European Journal of Foreign Language Teaching
ISSN: 2537 - 1754
ISSN-L: 2537 - 1754
Available on-line at: www.oapub.org/edu

doi: 10.5281/zenodo.2652036 Volume 4 | Issue 2 | 2019

THE EFFECT OF CLIL PROVISION ON DEVELOPMENT OF RECEPTIVE SKILLS AMONG LOWER SECONDARY SCHOOL STUDENTS AND THEIR USE OF METACOGNITIVE STRATEGIES

Marzena Foltyn[i]
University of Bielsko-Biala,
Poland

Abstract:
This research paper intends to examine the issue of CLIL's metacognitive dimension and its probable positive influence on receptive language skills among 51 teenagers in Polish lower secondary school in Bielsko-Biala. Peculiarly, it attempted at addressing the question whether there is a significant difference in receptive skills absorption and ones' use of metacognitive strategies between the two following experimental conditions, (CLIL-based approach applied in English classroom), compared to control condition (traditional EFL exposure to English). In order to investigate that, quasi-experimental design was applied. Having collected the data through administration of reading and listening pre-tests and post-tests and questionnaires, I was able to draw a conclusion that the experimental group (CLIL students) outperformed their non-CLIL peers and was more aware of metacognitive receptive strategies. It must be admitted that CLIL-based lesson has superiority over traditional approach in terms of raising students' awareness of their own learning and general receptive performance. Nevertheless, carrying out the experiment and analyzing the data from the statistical point of view also demonstrated that the progress made by the EFL group appeared to have been more recognizable.

Keywords: CLIL, content and language integrated learning, metacognitive strategies, receptive skills

1. Introduction

Introduction of the holistic concept of knowledge to formal education requires the creation of the teaching method that would combine various subject learning areas and, at the same time, support social, psychological, physical and cognitive development of the learners. In response to these needs, a teaching method was developed so as to promote the integration of diversified approaches. It can be characterized by the

[i] Correspondence: email marzenafn@gmail.com

synthesis of both linguistic components and cognitive skills. In didactics, this flagship example of the language teaching methodology, based upon holistic approach to education, is known as Content and Language Integrated Learning (CLIL). With the expanding significance of multilingualism across the globe, this approach has been conquering Europe exponentially for over 20 years. In Poland, there has also been a visible line of the empowering CLIL outcome. Thereupon, it is of high validity to uncover the term Content and Language Integrated learning (CLIL) and put it into context before the analysis of research material is to be executed. As Widdowson (2003: 3) claims:

> "*Once you have identified the idea about language or language learning that lies behind a particular classroom activity, then you are in a position to make a judgment about how valid it is*."

The acronym CLIL standing for Content and Language Integrated Learning is "*an innovative methodology that has emerged to cater to this new age*" (Mehisto, et.al. 2008). As Papaja (2014: ix) curiously enough points out, the main concept lying behind the Content and Language Integrated Learning is that not only it does introduce a wider cultural context, prepare the learner for international activities and exchanges, give access to international certification, develop multilingual interests but also increase one's motivation to learn a second or even a third language. The advantageous effects the approach brings might be particularly assigned to the so-called cognitive consequences, which result in greater metacognitive awareness and cross linguistic influence. With the link to Dalton-Puffer (2007), CLIL teachers are meant to ask metacognitive questions. At the beginning, the term "metacognition" simply indicates the knowledge about one's learning (Sajda 2008: 13). With the accordance to this definition, such questions mainly focus on making students reflect on their way of thinking and their consciousness (Llinares and Morton 2017: 223). As far as one of the CLIL's 4Cs methodology, cognition, is concerned, it consistently influences the model of knowledge creation, what in further development, results in metacognitive implications. According to Łyda and Szcześniak (2014: 193), this methodology elicits range of activities involving, mostly, reading. Moreover, most of them are goal-oriented so as to make students explore meaning and sense. Furthermore, such activities enable students to expand their content knowledge. Here, reading as such is the most complicated constructive process. Łyda and Szcześniak (2014: 194) underline that CLIL evokes strategic and dynamic reading. Thus, due to the fact that term "metacognitive awareness" is one's knowledge of cognitive functions, it might be related to the reading process and the reason that CLIL students are well-skilled readers. The before mentioned factors also enhance the process of listening comprehension. CLIL students, being exposed to wider content knowledge and meaning, sense-oriented tasks, are more likely to develop awareness of their own learning. To conclude, "metacognitive awareness" has been widely recognized among the individuals who are exposed to the

more linguistic resources than the mere monolinguals (Cenoz and Gorter 2011). Such an implication might have an additive positive influence on the whole process of second language acquisition. Dalton-Puffer (2007) also proposes the below depicted table which poses the representation of the particular useful or neutral implications of CLIL in the specific field of different language skills:

Table 1: Language Competences affected or unaffected by CLIL

Favourably affected	**Unaffected or Indefinite**
Receptive Skills	Syntax
Vocabulary	Writing
Morphology	Informal/non-technical language
Creativity, risk-taking, fluency, quantity	Pronunciation
Emotive/affective outcomes	Pragmatics

Adapted from: Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe, Dalton-Puffer, 2008.

What of the current study, the receptive skills, being at the top in the rank of the favourably affected skills might be interpretable as the effect of offering "*additional reasons for reading"* the CLIL might impose (Dulton-Puffer 2007). It has been observed that such a language teaching approach allows students to achieve the competences in listening and reading almost like native speakers. As far as the productive skills are concerned, CLIL students often outperform their non-CLIL peers in fluency of speaking. Especially differences in risk-taking, creativity and quantity are widely recognizable (Naiman 1995 quoted in Dalton-Puffer et al. 2010: 193). A parallel effect has been noticed in terms of morphology. Hence, not only low-level processes are used appropriately, but also the technical and semi-technical stock of words is much greater. Accordingly, the researchers claim that CLIL students perform better in academic language (Dalton-Puffer 2007). CLIL's implication might be also found in the most influential theories on Second Language Acquisition. Hence, Wilhelmer (2008: 21-32) attaches Krashen's "comprehensible input+1" to the provisions of CLIL. Such an input hypothesis and its idea of "i+1" represents the background knowledge the students have, plus, something beyond. Thus, this theory not only had been influential as a theoretical background in CLIL but also among many bilingual programmes such as immersion. Importantly, as far as CLIL offers the authentic context encouraging "acquistion" rather than "learning", Krashen's hypothesis is highly related to this distinction by focusing on unconscious, unintentional learning. When it comes to the "i+1", it might be realized through extra-linguistic information that the students are provided with in CLIL classrooms. As follows, alluding to the core of this work, the "comprehensible input" boost the improvement of comprehensions skills but not production. Working on this premise, the potential benefits to receptive skills might be recognized clearly. This claim has been proven by Genesee (1994) who reported such a relation in French immersion programmes in Canada. Although the research in CLIL area remains scanty, the more activation of passive skills than productive ones during this program may serve as the predictor of CLIL students success in the former

competences (Coonan 200: 627). These skills also might be advantageous in CLIL program due to the role of the teacher who appears to be a subject-matter teacher, producing language to convey the content which, subsequently, does facilitate comprehension (Costa 2016). Along these lines, the present study examines the background elements of CLIL and questions whether the above mentioned predictors of one's success is accurate or not among the lower secondary students in Bielsko-Biała, Poland. As far as the CLIL approach supports the development of cognitive functions and makes the learner the strategic participant of educational process, the current research is aimed at establishing the relationship between CLIL, one's level of receptive skill comprehension and their use of metacognitive strategies.

2. Literature Review

The topic of CLIL in teaching a foreign language has been widely investigated in literature. Researchers were particularly interested in the development of students' cognitive functions and receptive skills with the use of bilingual teaching methods. The significance of the already conducted research in the present field of study was nonmeasureable so as to to improve the present findings.

The very first paper's aim was to examine whether there is a possible positive effect of CLIL on Spanish-Catalan learners or not in the field of receptive skills comprehension. As far as the CLIL's motivating and challenging aspects have been widely recognised, the author attempts to examine the potential impact on the receptive skills. Nevertheless, the available literature indicates the limited amount of the productive resources in this research area. This work investigates the development of listening and reading skills in two groups of participants. The experimental group consisted of L3-English CLIL students and the second one was comprised of students from state-run schools. The performance of the secondary students from Balearic Islands has been examined upon their performance in two listening and two reading tests. In data collection procedures, the before mentioned tests served as the research instruments. The very first listening test has been taken from the so-called "News Listening" category. Thus, in that kind of listening, the students were supposed to listen to four broadcast news and, as follows, choose the correct answer from the ten-item multiple choice test. The second one, appeared to test one's ability in "Picture Listening." In this instance, the informants were listening to ten recordings, but, were expected to choose the answers, in the multiple choice test, out of visually provided answers. In the case of reading comprehension, each tests covered 15 minutes. To begin with the "General Reading Test", it consisted of ten short stories which were of different subjects and genres. As far as the textual background varied from post-it notes not fire emergency instructions, their lengths ranged from 3 to 29 words. The final test involved the "Specific Reading Test." What of this test, it was concerned with the CLIL content subjects taught at each school. Importantly, the correct answers were checked with the

help of true/false test and matching the words from the text with the enclosed definitions.

Pladevall and Vallbona (2016), on the other hand, investigate the CLIL-based methodology in terms of its learning outcomes and effects it has on the comprehensive input of receptive skills in Swiss primary school context where the students are taught English as an EFL. Interestingly, the authors attempt at measuring the results by comparing two groups of young learners. The very first group is to be exposed to EFL hours only, while the other is given the same amount of regular English classes, but, with the addition of one CLIL-hour per week. In order to achieve the validity of the results, the test was administered at four times in two academic years. The results the researchers achieved appear to be of prime importance to the current study. The two groups of children from primary school were participating in regular EFL classrooms and in the ones CLIL-oriented. The author investigates the lessons over 20 months. To achieve the reliability of the research, the students were divided into level groups. Such a division enables investigators to measure whether the students' level of English increased in terms of receptive skills. The comparability was guarantee by different data collection. To investigate the given subject the longitudinal study was applied responding to 5th and 6th grades.

Another work examines the level of discrepancy between lower secondary EFL students and their CLIL peers in the reading comprehension. With regard to the present work, the author underscores the variation and the impact of the strategies the students use when completing the reading tasks. The author decided to apply the three following research methods: an IELTS reading test, qualitative interviews and quantitative survey. Such a choice was justified by the statement that the mixed method enables all of them to complement each other. As far as the students, their achievement and the strategy were of prime core, the researcher collected data from CLIL and EFL students and teachers. To start with the construction of surveys. they were aimed at finding the information about participants' background, motivation and experience with regard to reading tests and CLIL. The questionnaire consisted of 33 multiple-choice questions. To check one's strategy and understanding while reading, the author used the combination of questionnaires with interviews. Moreover, the interviews were semi-structured, thus, the questions are formulated before the interview and by the interviewer. Due to the fact, that both students and teachers were of research interest, there were two guided interviews made. All of the questions were open-ended and comprised of five sections. Namely, the interviewees were expected to answer the questions. According to the test results, the CLIL group had scored a considerably higher results than the other groups. What is more, the students taught in CLIL context appeared to be more motivated while taking a test. As far as the strategies are concerned, likewise, the CLIL participants were more likely to apply the reading strategies. In contrast, the administered test seemed to be of high difficulty to the EFL students who found it challenging.

Reading all of the articles occurred to be useful for the purpose of designing my own research. Having analyzed different research methods and procedures, I was able to plan how to conduct my own research study.

3. Material and Methods

There were 51 teenagers (aged 15) taking part in my research study. They were attending second grade of lower secondary school. The experimental groups were made of students speaking Polish as a native language and learning English through CLIL method. The control groups (N=26) consisted of students speaking Polish as a native language and learning English as a second language. Participants in the experimental group (N= 25) were learning English through its exposure of 5 hours a week. The content of the lesson was basing upon different field of studies. During the first year of their studies they were being taught advanced structure of English. The second grade, however, was characterized by introducing technical vocabulary in the area of History, Maths, Chemistry, Biology, IT and Civics so as to start learning all of these subjects in English during their final grade. The non-CLIL group was exposed to the same amount of English but with no subject content-oriented vocabulary being introduced. According to their English teacher, level of language of these both groups might be represented by CEFR's[ii] B1. At the end of the year, they are expected to achieve B2. Hence, having consulted it with the leading teacher, we decided to check their receptive skills comprehension through FCE sample papers. During my teaching practice in lower secondary grades, I devoted 2 stages (16 hours), with the three-months interval between each, for conducting my research study. The participants did not receive any special instruction on strategies. They participate in their course and practiced receptive skills exercises as they appeared during the lessons.

The research questions were the following:

1) Is there a significant difference between the level of general reading proficiency in English among CLIL and non-CLIL students?
2) How do CLIL and non-CLIL students differ with regard to their use of metacognitive reading strategies while reading texts in English?
3) Is there a significant difference between the level of general listening proficiency in English among CLIL and non-CLIL students?
4) How do CLIL and EFL students differ with regard to their use of metacognitive listening strategies?
5) Do both CLIL and non-CLIL students' progress at the same pace in English reading and listening comprehension skills while listening to texts in English?

There were two sample papers which were retrieved directly from First Certificate in English (Cambridge English official website). During the first stage of my study, the tests were distributed to both control and experimental groups. It shall not be forgotten that the participants' level of proficiency was chosen after having had a

[ii] CEFR - Common European Framework of References

consultation with the form master teacher. As far as the students' level was supposed to indicate upper-intermediate (B1), CEFR recommends this kind of tests as being appropriate to this level of proficiency. Both reading and listening comprehension tests consist of three parts examining different dimensions of these receptive skills and focus on general reading. Both tests and questionnaires were distributed immediately to students and each of them were given 90 minutes to complete the particular test and the questionnaire. To start with reading comprehension test, it is composed of thirty questions which are distributed in three parts of which each question indicates one point. In the first part, students were expected to read an extract from a novel and then answer to multiple-choice items. Its aim is to test one's ability to pay attention to details and feelings. The second part focuses on reading an article with seven parts being removed and the students' task is to fill them appropriately. Accordingly, this kind of test checks students' ability to use language for specific purposes and being able to fill the extract. The last part serves a magazine article about four people to be read. The participants were expected to match the given 15 sentences with the proper person. Again, the tasks of this dimension are supposed to make students read for gist, details and main points. When it comes to listening comprehension, correspondingly, the test was made of thirty questions being divided into three parts and of which one question may have served one mark. Listening to people talking in different situations with the multiple-choice items to be asked for constitutes the first part. Through careful listening and choosing the best answer of three suggested, the students' ability to recognize the core and focus on details are checked. The second part serves short text with the missing parts to be filled after listening to an interview. This part offers checking the ability to manipulate one's language skills and matching information according to content and context. The last part, accordingly to reading paper, focuses on hearing five different people and then finding the speaker whose situation applies to the given sentence. Being able to read the feelings, details and proper understanding of the speaker is tested in the previously mentioned part. The survey of Metacognitive Treading Strategies Inventory (MARSI) designed by Mokhtari and Reichard (2002) was used to collect data. The MARSI questionnaire examines to what extent students are able to apply the given reading strategies. As far as validity and reliability is concerned, according to Mokhtari (2002), it has appropriately been field-tested. When it comes to the strategies, they comprise three subscales: global reading strategies, problem solving strategies and support reading strategies. More specifically, the very first category focuses on how the students deal with making interferences while facing difficulties in comprehension. As far as global reading strategies are concerned, they are supposed to measure one's ability to monitor and cope with the text directly. The last but not least, support reading strategies consist of what can aid reading comprehension as such.

Table 2: MARSI Questionnaire – Subscales

MARSI subscales	MARSI statements
Global	1, 3, 4, 7, 10, 14, 17, 19, 22, 23, 25, 26, 19
Problem-solving	8, 11, 13, 16, 18, 21, 27, 30
Support	2, 5, 6, 9, 12, 15, 20, 24, 28

As for the structure, the MARSI questionnaire comprises 30 items along with a 5-point Likert Scale, in which 1 stands for "never", 2 stands for "occasionally", 3 means "sometimes", 4 stands for "usually", and 5 means "always or almost always". The maximum score of the Metacognitive Reading Questionnaire is 150, when each of the items is responded to with "Always" and, on the contrary, the minimum score counts 30 and the answers are "never". In order to expedite the process of completing the questionnaire, it was translated into Polish. The second questionnaire which was applied during first stage was the one prepared by Vandergrift, Goh, Mareschal and Tafaghodtari (2006). According to them, it was designed on the basis of Flavell's theory on metacognition. It might be divided into five sub-categories: person knowledge, planning and evaluation, directed attention and mental translation. It might be said that this kind of strategy is devoted to check one's perception while being engaged in listening task and their confidence in completing listening tasks. In details, the person knowledge strategies refer to one's level of certainty in listening comprehension tasks. With regard to problem solving strategies, they check if students can stop for a moment if they do not understand what they hear so as to find a reference point. On the other hand, translating the information into one's first language is checked in mental translation strategies. As far as planning and evaluation is concerned with pre- and post-listening activities that are performed in students' minds before and after listening. Finally, directed attention strategies are supposed to measure students' span of attention on specific tasks.

Table 3: MALQ Questionnaire -Subscales

MALQ subscales	MALQ statements
Planning and evaluation	1, 10, 14, 20, 21
Directed attention	2, 6, 12, 16
Person knowledge	3, 8, 15
Problem solving	5, 7, 9, 13, 17, 19
Mental translation	4, 11, 18

When it comes to the questionnaire's structure, it consists of 21 items. The students had to respond to each of them by rating their answers on a 6-point Likert Scale, in which 1 stands for "strongly disagree", 2 stands for "disagree", 3 stands for "partially disagree", 4 stands for "partially agree", 5 stands for "agree", and 6 stands for "strongly agree". Nevertheless, there were three questiones that need to be coded reversely: 16, 3, 8, 4, 11. The questionnaire was translated into Polish to avoid misunderstanding and facilitate ones' performances. The maximum score of the Metacognitive Listening Questionnaire is 126 when each of the items is responded to with "Strongly agree" and, on the

contrary, the minimum score counts 21 when the answers are "strongly disagree". The results of the surveys were tabulated individually and later on, according to the total of the participants, the average scores and standard deviation were calculated for each of the groups.

4. Results and Discussion

4.1 Reading

The data displayed below were analyzed with the use of T-test calculation. As far as there was one variable needed, the two CLIL groups are to be analyzed as the whole experimental group and the two EFL groups are considered to be one control group, so as to compare the achievements of the given individuals. The table below portrays the results of the T-test calculation considering number of participants, mean, standard deviation p-value on the basis of pre-and post- reading comprehension test.

Table 4: The results of reading comprehension tests (pre- and post-test)

Group	Mean		Gain scores	Standard Deviation		N
	Pre-reading test	Post reading -Test		Pre-test	Post-test	
CLIL group 1	23, 61	25, 15	1, 16	2, 59	1, 49	13
CLIL group 2	26,08	26, 66		2, 39	2, 46	12
EFL group 1	22, 71	25,05	1,88	2, 36	2,10	14
EFL group 2	21, 46	23, 75		1, 85	1, 91	12
P value	0, 0001	0, 022				

As it might be clearly seen, the mean achieved on pre-test by two CLIL groups was much higher than the mean gained by the EFL students. On the other hand, it should be noticed that the control groups were more heterogeneous as standard deviation equals 2,36 and 1,86 in contrast to standard deviation of experimental group as it counted correspondingly 2,59 and 2,39. It might be suggested that among CLIL students the differentiation between ones' levels of reading comprehension might have also affected the results. As far as the p-value equals 0,0001, the examination seems to be notably statistically significant. The table depicted above also constitutes the answer to research question number 5 of my study, namely: "Do both CLIL and non-CLIL students' progress at the same pace in English reading and listening comprehension skills?" It shall not remain underestimated that in spite of achieving much higher scores in pre- and post-reading comprehension tests, the gain scores of CLIL groups, equaling 1,16, are lower than the gain scores of their EFL peers, whose gain scores equal 1,88. The result of the study might suggest that if the students were performing slightly better at the beginning, their pace of progress remains rather stable and not as visible as the one in the EFL groups. To conclude, the EFL students' pace of progress appears to be more recognizable. Having said that, it must be admitted that the CLIL groups significantly excelled over their non-CLIL peers. Nevertheless, it shall not be forgotten that the pace

of their progress appeared to be slower than the one of EFL groups as their gain scores were decidedly higher.

Table 5: The results of Metacognitive Awareness Reading Questionnaire

PRE-MARSI				
MARSI subscales	**Experimental Group**		**Control Group**	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Global	3, 02	0, 36	2, 78	0, 28
Problem-Solving	3, 68	0, 37	2, 98	0, 39
Support	2, 85	0, 53	2, 43	0, 28
POST MARSI				
MARSI subscales	**Experimental Group**		**Control Group**	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Global	3, 20	0, 33	2, 88	0, 26
Problem-Solving	3, 73	0, 32	3, 08	0, 33
Support	2, 93	0, 54	2, 53	0, 34

The table above, on the other hand, portrays the mean and standard deviation of the pre- and post-Metacognitive Awareness Reading Questionnaire results of the experimental group consisting of CLIL students and control group which represents EFL participants. Having calculated the average and the level of homogeneity of the groups, the results are to be interpreted according to the scale prepared by Mokhtari and Reichard, the authors of the survey. Particularly, the overall average indicates how often students use and to what extent they are aware of reading strategies. The key is the following: 3.5 or higher stands for high awareness, 2.5-3.4 equals medium and 2.4 or lower indicates low awareness. First, the results of pre-MARSI questionnaires of two groups are to be discussed, and, then, the results of post-MARSI and changes between these two stages. To start with the more detailed analysis and the global strategies, the experimental group gained higher average than the participants from the control group. Nevertheless, it must be said that the results of both groups ranked themselves in *medium* scale, gaining respectively 3,02 and 2,78. There was also a difference between these two groups when it comes to their homogeneity. Standard deviation of the results of experimental group, equaling 0,36 indicates lower heterogeneity than among EFL students. As far as the problem-solving strategies are concerned, the results seem to be surprising. First of all, this category of strategies compares well with the two others. Control group scored mean equaling 2,98 what suggests, again, medium awareness of metacognitive reading strategies. On the contrary, the experimental group of which mean was 3,68 draws conclusion that the use of the before mentioned strategies is *high.* For the matter of the level of standard deviation among these two, their homogeneity remained similar. Curiously enough, the knowledge and use of support reading strategies were the lowest in both groups. However, once more, CLIL students seemed to be more aware of them as their mean equaled 2,93 and the mean of the EFL students was 2,53. The two before mentioned values are considered to stands for *medium* awareness of reading strategies. Notwithstanding, it might be simply observed that the

second group was more homogeneous than the first one. Having distributed the questionnaires after three months the results were the following. The experimental group, again, gained more scores in each category of metacognitive strategies. The means gained in global category by CLIL students was 3,20 and 2,78 by control group – what remained categorized as *medium.* Here, again, the EFL students seemed to be less heterogeneous than their CLIL peers. With regard to the problem-solving category of strategies, the experimental group outperformed the control group by 0,65. The extent of usage of this category appeared to be *medium* in both cases. Correspondingly, these two groups remained similar according to their homogeneity. For the matter of support strategies, the result was comparatively similar since the mean gained by experimental group was higher (2, 93) than the one of control group (2,53). More importantly, the group that excelled was less homogeneous than the one consisting of EFL students. As it is clearly seen in the figure above, all of these three items of the Metacognitive Awareness Reading Questionnaire increased their values. Among participant of the experimental group the mean of global category increased by 0,18 – and in the control group – by 0,10. In this particular item, the boost was the most recognizable. As far as the problem-solving strategies are concerned, the gain scores achieved by the experimental group was 0,05 and by the second group it equaled 0,10. The means of the last category of strategies was also enhanced. Having calculated the difference between the means of pre-MARSI and post-MARSI, CLIL students achieved gain scores equaling 0,08 and their non-CLIL peers, again – 0,1. To conclude, it shall be said that the CLIL group seemed to be more aware of the usage of metacognitive reading strategies. Nevertheless, it is of worth saying that the difference was insignificant. Moreover, there was a boost in the means value between the pre-MARSI and post-MARSI among the two mentioned groups. It might be justified by the fact that the school year continued and therefore, both the level of proficiency and awareness of the students could have excelled. Furthermore, it might be reasoned by their verification and reflection stage while completing the questionnaire for the second time. More interestingly, the results might show that there is a striking correlation between CLIL-based lesson and one's learning awareness.

4.2 Listening

The table below portrays the effect of deployment of T-test calculation among the four groups of students, including the data on mean, standard deviation, number and p-value in listening comprehension tests.

Table 6: The results of listening comprehension tests (pre- and post-test)

Group	Mean		Gain scores	Standard Deviation		N
	Pre - listening test	Post- listening test		Pre-test	Post-test	
CLIL group 1	23, 38	25, 07	1,60	2, 39	2, 01	13
CLIL group 2	24, 5	26, 08		2, 23	1, 67	12
EFL group 1	22	24, 28	2,42	2, 68	2, 55	14
EFL group 2	21, 66	24, 25		2, 46	1, 81	12
P value	0,0039	0, 0049				

The first CLIL group gained mean equaling 23,38 on pre-test listening comprehension. With this result, they found themselves on the second position. It is worth mentioning that the standard deviation seemed to be quite high, what further suggests that the group represented higher level of homogeneity than usual. For the result of the second CLIL group, their mean (24,5) was the highest of all the groups taking part in the study. Since the level of standard deviation in this group equaled 2,23, their heterogeneity is the highest among other groups on pre-test. Both EFL groups scored similar amount of points, the first gaining mean of 22 value, and the second gaining 21,66. The first EFL group appeared to be more heterogeneous than the second group with the standard deviation equaling 2, 68. For the matter of the mean achieved by the first CLIL group on post-test, it was 25,07. Having compared these both means, the gain scores of the group equal 1,69. When it comes to their homogeneity, its level was higher while having completed pre-test. The mean gained by the second CLIL group on pre-test, equaling 26,08 was the highest of all the four. However, correspondingly to the result from reading comprehension, the gain scores, equaling 1,60 appeared to be the lowest of all the groups. Due to the fact that standard deviation was represented by numbers 2,23 (pre-test) and 1,67 (post-test), it must be said that this group was more homogenous while completing the pre-test. As far as the results of the EFL groups are concerned, the means gained by the first group on pre-test equals 22, and, on post-test, 24,28. Having calculated the difference between these two means, their gain scores equaled 2,28. The higher level was observed while completing pre-test as the values of standard deviation equaled accordingly 2,68 (pre-test) and 2,55 (post-test). According to the results concerning the last group to be discussed, the means gained by them on pre-test were: 21,66 and 24,25 on post-tests. The difference between these two values gave me the gain score equaling 2,59 what constitutes the highest level of progress of all these four groups.

As it is clearly visible above, the results of the CLIL groups were better than the ones achieved by their EFL peers. More importantly, similarly to the conclusions drawn from reading tasks, the pace of their progress seemed to be slower than the one of EFL groups as their gain scores were lower.

The following data table depicts the mean and standard deviation of the pre- and post-Metacognitive Awareness Listening Questionnaire.

Table 7: The results of Metacognitive Awareness Listening Questioannaire

PRE- MALQ				
MALQ subscales	**Experimental Group**		**Control Group**	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Planning and evaluation	3, 97	0, 58	3, 71	0, 68
Directed Attention	4, 71	0, 43	4, 68	0, 39
Person Knowledge	3, 51	0, 39	3, 38	0, 28
Mental Translation	3, 74	1, 02	3, 71	0, 91
Problem Solving	5, 00	0, 35	4, 86	0, 43
POST- MALQ				
MALQ subscales	**Experimental Group**		**Control Group**	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Planning and Evaluation	4, 06	0, 56	3, 81	0, 55
Directed Attention	4, 78	0, 39	4, 78	0, 38
Person Knowledge	3, 63	0, 36	3, 51	0, 33
Mental Translation	3, 88	0, 96	3, 81	0, 88
Problem Solving	5, 03	0, 33	4, 96	0, 37

Having calculated the average scores and the level of homogeneity of the groups, the results are to be discussed with regard to the scale prepared by Vandergrift, Goh, Mareschal and Tafaghodtari, the authors. In details, similarly to the MARSI questionnaire, the higher is the mean, the more aware of the strategies the students are. The key goes as follows: 4.5 or higher stands for high awareness, 4.5-3.4 equals medium and 3.4 or lower indicates low awareness. Regarding *planning and evaluation* category, the use and awareness of the experimental group seemed to be higher than the ones of the control group. Namely, the mean gained by the control group equaled 3,71 and that made them being excelled by the experimental group by 0,2 points. As far as the scale is concerned, both of these groups appeared to represent medium level. Interestingly, there was also difference between these two groups when it comes to their homogeneity. Since standard deviation of the control group was 0,68 and 0,58 of the experimental one, the CLIL students were more homogeneous than their non-CLIL peers. As far as the *directed attention* category is concerned, again, the CLIL students outperformed EFL students. The means were the following: 4,71 and 4,68. Curiously enough, such results indicate *high* awareness of this strategy. Regarding homogeneity, the control group, representing standard deviation equaling 0,39, was more homogeneous than the experimental one. The means gained by these two groups in *person knowledge* category were the lowest. The experimental group gained 3,51, entering into medium scale, and the control group with the result of 3,38, entered into *low* scale. Considering their heterogeneity, its level was quite low in both groups. Nonetheless, the experimental group was less homogeneous, with standard deviation equaling 0,39, than the control group of which result was 0,28. The results concerning *mental translation* categories were similar to the ones achieved in the first subscale. Similarly, the experimental group gained higher mean (3,74) than the control group

(3,71) which gained exactly the same mean as in the first category. Having put the means into Vandergrift's scale, the results seemed to be *medium* among participants of both groups. As standard deviation was the biggest while completing the questions of this category, equaling respectively 1,02 and 0,91, CLIL groups remained more heterogeneous than their EFL peers. Regarding the last kind of subscale, *problem solving*, the results, again, was positively surprising. Although, the CLIL participants excelled over EFL students, the difference was small. The means being the following: 5,03 and 4, 96 shall be considered as of *high awareness.* Meaningfully, both of the groups achieved low level of standard deviation. Here, the control group was less homogeneous, scoring 0,37, than the experimental group of which standard deviation equaled 0,33. Having discussed the pre-test results, it is of high value to pay attention to post-test effects. At the very beginning it must be said that both groups did advance in their scores. Regarding the first category, the mean gained by the experimental group was 4,06 and thus, was higher than the one of control group by 0,25 points. Having considered scales suggested by the authors, the results suggest medium awareness. Interestingly, gaining standard deviation equaling 0,55, the control group was much more homogeneous than while completing pre-questionnaire. There wasn't any significant difference when it comes to the homogeneity in the first group. As far as *directed attention* subscale is concerned, surprisingly, both groups achieved the same means (4,78). Due to the above, their awareness appeared to be high. The level of their heterogeneity with the values of standard deviation equaling respectively 0,39 and 0,38 seemed to be comparable. In connection with the means of *person knowledge* category, again CLIL students, whose results was 3,63, marginally outperformed their EFL peers who gained 3,51 points. Relating to the above mentioned scale, the two results show medium usage. However, the control group with the standard deviation equaling 0,33 demonstrated higher level of homogeneity. In respect to *mental translation*, it must be said that the means were almost the same. With the results of 3,88 points, the experimental group was better by 0,07 points than the control group. Consistently, both means are to be considered as *medium.* Another time, the EFL students representing 0,88 standard deviation were less heterogeneous than their CLIL peers who gained 0,96. Relating to the last category, the first group from the table scored mean equaling 5,03 and, at the same time, constituting *high* level of awareness. The second one, through gaining mean totaling 4,96 also indicates the highest level of the given concern. With reference to the final indicator, the control group representing slightly higher level of standard deviation, remained less homogeneous than the experimental one. With an attempt to conclude the answer to the research question number four, CLIL students seemed to marginally excelled over non-CLIL participants. Nevertheless, it is crucial to be noticed that in one category the results were exactly the same. Additionally, once more, there was a slight regular progress noticed among participants in each group. Correspondingly to MARSI questionnaire, the justification might lie in the continuation of education throughout three months break between two stages of the study and previous knowledge of questions from the first stage.

The first research question was addressed to compare CLIL and their non-CLIL peers' reading skill being gained in general English. Additionally, taking into consideration reading strategies, the second question enquired into the use and awareness of metacognitive strategies among the same participants. The third question was devoted to investigate the possible outperforming outcomes in listening comprehension test among experimental and control groups. Adequately, in order to check the probable correlation between listening skills and awareness of metacognitive listening strategies, the fourth question was implied. The fifth research question aimed at exploring the differences in the pace of progress in receptive skills' increase among these two groups. Finally, having discussed the performance of the CLIL and EFL students, the question considering correlation between one's receptive skill performance, CLIL and their use of metacognitive strategies appeared. As expected, CLIL students significantly excelled over EFL participants on general reading comprehension tests at the two collection times. Overall, the findings might indicate that the participants having been taught through CLIL approach might have learnt a wider range of lexical items which do not appear in regular classes. The result may also show that reading skill is not sufficiently practised in the EFL classroom. As for the second question, both groups were analyzed in terms of their use of metacognitive reading strategies. Results showed that CLIL participants were more likely to use them. Regarding the difference between pre- and post- MARSI, it should be aforementioned that both groups increased their average scores. Aiming at considering three subscales, it was observed that the average scores increased in each of the scale. However, there was a particular trend noticed. Out of three categories: global, problem-solving and support strategies, CLIL students seemed to score the highest in the second category. This finding might indicate that CLIL, through being content-focused, is meant to elicit problem solving because specific information needs to be constructed. Regarding the third research question, it can be stated that, predictably, CLIL students outperformed their non-CLIL peers in general listening comprehension tests that were collected two times. In the main, both groups seemed to gain less scores on listening tests than on reading ones. Considering the difference between pre- listening test and post- listening test, it shall be noticed that both groups moderately increased their average scores. The results prove that CLIL students are more likely to be exposed to more lexical items than the EFL students. Accordingly, the next research question was supposed to investigate the students' awareness of metacognitive listening strategies. The results were adequate to the ones achieved on MARSI. Namely, CLIL students represented higher average than the EFL participants. It shall be stated that both groups improved their awareness in the span of three months in each of the five categories. More importantly, again, the *problem-solving* subscale gained the highest average among CLIL participants. It evokes the conclusion that being immersed in CLIL methodology might enhance this skill also by necessity of solving problems through foreign language. When it comes to the next to last research question of my study, the receptive skills of both groups seemed to progress at very similar, moderate pace. Nevertheless, it must

not be omitted that, although gaining less scores, the average improvement between pre- and post- receptive comprehension tests was higher among EFL students. It might be justified by the fact that CLIL participants scored higher on all tests so the progress might have been more regular than among the EFL groups. Finally, focusing on the relationship between the results of receptive comprehension skills, CLIL and metacognitive awareness listening and reading questionnaires, the following trends are spotted. For the CLIL participants, the higher averages in tests and questionnaires might be explained by the fact that CLIL not only does positively influence receptive dimension of language skills but also, in furtherance of improving reading and listening, allows students to be more strategic listeners and readers. Additionally, the following correlation might exist: the more advance are students' receptive skills, the more likely the strategies are to be applied.

5. Recommendations

Due to the fact that there were three main limitations while conducting my research, I find the following recommendations useful for further investigation in the area of the given subject. As far as the first limitation was sample size, it would be worth researching more participants than I did. Obtaining the results of much more students would provide more complete picture of CLIL's effect. The second issue is strictly connected with the time the research was being investigated on. I believe that longitudinal studies lasting at least one year would be of prime importance to the relevance of the study. The last but not least, the measurement of one's level of English language shall be done at the very beginning of the researching process. I notice that there was a slight discrepancy between the level that the teacher did assume and the students' actual level. Thus, the limitations and the findings of the study could possibly provide the background for further research of CLIL and other bilingally-related methods of teaching on both receptive skills and the way the student think. Based on the result from the given study, future research shall conduct a similar study but with the addition of the qualitative data and some observations along with interviews with the students. The focus also could be on the way the teacher conduct the CLIL lesson and whether there is former instruction on how to listen and read. Obviously, these recommendations constitute only a starting point to explore more widely the phenomenon of bilingual education and its further benefits in the process of learning.

6. Conclusion

Principally, the results report that CLIL groups generally excelled over EFL ones on every tests. This finding might indicate that CLIL approach favorably affects the development of receptive skills comprehension. Notwithstanding, since the scores were higher than on listening tests, such a result is more recognized in reading comprehension. It might be explained through the fact that Content and Language

Integrated Learning makes students exposed to "additional reasons for reading" (Dulton-Puffer 2008: 6). Additionally, the findings show that receptive skills progress similarly among these two groups. As regards the use of metacognitive strategies, it could be worth taking into consideration that this factor is also determined by CLIL and level of receptive kills comprehension. Thus, another crucial aspect of research would be to investigate the relationship between language level, context and strategies. Finally, the gain scores achieved in both reading and listening comprehension tasks indicate that overall bigger progress apply to the non-CLIL participants. Nonetheless, it shall not be forgotten to add that the scores of CLIL group are regularly higher than the ones of their peers in the control group. The research evidence shows that the teenagers exposed to additive language context, attain relatively higher level of L2 receptive skills and do benefit metacognitively from their CLIL learning experience. On the other hand, students in EFL environment represent lower level of language establishment and might experience some cognitive disruptions what, in further, does affect their use of metacognitive strategies. In closing, the present study constitutes just a limited attempt to investigate this research field. On the whole, various limitations appeared during the research. The main limitation of the present study is the limited number of participants. The short duration of the study, equaling three months between each stage, might have also affected the results negatively. Taking these into consideration, there are some improvements that could have been made. For instance, the sample size could be larger and the period between two stages shall be longer to make the research more significant. At last, the present work may suggest and promote subsequent examining of one of the bilingual programmes and its effect on one's both cognitive and language through more rational, longitudinal studies.

About the Author

Marzena Foltyn is a Polish passionate of language learning and a propagator of education's significance. The author has been teaching English and Czech language in Polish schools. Her teaching and research interests include also cross-cultural perspectives, applied linguistics, creativity and literature.

References

Coonan, C. M. Insider Views of The CLIL Class Through Teacher: Self-observation-introspection. Internal Journal of Bilingual Education and Bilingualism, P.10.

Costa, F. CLIL (Content and Language Integrated Learning) through English in Italian Higher Education. Milano, LED, 2016.

Dalton-Puffer, Ch. (2007) Discourse in Content and Language Integrated Learning (CLIL) Classrooms. John Benjamins Publishing Company, 2007. P. 1; 283.

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

Krashen, S. Second Language Acquisition: Theory, Applications, and Some Conjectures, Mexico: Cambridge University Press, 2013.

Llineares, A., Morton, T., Applied Linguistics Perspectives on CLIL. Amsterdam/Philadelphia, John Benjamins Publishing Company, 2017.

Mehisto, P., Marsh, D. and Frigols, M. Uncovering CLIL: Content and Language Integrated Learning in Bilingual and Multilingual Education, Oxford: Macmillan, 2008.

Mokhtari, K., Reichard, C. A. Assessing Students' Metacognitive Awareness of Reading. Oklahoma State University, Journal of Educational Psychology, 2002.

Papaja, K. Focus on CLIL: A Qualitative Evaluation of Content and Language Integrated Learning (CLIL) in Polish Secondary Education. Cambridge Scholars Publishing, 2014. P. 1.

Pladevall-Ballester, E. and Vallbona, A. (2016). "CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills", in System, 58

Sajda, L. Content and Language Integrated Learning - CLIL. Auflage: GRIN Verlag, 2008.

Szcześniak, K., Łyda, A. Second Language and Teaching, Awareness in Action, The role of Consciousness in Language Acquisition. Switzerland, Springer International Publishing, 2014.

Wilhelmer, N. Content and Language Integrated Learning (CLIL). Teaching Mathematics in English. Saarbrücken, VDM Verlag, 2008.

Widdowson, H. Defining Issues in English Language Teaching. Oxford University Press, 2003.