the research review made by Dalton-Puffer (2008), it mostly supports the findings by Lasagabaster (2008), Ruiz de Zarobe (2010) and our own.

When examining factors beyond learner linguistic progress, such as attitude, motivation and willingness to communicate, results from the COLE project, as reported in Chapters “[Exploring Affective Factors in L3 Learning: CLIL vs Non-CLIL](#)” and “[English Learners’ Willingness to Communicate and Achievement in CLIL and Formal Instruction Contexts](#)” in this volume, reflect that CLIL students tend to have more positive attitudes and beliefs towards English than their non-CLIL peers, albeit not significantly so. Their motivation to learn is also higher. However, it must be remembered that this is true even before the CLIL experience starts, most probably as a consequence of the fact that CLIL students are screened for good marks before entering the programme, as Chapter “[Exploring Affective Factors in L3 Learning: CLIL vs Non-CLIL](#)” clearly describes. The CLIL group also shows lesser anxiety and higher WTC, the latter being related to higher levels of achievement in EFL.

Taken together, the results presented in Part II of this volume point to a general beneficial effect of the CLIL programme over the non-CLIL programme, as it raises the level of learners’ ultimate attainment. Interestingly, it might be argued that, by being often first offered to the most advanced learners (or perceived as more attractive by them), it proves effective even before the programme begins! It has often been claimed that education does not present enough interesting and challenging opportunities for those learners who find themselves in the upper levels. CLIL does indeed provide such an opportunity for them, while at the same time, since the approach can be versatile in the hands of properly trained skilled teachers, it also provides fertile ground for lower level learners to make greater progress than on non-CLIL programmes (see Escobar 2004).

Acknowledgements This research received financial support through HUM2007-66053-C02-01/02, FFI2010-21483-C02-01/02, FFI2013-48640-C2-1/2-P and ALLENCAM (SGR2005-01086/2009-140/2014-1563) from the Spanish Ministry of Economy and Competitiveness and the Catalan Government respectively. Special thanks go to the school where data were collected, and to Cristina Escobar whose insights into CLIL have always been a source of inspiration.

References

- Ackerl, C. (2007). Lexico-grammar in the essays of CLIL and non-CLIL students: Error analysis of written production. *ViewZ (Vienna English Working Papers)*, 16(3), 6–11.
- Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation*, 12, 75–93.
- Alejo, R., & MacArthur, F. (2009). *Linguistic, cognitive and affective outcomes of CLIL experiences: The case of bilingual education programmes in Extremadura (Spain)*. Paper presented at the Gurt conference 2009: Implicit and explicit conditions, processes, and knowledge in SLA & Bilingualism. Washington, DC: Georgetown University.
- Allen, P., Swain, M., Harley, B., & Cummins, J. (1990). Aspects of classroom treatment: Towards a more comprehensive view of second language education. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 57–81). Cambridge: Cambridge University Press.
- Amengual-Pizarro, M., & Prieto-Arranz, J. L. (2015). Exploring affective factors in L3 learning: CLIL vs non-CLIL. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 197–220). Berlin: Springer.
- Brinton, D. M., Snow, M. A., & Bingham Wesche, M. (1989). *Content-based second language instruction*. Cambridge, MA: Newbury House.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39, 523–532.
- Cenoz, J., & Genesee, F. (1998). *Beyond bilingualism: Multilingualism and multilingual education*. Clevedon: Multilingual Matters.
- Cenoz, J., & Jessner, U. (2000). *English in Europe: The acquisition of a third language*. Clevedon: Multilingual Matters.
- Collentine, J., & Freed, B. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26(2), 153–171.
- 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.
- Dafouz, E., & Guerini, M. (2009). *CLIL across educational levels* (pp. 3–16). Madrid: Santillana Educación/Richmond publishing.
- Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning*. Amsterdam: John Benjamins.
- Dalton-Puffer, C. (2008). Outcomes and processes in content and language integrated learning (CLIL): Current research from Europe. In W. Delanoy & L. Volkman (Eds.), *Future perspectives for English language teaching* (pp. 139–157). Heidelberg: Winter.
- Dalton-Puffer, C. (2011). Content and language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, 31, 182–204.
- Dalton-Puffer, C., Hüttner, J., Jexenflicker, S., Schindelegger, V., & Smit, U. (2008). *Content and language integrated learning an Österreichs Höheren Technischen Lehranstalten. Forschungsbericht*. Vienna: Universität Wien & Bundesministerium für Unterricht, Kultur und Kunst.

- DeKeyser, R. M. (2007a). *Practice in a second language. Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.
- DeKeyser, R. M. (2007b). Skill acquisition theory. In J. Williams & B. VanPatten (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–113). Mahwah: Lawrence Erlbaum.
- DeKeyser, R. M., & Sokalski, K. (1996). The differential role of comprehension and production practice. In R. Ellis (Ed.), *Form-focused instruction and second language learning* (pp. 81–112). Oxford: Blackwell Publishers.
- Escobar Urmeneta, C. (2004). Content and language integrated learning: Do they learn content? Do they learn language? In J. M. Oro Cabanas, J.D. Anderson, & J. Varela ZapataL (Eds.), *Linguistic perspectives from the classroom: Language teaching in a multicultural Europe* (pp. 27–38). Santiago de Compostela: Universidad de Santiago.
- Escobar Urmeneta, C., & Pérez-Vidal, C. (2004). Teacher education for the implementation of a content and language integrated learning approach (CLIL) in the school system. In B. Wilkinson (Ed.), *Integrating content and language. Meeting the challenge of multilingual education* (pp. 402–415). Maastricht: Universitaire Pers Maastricht.
- European Commission. (1995). *White paper on education and learning* 449. Brussels: European Commission.
- Eurydice. (2006). *Content and language integrated learning (CLIL) at school in Europe. Spain national description 2004/05*. Brussels: Eurydice.
- Eurydice. (2008). *Key data on teaching languages at school in Europe. Special Eurobarometer 243*. Brussels: European Commission.
- Friedl, G., & Auer, M. (2007). Rating scale used for assessment of the writing task. *Erläuterungen zur Novellierung der Reifeprüfungsverordnung für AHS, lebende Fremdsprachen*. BIFIE. Wien – St. Pölten. (Gültig ab dem Sommertermin 2009) <http://www.bifie.at/publikationen>
- Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Writing development under CLIL provision. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 139–161). Berlin: Springer.
- Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Cambridge: Newbury House.
- Genesee, F. (1994). *Integrating language and content: Lessons from immersion*. Educational practice report II. Washington, DC: Center for Research on Education Diversity and Excellence (CREDE).
- Genesee, F. (2013). Insights into bilingual education from research on immersion in Canada. In C. Abello-Comtesse, P. Chandler, M. D. López-Giménez, & R. Chacón-Beltrán (Eds.), *Bilingual and multilingual education in the XXI century* (pp. 24–41). Clevedon: Multilingual Matters.
- Grenfell, M. (Ed.). (2002). *Modern languages across the curriculum*. London/New York: Routledge Falmer.
- Harley, B., Allen, P., Cummins, J., & Swain, M. (1990). *The development of second language proficiency*. Cambridge: Cambridge University Press.
- Hellekjaer, G. (2010). Language matters: Assessing lecture comprehension in Norwegian English-medium higher education. In C. Dalton-Puffer, T. Nijula, & U. Smit (Eds.), *Language and use and language learning in CLIL classrooms* (pp. 253–258). Amsterdam: John Benjamins.
- Jexenflicker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nijula, & U. Smit (Eds.), *Language and use and language learning in CLIL classrooms* (pp. 169–189). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Johnson, K., & Swain, M. (1997). *Immersion education: International perspectives*. Cambridge: CUP.
- Juan-Garau, M., & Salazar-Noguera, J. (2015). Learning English and learning through English: Insights from secondary education. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 105–121). Berlin: Springer.
- Juan-Garau, M., Prieto-Arranz, J. L., & Salazar-Noguera, J. (2015). Lexico-grammatical development in secondary education CLIL Learners. In M. Juan-Garau & J. Salazar-Noguera (Eds.),

- Content-based language learning in multilingual educational environments* (pp. 179–195). Berlin: Springer.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1, 31–42.
- Lasagabaster, D. (2015). Different educational approaches to bi- or multilingualism and their effects on language attitudes. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 13–30). Berlin: Springer.
- Lasagabaster, D., & Ruiz de Zarobe, Y. (Eds.). (2010). *CLIL in Spain: Implementation, results and teacher training*. Newcastle: Cambridge Scholars Publishing.
- Linares, A., Morton, T., & Whitakker, R. (2013). *The roles of language in CLIL*. Cambridge: Cambridge University Press.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections. Evaluation Project. *Applied Linguistics*, 20(11), 1–25.
- Lyster, R. (2007). *Learning and teaching languages through content. A counterbalanced approach*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Menezes, E., & Juan-Garau, M. (2015). English learners' willingness to communicate and achievement in CLIL and formal instruction contexts. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 221–236). Berlin: Springer.
- Moore, P. (2009). *On the emergence of L2 oracy in bilingual education: A comparative analysis of CLIL and mainstream learner talk*. Sevilla: Universidad de Sevilla. Unpublished doctoral dissertation.
- Muñoz, C. (2015). Time and timing in CLIL: A comparative approach to language gains. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 87–102). Berlin: Springer.
- Muñoz, C., & Navés, T. (2007). Spain. In A. Maljers, D. Marsh, & D. Wolff (Eds.), *Windows on CLIL. Content and language integrated learning in the European spotlight* (pp. 160–165). Graz: European Platform for Dutch Education and European Centre for Modern Languages.
- Nikula, T. (2007). Speaking English in Finnish content-based classrooms. *World Englishes*, 26, 206–223.
- Pérez-Vidal, C. (2009). The integration of content and language in the classroom: A European approach to education (The second time around). In E. Dafouz & M. C. Guerrini (Eds.), *CLIL across educational levels: Experiences from primary, secondary and tertiary contexts* (pp. 3–16). Madrid: Santillana Educación/Richmond Publishing.
- Pérez-Vidal, C. (2013). Perspectives and lessons from the challenge of CLIL experiences. In C. Abello-Comtesse, P. Chandler, M. D. López-Giménez, & R. Chacón-Beltrán (Eds.), *Bilingual and multilingual education in the XXI century* (pp. 52–82). Clevedon: Multilingual Matters.
- Pérez-Vidal. (2015). Languages for all in education: CLIL and ICLHE at the crossroads of multilingualism, mobility and internationalisation. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 31–50). Berlin: Springer.
- Pérez-Vidal, C., & Juan-Garau, M. (2010). To CLIL or not to CLIL? From bilingualism to multilingualism in Catalan/Spanish communities in Spain. In Y. Ruiz de Zarobe & D. Lasagabaster (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 115–138). Newcastle: CUP/Cambridge Scholars.
- Prieto-Arranz, J. I., Rallo Fabra, L., Calafat-Ripoll, C., & Cattrain-González, M. (2015). Testing progress on receptive skills in CLIL and non-CLIL contexts. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based language learning in multilingual educational environments* (pp. 123–137). Berlin: Springer.
- Rallo Fabra, L., & Jacob, K. (2015). Does CLIL enhance oral skills? Fluency and pronunciation errors by Spanish-Catalan learners of English. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-*

- based language learning in multilingual educational environments* (pp. 163–177). Berlin: Springer.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque country. *International CLIL Research Journal*, 1(1), 60–73.
- Ruiz de Zarobe, Y. (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). Berlin: Springer.
- Swain, M. (2000). French immersion research in Canada: Recent contributions to SLA and applied linguistics. *Annual Review of Applied Linguistics*, 20, 199–212.
- Sylvén, L. K. (2004). *Teaching in English or English teaching? On the effects of CLIL on Swedish learners' incidental vocabulary acquisition*. PhD dissertation, Göteborg University.
- Van de Craen, P., Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in primary schools: Teaching maths in a CLIL environment. In D. Marsh & D. Wolff (Eds.), *Diverse contexts: Converging goals – CLIL in Europe* (pp. 185–200). Frankfurt am Main: Lang.
- VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225–241.
- Villarreal Olaizola, I., & García Mayo, M. P. (2009). Tense and agreement morphology in the interlanguage of Spanish/Basque bilinguals. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from the research in Europe*. Bristol: Multilingual Matters.
- Wesche, M. (2002). Early French immersion: How has the original Canadian model stood the test of time? In P. Burmeister, T. Pieske, & A. Rhode (Eds.), *An integrated view of language development: Papers in honor of Henning Wode*. Trier: Wissenschaftlicher Verlag Trier.
- Whittaker, R., & Llinares, A. (2009). CLIL in social science classrooms: Analysis of spoken and written productions. In Y. RuizdeZarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from the research in Europe* (pp. 215–234). Clevedon: Multilingual Matters.
- Wilkinson, B. (2004). *Integrating content and language. Meeting the challenge of multilingual higher education*. Maastricht: Maastricht University.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. Hawai'i: University of Hawai'i at Manoa.
- Wolff, D. (2002). Germany. In M. Grenfell (Ed.), *Modern languages across the curriculum* (pp. 78–87). London/New York: Routledge Falmer.
- Zydatiss, W. (2007). *Deutsch-Englische Züge in Berlin (DEZIBEL). Eine Evaluation des bilingualen Sachfachunterrichts in Gymnasien: Kontext, Kompetenzen, Konsequenzen*. Frankfurt am Main: Peter Lang.
- Zydatiss, W. (2012). Linguistic thresholds in the CLIL classroom? The threshold hypothesis revisited. *International CLIL Research Journal*, 1(4), 16–28.