Thinking and Content Learning of Mathematics and Science as Cognitional Development in Content and Language Integrated Learning (CLIL): Teaching Through a Foreign Language in Finland

Aini-Kristiina Jäppinen
Institution for Educational Research, University of Jyväskylä, Finland

This paper presents a study on thinking and learning processes of mathematics and science in teaching through a foreign language, in Finland. The entity of thinking and content learning processes is, in this study, considered as cognitional development. Teaching through a foreign language is here called Content and Language Integrated Learning or CLIL. CLIL refers to all those diverse programmes, including some forms of immersion and bilingual education, where a foreign language is a medium of instruction, affecting the entire learning process of the learner. Thinking and content learning in CLIL manifests itself as analogical CLIL reasoning systems, based on two languages, and is assumed to affect cognitional development. Cognitional development was studied with 669 Finnish mainstream L1 learners aged 7–15 in the public comprehensive school. The experimental group, 335 learners, was taught through English, French or Swedish. The experimental group was compared with a control group of 334 learners, taught through the mother tongue, i.e. Finnish. Cognitional development was studied in terms of individual concepts and conceptual structures that are here called meaning schemes. The results of four measurements in 2002–2003 are presented in which statistical differences were found between the experimental and the control group in cognitional development.

Keywords: CLIL, content and language integrated learning, teaching through a foreign language, thinking, content learning, cognitional development

Globalisation and continuous social change in the field of economics, culture and politics (e.g. Robertson, 2001; Urry, 2003) make particular demands on European education (European Commission, 2000, 2001; Reding, 2001). This change includes the question of increasing the efficiency and scope of language learning. Teaching a content through a foreign language can be regarded as an approach to rise to this challenge.

There is a growing tendency to teach through a foreign language in Europe (Marsh & Hartiala, 2001). Teaching through a foreign language can be seen as an important medium in advancing intercultural knowledge and understanding (Hughes, 1991; Trim, 1995; Viqueira, 1991) and communication (Maclntyre et al., 2002; Paulston, 1992), and in providing a wider choice of languages to be learned. It can also be considered to benefit the mobility of labour and increase social justice all over Europe, particularly in public and mainstream education. It also
allows a temporal and quantitative increase in foreign language usage, normally without any curricular changes or extra investments.

European approaches of teaching a content through a foreign language originate from Canadian immersion (e.g. Cummins, 1995, 2001; Swain & Johnson, 1997) but are strongly adapted to European cultural, linguistic or social circumstances. The European ways of using a foreign language as a tool in teaching contents vary to a great degree, as a result of the multicultural and multilingual diversity in Europe (Hartiala, 2000; Marsh & Hartiala, 2001).

**Modes of Content and Language Integrated Learning or CLIL**

The forms of teaching through a foreign language are often called Content and Language Integrated Learning or CLIL. The term is widely adopted, for example, by the European Commission (de Bot, 2001; European Commission, 2000, 2001; Reding, 2001). In this paper, CLIL refers to all those situations where there is a connection between a foreign language used as a medium of instruction and a content taught, including immersion and some forms of bilingual education (e.g. Baetens Beardsmore, 1982; Baker & Prys Jones, 1998; Met, 1991).

In CLIL, learning the foreign language is not the direct objective of education but a natural part of the whole learning process. Because of their diversity, European CLIL programmes have various aims related to culture, environment, language, content, and/or learning. This means that, in many cases, language learning or teaching is not the focus point of the CLIL programmes although language is always one of the key features of a CLIL learning environment. European CLIL can be carried out from early childhood to higher education, by native or non-native teachers, in any language, subject or quantity, and during any period (Hartiala, 2000; Marsh & Hartiala, 2001).

Contemporary European CLIL programmes are a fairly new and prolific phenomenon. Consequently, there is a considerable scarcity of CLIL research although studies on immersion and bilingual education (Baker, 1996; Bialystok, 1999; Bialystok & Majumder, 1998; Cummins, 2001; Mohan & Beckett, 2003) provide useful knowledge for CLIL approaches.

**CLIL in Finland**

Finland is one of the pioneers in the European CLIL, especially in the public mainstream education where, in 1996, 8% of the Finnish primary and 15% of the secondary mainstream schools were assumed to use a foreign language as a medium of instruction (Nikula & Marsh, 1996). Since then, the number of CLIL programmes has increased.

In Finland, CLIL is not normally restricted to a cultural or linguistic minority, except the immersion in Swedish, but aimed at the public mainstream education of L1 learners to provide everyone with better opportunities for the working life and an access to the multicultural and plurilingual world (cf. Leung, 2003). Most Finnish CLIL learners also get formal teaching of the target language. At the same time, the instruction of the mother tongue is well organised (Adams & Wu, 2002; Välijärvi et al., 2002).

Because of the rapid increase of CLIL, some Finnish educational institutes and experts became worried about the learning process of the content taught and the
development of mother tongue skills, although most Finnish CLIL learners get instruction in Finnish and of Finnish, along with the foreign language usage. Therefore, it seemed essential to study the thinking and content learning processes of the Finnish CLIL learners.

English is the most common CLIL language because it is used as a lingua franca in many Finnish enterprises and international contexts. The second most common is Swedish but there is also a lot of CLIL in German, French and Russian. The most common subjects taught are mathematics, biology, geography, music, drawing, physical education, handicrafts, history, domestic science, physics, chemistry, and art.

In Finland, immersion takes place in Swedish, normally as early total immersion (e.g., Swain & Johnson, 1997). This means that learners are first taught solely through Swedish, the portion of Finnish increasing with time. The teachers are native speakers of Swedish who are not allowed to use Finnish in instruction, although most of them are fluent in Finnish (Björklund, 1997). Other Finnish CLIL programmes of English, French, German or Russian can embody different quantities of both languages, depending on the subject or topic taught. The teacher can normally use Finnish with the same pupils in settings other than CLIL. A large number of the teachers are not native speakers of the foreign language although they have a good command of it (Hartiala, 2000; Nikula, 1997).

**Research Objectives and Research Questions**

The aim of the study was to examine Finnish mainstream L1 learners’ thinking and content learning processes in CLIL environments of English, French or Swedish. Thinking and content learning processes were, in this study, considered as a whole in order to be able to study how the CLIL learners made use of the contents taught through a foreign language. This entity is here referred to as

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Contents to be taught through the foreign language</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOPICS</strong></td>
<td>Mathematics</td>
</tr>
<tr>
<td>magnitude, ratio, proportion, denomination, estimation</td>
<td>climate</td>
</tr>
<tr>
<td><strong>CONCEPTS</strong></td>
<td>(1) number meanings: rational, whole, and natural numbers</td>
</tr>
<tr>
<td>(2) power; percentage</td>
<td></td>
</tr>
<tr>
<td>(3) recognition of diagrams and tables</td>
<td></td>
</tr>
<tr>
<td><strong>MEANING SCHEMES</strong></td>
<td>(1) number systems: different systems based on 10</td>
</tr>
<tr>
<td>(2) addition, subtraction, multiplication, and division with rational, whole, and natural numbers</td>
<td>(2) weather signs/maps/forecast/types</td>
</tr>
<tr>
<td>(3) construction and interpretation of diagrams and tables</td>
<td>(3) wind</td>
</tr>
<tr>
<td></td>
<td>(4) atmosphere, air protection, and global climatic zones</td>
</tr>
</tbody>
</table>
cognitional development. ‘Cognitional’ is used here to refer to both thinking and content learning and to separate it from the established term ‘cognitive’ that covers, according to Encyclopedia Britannica, ‘every mental process that can be described as an experience of knowing as distinguished from an experience of feeling or of willing’ (see also the use of the term ‘cognitive’ in The International Encyclopedia of Education, 1994).

Cognitional development is assumed to manifest itself in understanding, using and applying concepts and conceptual structures of the contents taught through a foreign language in mathematics and science. Different conceptual structures when concepts are related to each other are here called meaning schemes. Table 1 presents the contents taught that were covered in four measurements as concepts and meaning schemes.

Three research questions were formed: (1) What kind of cognitional development is to be found in CLIL in mathematics with regard to magnitude, ratio, proportion, denomination and estimation, and in science as to the climate? (2) What are the advantages and problematic areas of CLIL? (3) How to support CLIL learners’ thinking and learning processes?

Characteristics of CLIL Learning Environments

CLIL environments are assumed to consist of four key characteristics that differ from the settings where learning occurs through the mother tongue (Jäppinen, 2002). This makes it possible to study differences between CLIL and education realised solely through the mother tongue.


The first of the four key characteristics is the great difference between what the learner is able to do without guidance and what s/he can do with help from others. This kind of ‘large’ ZPD is typical of CLIL. CLIL learners need much support to reach the upper limit of their ZPD in terms of extra explanations and help from the teacher and fellow learners, in terms of special gesticulation and movement, in terms of the special features of spoken language, and in terms of supportive materials. The second element is a variety of socio-culture-psychological factors, intertwined with each other. The foreign language used has an effect on the entire learning situation. It opens a wider view to other kinds of societies and cultures that the learners interpret in a very personal way. The third characteristic is special discovery learning related situations. These manifest themselves in various and contradictory connections between the foreign language used and the learner’s mother tongue which the learner discovers and makes use of in meaning making. These situations are explained in more detail in the next section. Finally, the CLIL environment provide informal and natural language-learning opportunities because CLIL learners learn and acquire the
foreign language in much the same way as they have learned their mother tongue.

Because the CLIL environments differ significantly from those in teaching through the mother tongue, CLIL learners’ thinking and content learning processes can be compared with those in teaching through the mother tongue. In this way, it is possible to uncover the consequences and effects of the CLIL learning process.

**CLIL Thinking and Content Learning Studied as Cognitional Development**

Although the learning process of the foreign language or the mother tongue has been focused on in some studies, thinking in connection with content learning is an almost unexplored area in Europe (cf. Walker & Tedick, 2000). Consequently, many questions remain unanswered (cf. Mohan et al., 2001): Do these types of language learning environments support individual thinking and learning processes? What are the problematic features of CLIL? How could administrators, politicians, and teachers develop CLIL environments so that learners would be supported in the best possible way?

This paper focuses on the thinking and content learning processes of Finnish CLIL learners in comparison with learners taught through the mother tongue (Finnish), not on language learning or linguistic issues. How can we study, then, the effect of foreign language usage on CLIL learners’ thinking and content learning processes, that is, on their cognitional development?

In its current forms, CLIL is a new creation. There is not yet any satisfactory theoretical framework to explain how CLIL learners make sense of the contents taught in diverse CLIL environments, that is, how they understand, use, and apply concepts and meaning schemes within two semantic systems of two different languages. In other words, what is CLIL learning? This article will present some answers to this question. The research methods of the study did not follow those recently used for studying human cognition (cf. Eysenck & Keane, 1999). The scarcity of CLIL research and the special nature of CLIL environments, due to simultaneous and multiple effects of two languages and two semantic systems, required new and innovative approaches.

CLIL learning is a mixture of languages, cultures, and contents. This study is based on the idea of Vygotsky (1986) and Bruner (1971, 1983, 1986, 1990, 1996) that the relationship between language and thought is basically related to environments and cultures. Language is not only something that the learner thinks about but also part of the learner’s thinking process itself. CLIL learning also includes Halliday’s (1978) idea that the learner thinks with language, that is, reflects on experience and achieves understanding. The foreign language is not just a tool of instruction but an important tool in the learner’s thinking process.

The study is also based on the widely accepted idea (e.g. Cole & Cole, 1996) that language and thought are ultimately separable psychological functions that cannot be reduced to each other although they are profoundly intermixed in all normal development. The extent and the way they influence each other is, however, a complicated issue. It becomes ever more complicated when a foreign language is a tool of learning and has a remarkable effect on thinking. Instead of
two factors we now have three: two language systems, thinking, and content learning. However, the two languages and content learning form a unity because the foreign language and the mother tongue are both involved in the learning process of the content. This makes it possible to study content learning through thinking processes and thinking through content learning.

CLIL learners are assumed to interpret CLIL environments by creating analogical reasoning systems based on the two languages that influence, often unconsciously, the learning situation and are part and parcel of it. As Figure 1 shows, analogical CLIL reasoning is the basis for CLIL thinking and content learning. The reasoning process is analogical because it is based on comparing similarities and differences between the two languages in a semantic, cultural, and social sense. Analogical CLIL reasoning forms the basis for content learning as representational thinking, that is, meaning making of concepts and meaning schemes, and brings about the cognitional modifiability of concepts and conceptual systems.

To study the consequences of analogical CLIL reasoning, learning and thinking had to be linked with each other. For this purpose, 18 unities were created and modified (Jäppinen, 2002). The unities of analogical CLIL reasoning are here simply called ‘cognitions’ to be able to operate with a single term and to connect it with cognitional development. These cognitions were pairs of five critical discovery learning areas, adapted from Bruner (1971) and 10 thinking categories, adapted from Piagetian operations (e.g. Cole & Cole, 1996; Keating, 1980 in Cole & Cole, 1996). The study focused on the development of these cognitions as cognitional development.

The critical discovery learning areas to be studied were: (1) awareness of existing concepts and the ability to understand, use, apply, and explain them, (2) awareness of meaning schemes and the ability to create links between them, (3) the ability to exploit information in problem-solving situations and to formulate hypotheses, (4) the ability to solve problems that are difficult to explain to oneself.
or when concepts that one has not yet acquired are involved, and (5) the ability to exploit the flow of information, to make comparisons and form antitheses, and to choose between two or more alternatives.

The 10 thinking categories through which the critical discovery learning areas were studied included: (1) classifying according to more than one concurrent factor, (2) realising the constancy of properties, (3) realising the similarity of a change, (4) realising the compensation or equivalence of a change, (5) realising the reciprocity or reversibility of a change, (6) noticing and charting alternatives for action, (7) thinking ahead the progress of a process in the light of different alternatives and choices, (8) changing possibilities into hypotheses and testing the hypotheses, (9) becoming conscious of one’s own thought processes, and (10) thinking beyond conventional limits.

The cognitions remained the same in each of the four measures to make it possible to study cognitional development. Only the content of the measurements varied, in accordance with the topics in the curriculum, Table 1. The cognitions were chosen on the basis of pre-testing that took place in autumn 2001. Its purpose was to compare different combinations of thinking categories and critical discovery learning areas and to select and modify the cognitions suitable for each age group.

Method

The study was carried out in 12 Finnish mainstream comprehensive schools that provided both experimental and control groups. The age range of the 669 learners was 7–15 years. The schools were located in Helsinki (the capital, over 500,000 inhabitants), Tampere (about 200,000 inhabitants), and Turku (about 190,000 inhabitants) that together represent about 20% of the population of Finland. The experimental group, 335 learners, were taught through English, French or Swedish. They were compared with a control group of 334 learners taught through Finnish that was the mother tongue of all the learners involved in the study. The amount of CLIL in English was about 60%, the amount of French 30%, and the amount of Swedish about 10%.

This article presents the results of four measurements. The first measurement M1 provided the starting level of the research. It was carried out in spring 2002. M2, M3, and M4 concerned cognitional development (Jäppinen, 2003). The first follow-up measurement M2 was made in autumn 2002, the second M3 in spring 2003, and the third follow-up measurement M4 in autumn 2003. This article presents only those cognitions in which there were significant differences between the experimental and the control group in all three or in the two last follow-up measurements after a similar starting level. The cognitions were the same in the tests of mathematics and science, and the majority of the learners were tested in both subjects.

Subjects

Mathematics and science are studied at all levels of the Finnish comprehensive school. This allowed a continuous follow-up of cognitional development. Both subjects provide enough CLIL experiences for research purposes, and are common in the Finnish and in the European CLIL (e.g. Met, 1991). For studying
cognitional development in mathematics, research on learning mathematics was applied. For studying science, research on knowledge representations and concept development was applied to and exploited for CLIL environments (Behr et al., 1992; Bisanz et al., 1994; Cole & Cole, 1996; Elkind, 1976; Ellis & Siegler, 1994; Eysenck & Keane, 1999; Fuson, 1992; Gerrig & Banaji, 1994; Greer, 1992; Hunt, 1994; Lubart, 1994; McNamara, 1994; Nunes & Bryant, 1996; Nunes et al., 1993; Ross & Spalding, 1994; Sowder, 1992; Veel, 2000; Wells, 1999; Wood, 1998)

Learners and schools

There is no database about Finnish CLIL programmes. For this reason, it was not possible to use random sampling in choosing learners for the study. The three cities, Helsinki, Tampere and Turku, were chosen because they provide a large research field and enough CLIL schools for the study. The selection of the schools was organised by collecting all available information from the administration. The study covered all nine grades of the Finnish comprehensive school. The learners were divided into three age groups: the first age group with learners aged 7–9, the second with learners aged 10–12, and the third with learners aged 13–15. The first age group had five, the second seven, and the third group six cognitions to be studied. Table 2 shows the number of the learners in each group.

<table>
<thead>
<tr>
<th>Learners</th>
<th>N = Experimental</th>
<th>N = Control</th>
<th>N = Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>First age group</td>
<td>131</td>
<td>138</td>
<td>269</td>
</tr>
<tr>
<td>Second age group</td>
<td>98</td>
<td>89</td>
<td>187</td>
</tr>
<tr>
<td>Third age group</td>
<td>106</td>
<td>107</td>
<td>213</td>
</tr>
<tr>
<td>N = Total</td>
<td>335</td>
<td>334</td>
<td>669</td>
</tr>
</tbody>
</table>

Teaching and Testing

The contents taught originated from both the Finnish national curriculum and the particular curricula of the schools. The topics and concepts of the study provided a logical continuum from one grade to another, and allowed the observation of cognitional development. The 46 teachers had freedom to choose materials and methods. Only the language of instruction, the topics, and the schedule were predetermined. There was no follow-up on teaching. The possible bias due to the idiosyncrasies of the teachers was reduced by the number of the teachers involved. The amount of teaching through a foreign language of the total teaching time varied in the first and second age group from 50% to nearly 100% and in the third age group from 20% to 30%.

There were four predetermined teaching periods, each of which took 3–5 weeks. The four measurements in the form of four tests in mathematics and four in science, followed the teaching periods. The language of the tests was Finnish because the control group was taught in Finnish. The tests included problem solving, generalisations, picture tasks, metaphors, similes, proverbs, tables, graphics and planning. The viewpoints for examining the concepts and meaning schemes were as follows: nesting classes; classifying according to multiple criteria; kinship relations; conceptual similarity, reversibility, seriation,
combination, linkage, and/or invariance; level of abstraction; and/or part-whole constructs.

**Analyses of the tests**

The analyses were made on the basis of a taxonomy of five levels. Its criteria were created, originally, on the basis of the SOLO Taxonomy (Biggs & Collis, 1982), however, without an age-group classification. The modified contents of the SOLO Taxonomy were adapted to define the cognitional level of the learner. The taxonomic levels were: 0 – no apparent cognition, 1 – weak cognition, 2 – advanced cognition, 3 – almost complete cognition, and 4 – complete cognition. The exact criteria for each level were theoretically predetermined in terms of the extent of logicality of thinking processes, the level of generalisation, the quality of inductive or deductive reasoning, and the quality of problem-solving strategies. All the learners’ solutions of a test item were listed according to these theoretical criteria and adapted in a concrete form to the five taxonomical levels so that all the answers of the same type were on the same level. In M1, this process was triangulated so that at least two different persons placed the answers independently to the taxonomical levels. The result of the triangulation was 85–95%.

The first analysis focused on the differences between the experimental and the control group in overall performance. The differences were detected as equality of means as sum variable: sum scores for the learners were calculated from each test. These were used for calculating the means and standard deviations. A sum score referred to the total score of a learner in a test; the first age group max. 20, the second 28, and the third 24. The differences between the groups, based on sum scores, were analysed by means of a \( T \)-test when \( p < 0.01 \).

The second analysis concerned the differences between individual cognitions. They were analysed by means of Chi-square. If there were, after an equal starting level, significant differences at the level \( p < 0.01 \) between the experimental and the control group in all the follow-up measurements, the differences were interpreted as resulting from different cognitional development. The tests are not included in this paper, except for some contents of those test items in which there were statistically significant differences between the experimental and the control group, Appendix 1, 2.

All the learners did not participate in all the tests. Some learners were absent at the time of testing. Some classes could not participate in some of the tests because of curricular difficulties or the school could not organise a course on the topic taught through the foreign language during the teaching and testing period. Further, some learners had moved away or entered the classes after the study begun. Therefore, to study the differences, only those learners were included in the analyses of \( T \)-tests who had participated at least in two measurements of which the other was either M3 or M4.

**Results**

**First age group: Mathematics**

The topic of M1 was the number system 0–20 and particularly those additions where there was a surpassing of 10. M2 concerned additions and subtractions of the number system 0–100 and the interpretation of a table. The content of M3 was
simple multiplications, operations to estimate time, and the interpretation of a bar chart. M4 concerned additions and subtractions of the number system 0–10,000, the interpretation of a linear diagram, and the concept of division.

No significant differences between the experimental and the control group in cognitional development were found in any of the measurements, Table 3, or in any single cognition. Thus, CLIL-environments seemed to have supported the mathematical thinking and learning processes of the learners in the first age group in the same way as the mother tongue environment.

First age group: Science

M1 had to do with the states of water, that is, vapour, water, and ice. The topic of M2 was the sun and the planets and their influence on the succession of the seasons. M3 focused on weather observation and forecast, and M4 on the circulation of water.

There were statistically significant differences between the experimental and the control group in M2 (t = −2.659, df = 190, p < 0.01) where the cognitional level of the control group was higher than that of the experimental group, and in M4 (t = 3.973, df = 157, p < 0.01) where the control group was behind the experimental group, Table 4. These results support the results in mathematics that analogical CLIL reasoning systems work for the younger CLIL learners. CLIL supports the thinking and content learning processes resulting in the same kind of cognitional development as in teaching through the mother tongue.

The results from the single cognitions clarify the results above. It was found that some very abstract topics may not be very well suited for young CLIL learners, Appendix 1. In the cognition that concerned conceptual invariance with an unknown factor, there was a statistically significant difference between the experimental and the control group in M2 (χ² = 26.482, df = 4, p < 0.01), the experimental group being behind the control group in cognitional development. In the cognition that focused on problem-solving based on alternatives, the cognitional
level of the experimental group was lower in M2 ($\chi^2 = 15.659$, df = 4, $p < 0.01$) and in M3 ($\chi^2 = 18.843$, df = 4, $p < 0.01$).

However, when the children got older, the CLIL learners were in M4 on a higher cognitional level than the control group in the cognition that had to do with equivalence or similarity of a conceptual change ($\chi^2 = 21.220$, df = 4, $p < 0.01$) and in the cognition concerning multifactorial classification of a schema ($\chi^2 = 12.900$, df = 4, $p < 0.01$), Appendix 1.

**Table 4** First age group: Science

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>experimental</strong></td>
<td>10.5</td>
<td>9.75</td>
<td>8.8</td>
<td>12.27</td>
</tr>
<tr>
<td><strong>control</strong></td>
<td>9.97</td>
<td>10.9</td>
<td>9.35</td>
<td>10.68</td>
</tr>
</tbody>
</table>

**Table 5** Second age group: Mathematics

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>experimental</strong></td>
<td>13.18</td>
<td>10.01</td>
<td>15.92</td>
<td>13.55</td>
</tr>
<tr>
<td><strong>control</strong></td>
<td>11.56</td>
<td>7</td>
<td>10.24</td>
<td>9.27</td>
</tr>
</tbody>
</table>
Second age group: Mathematics

The content for M1 was the number system of 0–1,000,000, the number units, and the mental calculations of addition and subtraction. M2 had to do with multiplication and division with the numbers 0–1,000,000 and with understanding a line diagram. M3 focused on the continuum, on its applications with decimals, and on additions and subtractions with decimals. M4 was concerned with addition and subtraction of rational numbers and fractions and the construction and interpretation of a bar chart.

The cognitional starting level of M1 was equal in both groups. However, the experimental group was ahead in M2 ($t = 3.787, df = 127, p < 0.01$), in M3 ($t = 5.664, df = 129, p < 0.01$), and in M4 ($t = 4.192, df = 108, p < 0.01$), Table 5. Teaching through a foreign language seemed to support or even promote the mathematical thinking and learning processes of the learners in this age group.

As to single cognitions, two of them proved special after a equal starting level of M1, Appendix 2. The first of these was charting conceptual alternatives and making choices between congruent concepts. In M2 ($G^2 = 17.311, df = 4, p < 0.01$), in M3 ($G^2 = 16.039, df = 4, p < 0.01$), and in M4 ($G^2 = 13.151, df = 4, p < 0.01$) the cognitional level of the experimental group was higher than that of the control group. The second was charting possible alternatives within existing meaning schemes in spite of ambiguous concepts. Again, the experimental group was ahead in M2 ($G^2 = 27.889, df = 4, p < 0.01$), in M3 ($\chi^2 = 16.122, df = 4, p < 0.01$), and in M4 ($G^2 = 21.722, df = 4, p < 0.01$). This kind of an ambiguous cognition seemed to be demanding for the younger learners but not anymore in this age group.

Second age group: Science

M1 focused on interpreting and completing a weather map, on weather signs, on the composition of the atmosphere, and on the effects of the Sun’s radiation on the Earth. M2 dealt with the significance of the atmosphere for the Earth, and M3

Table 6 Second age group: Science

<table>
<thead>
<tr>
<th>measurements</th>
<th>experimental</th>
<th>control</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>12.35</td>
<td>11.68</td>
</tr>
<tr>
<td>M2</td>
<td>14.95</td>
<td>14.09</td>
</tr>
<tr>
<td>M3</td>
<td>14.75</td>
<td>12.89</td>
</tr>
<tr>
<td>M4</td>
<td>14.97</td>
<td>11.94</td>
</tr>
</tbody>
</table>
with climatic zones. M4 was concerned with the origin and mechanism of the wind and trade winds and the monsoons.

After an equal starting level, M1, and the first follow-up measurement M2, there was a statistically significant difference between the groups in M3 ($t = 2.549, df = 131, p < 0.01$) and in M4 ($t = 4.980, df = 101, p < 0.01$). The experimental group has reached a higher cognitional level. The findings suggest that teaching through a foreign language in science gives support to or even promotes the cognitional development of the CLIL learners in this age group.

There was no single cognition where there was a statistically significant difference between the groups in the three follow-up measurements after an equal starting level. Thus, the different cognitional development was not dependent on a single cognition.

**Third age group: Mathematics**

M1 focused on the inverse of fractions and decimals, on their relationship to percents, and on interpreting bar charts. In M2, the learners dealt with the operations of decimals, with a decimal as a divider, and with diagrams as lines. The areas in M3 covered the concept of equation, the formation of an equation from a verbal problem, and the interpretation and construction of a table. In M4, the learners worked with real numbers, the power, and the drawing up and interpreting tables.

The experimental group and the control group were very equal and there was no statistically significant differences between the groups in any of the measurements, Table 7, or in any single cognition in all the follow-up measurements after an equal starting level.

**Table 7 Third age group: Mathematics**

<table>
<thead>
<tr>
<th>means of the sum variable</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.44</td>
<td>8.6</td>
<td>10.93</td>
<td>11.68</td>
</tr>
<tr>
<td>(experimental)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(control)</td>
<td>10</td>
<td>9.1</td>
<td>10.79</td>
<td>9.27</td>
</tr>
</tbody>
</table>
Third age group: Science

Table 8 Third age group: Science

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>8.86</td>
<td>10.94</td>
<td>11.71</td>
<td>14.46</td>
</tr>
</tbody>
</table>
| means of the sum variable
| experimental | 8.86 | 10.94 | 11.71 | 14.46 |
| control      | 10.09| 9.75 | 10.03 | 12.76 |

The results from the science tests concerned, partly, different learners than those in mathematics. The topic for M1 was the climate of the tropical regions, the trade winds, the protection of rainforests, and their characteristics. M2 concerned the influence of human actions and the use of natural resources on the atmosphere and on sustainable development. M3 focused on the continental and maritime climate and on global weather observation. In M4, the topic was the global importance of climatic conditions for human beings.

There were no statistically significant differences between the experimental and the control group which indicates that CLIL-environments had supported the learners’ cognitional development in the third age group in the same way as the learning environment of the mother tongue, Table 8. As regards the individual cognitions, there was no systematically different cognitional development between the experimental and the control group. It was very difficult to organise the CLIL courses at the same time in different schools. Therefore, in M4, only a small part of the learners participated which might have had an influence on the results.

Discussion

The aim of the study was not to examine how well a content was taught through a foreign language. Instead, the purpose was to see how the learners exploited, in their thinking processes, the concepts and meaning schemes they had learned through a foreign language. In other words, the aim was to find out about the consequences of analogical CLIL reasoning systems as cognitional development.

According to this study, the Finnish CLIL environments in public mainstream L1 education have succeeded, in general, in offering favourable conditions for
thinking and content learning in mathematics and science. In most cases, the
cognitional development in the CLIL environments resembled the development
in teaching through the mother tongue. In some cases, in the second age group, the
cognitional development of the experimental group seemed to be even faster than
that in the control group. One thus can conclude that analogical CLIL reasoning
systems work in the Finnish CLIL environments where the mother tongue of all
learners is Finnish and they are taught through English, French or Swedish.

In particular in the second age group, CLIL learners aged 10–14, cognitional
development in the experimental group was occasionally even faster than in the
control group. However, younger CLIL learners, aged 7–9, had some difficulties
with certain more abstract scientific topics that may not be very well suited for
being taught through a foreign language. In the third age group, the experimental
and the control group differed very little from each other. One possible explana-
tion for this is that the amount of teaching through a foreign language was
smaller in the third age group than in the first or the second. This was due, first, to
the level of difficulty of the subjects which restricts the amount of CLIL or,
second, to the fact that eligibility for further studies that are, in general, in Finnish
must be guaranteed.

The positive outcomes from Finnish CLIL environments mean that teaching
through a foreign language supports CLIL learners’ thinking and content learning.
There are factors that are assumed to have positively influenced the cognitional
development of the Finnish CLIL learners. These learners have a large number
of mother tongue lessons during their studies in the comprehensive school, in addition
to participating in teaching contents through a foreign language. Most CLIL
learners also attend subject lessons that are conducted in Finnish only. Finnish
CLIL learners in the second and third age groups are also normally involved in
formal language learning, in addition to CLIL lessons, which guarantees the learn-
ing of grammar and other necessary features of the target language.

Teaching and learning of the mother tongue and mathematics are, in Finland,
of high quality, as the outcomes of the PISA Study 2000 show (OECD, 2002). This
follows mainly from two factors that are typical of Finnish learning environ-
ments. First, learning the mother tongue is highly valued and instruction is
well-organised (e.g. Adams & Wu, 2002; Kirsch et al., 2002; Välijärvi et al., 2002).
Secondly, Finnish teacher education takes place at the universities and is of high
quality. The social status of teachers is also high.

One must also consider the possibility of there having been other variables, for
example, home environment, the socio-economic status of parents, possible
entrance exams for CLIL classes, or the nature of learners entering CLIL that
might all have influenced the results. These factors were, however, taken into
consideration. It was assumed that the influence of these factors would be appar-
ent in the cognitional starting level, and the follow-up measurements were
expected to uncover similar or different cognitional development. Some other
factors were also expected to decrease the bias, one of them being that the experi-
mental and control groups came from similar schools in similar kinds of urban
social settings.

One possible factor to increasing the bias might have been the lack of teacher
and material observation. For example, teaching methods that emphasise critical
and creative thinking support learning styles and processes that help the learners
to better manage in the tests requiring more multiform and sophisticated thinking. However, the large number of teachers using diverse teaching methods and a similar professional background of the CLIL and mother tongue teachers was assumed to reduce the bias. Another critical point resulting in bias might have been the turnover of some learners during the study. Although most of the teachers were inspired by the study in all four measurements, a few were lacking necessary commitment and dropped out. To avoid this type of bias, the means as sum variables included only those learners that had participated in at least two measurements of which one should be either M3 or M4.

Although the CLIL learners’ general cognitional level and development was high, compared with the control group, some difficulties were found in the first age group with very abstract or less central topics. In this study, these types of topics related, for example, to spatial contents. This suggests that, in the beginning, in CLIL environments, teachers have to consider very carefully the contents to be taught through a foreign language with younger learners. They have to make a choice between what is vital to be taught in the mother tongue and what is wise to teach through the foreign language. With young learners, it would be good if the topics taught through a foreign language related mainly to the immediate environment of the learner. Later, on the other hand, when thinking processes have developed, the use of a foreign language as a medium of learning seems to be an advantage in cognitional development.

According to the findings of this study, it seems that Finnish CLIL environments support thinking and content learning, in particular, in situations where the learner has to compare different concepts and meaning schemes with each other. This is assumed to be due to the analogical CLIL reasoning systems that are based on exactly the kinds of situations where the learner makes comparisons between two semantic systems of two languages and two or more underlying cultures. CLIL learners may get special practice in classifying concepts and meaning schemes, in noticing and creating links between concepts and meaning schemes, and in hypothesising diverse things.

Learning in CLIL environments seems to be, in the beginning, more demanding than in environments where the mother tongue is the medium of learning. However, as the study shows, Finnish CLIL learners seemed to attain, over time, the necessary abilities and cognitional level. The main conclusion of the study is that a demanding and language-enriched learning environment has, in general, a positive effect on the Finnish mainstream CLIL learner’s cognitional development.

**Correspondence**

Any correspondence should be directed to Dr Aini-Kristiina Jäppinen, Ahde 4 P, 40520 Jyväskylä, Finland (jappinen@ktl.jyu.fi).

**Notes**

1. In 2002, there were in the Finnish comprehensive school 582,500 learners of whom about 10–15% was assumed to participate in CLIL, that is, about 60,000–90,000 learners. Although the sample is relatively small, triangulation of several factors increases the validity, such as three languages, two subjects, three cities, 12 schools, over 40 teachers, and eight tests in the course of almost three years.
2. The Finnish educational system consists of three different curricula: the national curriculum, the local curricula, and the individual curricula of the Finnish comprehensive schools.

References
Teaching Through a Foreign Language in Finland


Appendix 1. First Age Group: Science

Cognition: Conceptual invariance with an unknown factor
thinking category: realizing the constancy of properties
critical discovery learning area: the ability to solve problems that are difficult to explain to oneself or when concepts that one has not yet acquired are involved

M2: The topic of the test item was the phases of the moon. There was an explanation, ‘Although the moon is always round it seems to change its shape in the sky. This is due to the fact that sunlight is not always able to light up the whole surface of the moon because of the Earth. We call this transformation the phases of the moon.’ The learners were asked to connect the right picture, from four alternatives, with one of the concepts ‘full moon’, ‘waning moon’, ‘new moon’, or ‘waxing moon’. The explanation and the pictures showing the apparent transformations of the moon were regarded as conceptual invariance. The phases of the moon were not included in the curriculum, and it was expected that at least some of the four concepts would be unfamiliar to the learners because their teachers were not told to teach this particular content.

Cognition: Problem solving based on alternatives
thinking category: classifying according to more than one concurrent factor
critical discovery learning area: the ability to exploit the flow of information, to make comparisons and form antitheses, and to choose between two or more alternatives

M2: There was a picture of a rocket flying through the space. Around the rocket, there were some little stars, a small planet and a bigger planet with a circle around it. The learners were presented four groups of explanations, with three alternatives in each of them. They were asked to put the three explanations in order according to their probability by numbering, in each group, the explanation they considered the most plausible with number one, the next with number two, etc. For example, the first group concerned the recognition of the orbs. The three explanations were ‘in the picture, there are Jupiter and the Sun’, ‘in the picture, there are Saturn and Neptune’, and ‘in the picture, there are the Earth and Venus’. The other groups focused on the other elements of the space, such as, the relationship between time and distance or the relationship between the Sun and the planets.

M3: The learners were introduced to a weather forecast for a week in terms of visual or numerical symbols. They also got four verbal explanations relating to a lesser or greater degree to these symbols. The learners had to put the explanations in order according to their probability, as in the first follow-up measurement. One of the explanations fitted the symbols completely, and another diverged from it with regard to one fact. One explanation contained two diverging facts and one several diverging facts.

Cognition: Equivalence or similarity of a conceptual change
thinking category: realizing the compensation or equivalence of a change
critical discovery learning area: awareness of existing concepts and the ability to under-
M4: The learners were introduced to a picture about the circulation of water. They were asked to complete five phrases with the help of prompts in brackets: ‘when air rises, it…(what kind of change in temperature?)’, ‘the rising air is caused by … (what?)’, ‘humidity in the air becomes … (which state of water?)’, ‘when they have grown bigger, they … (what kind of change in the process?)’, and ‘water will never run out, because …’.

Cognition: Multifactorial classification of a schema

thinking category: classifying according to more than one concurrent factor

critical discovery learning area: awareness of meaning schemes and the ability to create links between them

M4: The learners were asked to link the words ‘humidity’, ‘gravitation’, ‘evaporate’, ‘groundwater’, ‘energy’, ‘temperature’, ‘water vapour’, and ‘nutrients’ to the most appropriate of the three alternatives ‘the Sun’, the Earth’, or ‘the air’.

Appendix 2. Second Age Group: Mathematics

Cognition: Charting conceptual alternatives and making choices between congruent concepts

thinking category: noticing and charting alternatives for action

critical discovery learning area: awareness of existing concepts and the ability to understand, use, apply, and explain them

M2: The multiplications $25 \times 5$, $15 \times 3$, $3 \times 44$, $82 \times 5$, $11 \times 35$, $120 \times 3$, $5 \times 45$, and $74 \times 11$ were presented. The learners were first asked to put them in order according to magnitude. Then they were asked to explain the strategies that they had used to solve the multiplications $82 \times 5$ and $74 \times 11$. Thirdly, they had to choose one of the multiplications and transform it into division(s). Finally, they were asked to choose two multiplications that they considered to most closely resemble each other, and to explain the reasons for their choice.

M3: The learners were presented a series of decimals: $9.2 – 4.3 – 1.05 – 9.45 – 1.4 – 12.6 – 19.9 – 4.15 – 6.7 – 13.8 – 15.2$. They had to place the decimals on a continuum on which the numbers 1, 5, 10, 15, and 20 were already marked. After that, they were asked to name those decimals the difference of which was 5.3, those the sum of which was closest to 10, and those the difference of which was closest to 5. Finally, the learners were asked which decimals in the series resembled each other most closely and on what grounds.

M4: The learners were asked to write the rational numbers $13 – 5.5 – 1/4 – 20 – 1/8 – 41/2 – 1.75 – 2/3 – 1/2 – 0.04 – 1/3$ in an order of magnitude starting with the smallest one. Then they should define, in their own words, a rational number. This was followed by the question ‘which rational numbers in the task are referred to here?’. The cues given were ‘their sum is 4.75’, ‘their difference is 5’,
‘their sum is closest to 0.3’, and ‘their difference is closest to 1.7’. Finally, the learners were asked which three rational numbers on the list, in their opinion, resembled each other most closely and why.

Cognition: Charting possible alternatives within existing meaning schemes in spite of ambiguous concepts

thinking category: changing possibilities into hypotheses and testing the hypotheses
critical discovery learning area: the ability to solve problems that are difficult to explain to oneself or when concepts that one has not yet acquired are involved

M2: The learners were presented a table where they found the prices for foreign calls to Sweden, Britain, United States of America/Canada, Germany, Italy, Estonia, and Australia. The prices were quoted by four different operators. The learners should answer the following questions ‘how much does the cheapest 3 minute call to Britain cost?’, ‘how much does the most expensive 5 minute call to Australia cost?’, ‘which operator offers an 8 minute call to Canada for €2.24?’, and ‘how much is the difference between the cheapest and the most expensive 10 minute call to Italy?’.

M3: A diagram showed the number of new computer games invented between 1998 and 2002. The years were presented in the diagram, but no exact figures about the number of games were given. The learners were asked to fill in the diagram with the figure that they thought to denote the number of games, with an accuracy of 10, on the basis of the information ‘in 2001, some 775 games were invented’. After that, the learners were asked ‘what is your estimate of the exact number of games invented in 2000?’ and ‘can you explain how you arrived at this alternative?’.

M4: The learners had to look at three recipes. The first was for waffles (about 8): 2 eggs, 5 dl milk, 3 dl flour, 1/2 tsp baking powder, 3/4 tsp salt, and 2.5 tbsp margarine when 1 tbsp equals 50 g. The second was blueberry pie (for 6): 100 g butter, 0.75 dl sugar, 2 1/2 dl flour, 1 tsp baking powder, 1/2 cream, 1 tsp vanillin sugar, and 0.5 l blueberries. The last was coffee-bread loaf (for 10): 500 g butter, 2.5 dl sugar, 1 egg, 1/2 l sour milk, 3 tsp soda, and 1.1 kg flour. Then the learners were asked to form three groups from the information given in such way that in one group all the quantities of the ingredients would mean approximately the same.
Copyright of Language & Education is the property of Multilingual Matters and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.