# Learning biology in Dutch bilingual education:

# Any reason for concern?

Linda W. Mulder

Utrecht University

3517616

30 ECTS

February 2016

Supervisor (FI): dr. Arthur Bakker

Advisor: prof. dr. Rick de Graaff

**Abstract**

Dutch bilingual education has been under investigation since the early nineties. Content and Language Integrated Learning (CLIL) is an educational approach that aims to help teachers teach the content of a school subject through a second language. Previous studies found that the final exam grades of students in bilingual streams (CLIL students) did not deviate from students in regular Dutch education (non-CLIL students), but it is still unclear what the influence of bilingual education is on students' comprehension of the sciences. In this study, research was conducted with 16 CLIL students and 49 non-CLIL students. Both groups were in their third year of secondary education and were taught by the same biology teacher. The CLIL and non-CLIL students did a biology test in English and Dutch, respectively. The answers were then analysed on: number of points (test scores), time needed to finish the test, the number of words used in the answers, the number of subject-specific words that were used and the percentage of correctly used subject-specific words. A significant difference was found in the test scores for CLIL ($M$ = 8.3, $SD$ = 2.72) and non-CLIL students ($M$ = 10.3, $SD$ = 2.95); $t(63) = -2.343$, $p = 0.022$, $d = 0.69$. However, none of the differences found for the other variables were significant. These findings point to a possible concern for the comprehension of biology content for CLIL students when taught and assessed in their second language, but more qualitative research on possible causes is needed.

*Keywords*: Dutch Bilingual education, TTO, CLIL, Science education, Biology education, Secondary school

There are many countries in Europe that have implemented some form of bilingual education in the past because of regional subdivisions, for example Belgium and Switzerland but also the Danish-German border (Baetens Beardsmore, 1993; Lorenzo, Casal, & Moore, 2010). In the more recent past, monolingual countries decided to offer bilingual education as well since they saw a need for English proficiency. Bilingual education has become increasingly more important in the Netherlands since its implementation in the early 1990s. At this moment, there are about 130 schools in the Netherlands offering bilingual education (Europees Platform, n.d.) in addition to regular education. Being taught in not only their first language (Dutch, their L1) but also in English (their L2) might cause problems for students in comprehending other subjects than English. Since the start of bilingual schools, researchers have been giving mixed images on the effects of bilingual education. The effects of bilingual education on students' grades have been studied multiple times and it was found that the final exam grades of students enrolled in this type of education did not deviate negatively from those of students in regular education (Admiraal, Westhoff, & de Bot, 2006; Huibregtse, 2001; Verspoor, de Bot, & Xu, 2015).

Most of the studies on bilingual education focussed mainly on the humanities and language subjects (such as history, geography and French), but not so much on the natural sciences (such as chemistry, biology and physics). Since language is important for learning science subjects, the effects of bilingual education on these subjects should be studied as well. This study aims to investigate whether the students in bilingual education are just as capable of answering biology test questions in English as regular students are in Dutch.

## Theoretical Framework

### Bilingual education in the Netherlands

In Dutch society, foreign language skills have been essential for many people since there have been increasingly more possibilities of mobility and communication across the border (Admiraal et al., 2006). Living, studying and working abroad has made it essential to learn at least one other European language. The noticeable place English has in Dutch society and, for example, the many universities in the Netherlands that offer Bachelor and Master programmes in English, make most students highly motivated to learn this language. The Dutch are known for high proficiency in English, which can be attributed to the educational system but also to the high exposure to English the students have outside of the classroom (Llinares, Morton, & Whittaker, 2012; Verspoor et al., 2015). The Ministry of Education recognised this as early as 1989, and appointed a task force to address the need for foreign language proficiency (Admiraal et al., 2006). This task force constructed a programme in which they suggested exploring the possibilities of bilingual secondary education. By 1993 there were already six schools that were offering a form of bilingual education (Huibregtse, 2001). Those schools were originally offering International Education to students who were temporarily in the Netherlands and decided to foster to the needs of parents and students and start offering bilingual education for Dutch students as well (Admiraal et al., 2006).

Bilingual education in the Netherlands is mostly implemented at the VWO level (pre-university education) and the second language is almost always English, although there are some German bilingual schools at the Dutch-German border (Huibregtse, 2001). The Dutch bilingual schools offer a type of education that is called Tweetalig onderwijs (TTO). They are fairly unrestricted in how they design their bilingual curricula, although they are regulated by some general rules set up by the European Platform (EP-Nuffic, n.d.). For example, they must ensure that the English bilingual education is not given at the expense of the subject of Dutch

language and literature or any other subjects. Therefore, students' grades for their final exams on other subjects, especially Dutch, should not deviate negatively.

The increased emphasis on language is apparent in bilingual education. On average, the students receive three additional hours of English as a subject and about 10 hours of other subjects taught through the English language. These extra English lessons could help them understand the linguistic aspects of these other subjects as well. In addition, the (other) subject teachers take extra time to explain subject-specific language during their lessons. For TTO schools, some 50% of the school subjects are allowed to be taught in English, and most schools decide to only teach subjects in English to the lower secondary levels (Admiraal et al., 2006). This is usually decided by schools due to scheduling issues since students in the Netherlands are enrolled in subject profiles from the fourth year of secondary school until their final exams. However, there are some schools in the Netherlands that offer their students bilingual education up to the fifth year of Dutch secondary school, equivalent to 11$^{th}$ grade.

**Evaluations of TTO in the Netherlands**

In 2006, a study by Admiraal et al. was published that was financed by the Dutch Ministry of Education to evaluate bilingual education in the Netherlands. Here, the focus was on TTO as a type of education in Dutch schools. No negative deviations were found for the subjects Dutch, history and geography at the final exams for students who received bilingual education during the lower years. However, these results were based solely on the results of the final exams (in the sixth year), and all subjects had been taught in Dutch in the three years prior to these exams. Moreover, Verspoor et al. (2015), argued that this study was set in the early years of the implementation of bilingual streams and might have been influenced by the introduction of an innovation (some sort of a Hawthorne effect). Therefore, Verspoor et al. decided to use mostly the same design and instruments, and replicate the Admiraal study. They compared students who received bilingual education to students who received education

in Dutch. Verspoor et al. found that the TTO students outperformed the other groups in English proficiency. Most of the progress was made in the first year of bilingual education and overall the development can be observed as an S-curve. This indicates that the development of English proficiency for students in their third year of bilingual education reaches a ceiling effect and by then, their English is of a higher level than students in regular education. The English proficiency of the students while learning other subjects was not tested by Verspoor et al., although other studies (de Goede, 2015; Gablasova, 2012) show that students in bilingual streams have problems with using their second language when studying a particular domain.

**Concerns about bilingual education**

In 2011, Bruton evaluated some of the studies on bilingual education in Hong Kong, Basque country and in Finland. He questions the positive outcomes of these studies and argued that most studies made a big error in their study designs. First of all, most of the studies did not take into account the possibility of higher initial proficiency scores of the students in bilingual education compared to regular education students. Rumlich (2013) did a study with German students and pointed out that students in the bilingual stream have higher initial English proficiency than students in regular streams. Furthermore, Bruton claims that "the students who opt for, and are very often encouraged into, the bilingual programmes are the highly motivated ones, whose parents are generally in the higher socio-economic classes" (p. 529). This would especially be the case when the students in the experimental and control groups where from the same school. Bruton argues that self-selection of students might take place when highly motivated students opt for the bilingual stream instead of the regular stream on the same school. Verspoor (et al., 2015) acknowledged Bruton's arguments, but also pointed out that:

> L2 development is a dynamic, complex process in which several factors interact over time with non-linear effects, and that the positive effect of bilingual education is more than the sum of its parts (p. 6).

To deal with the issues raised by Bruton (2011) and Rumlich (2013), Verspoor et al. (2015) added a third group to their study, where most studies only compared two groups. They studied students who were enrolled in the TTO stream (called bilinguals), regular VWO students (regulars) and Gymnasium students (controls). This last group consisted of students in a stream called Gymnasium, a grammar stream in which the students (in addition to the regular subjects) have the subjects Greek and Latin. This stream is only meant for the brightest students and is considered very challenging, therefore most schools decide to select students for this stream based on scholastic aptitude. This stream is very comparable to the TTO stream that also has mainly high scholastic aptitude students and is similarly considered selective for motivation. By adding this last group, Verspoor et al. showed that the difference they found could not be attributed to higher scholastic aptitude of the TTO students. Bruton's and Rumlich's arguments, however, do remain relevant for future research on bilingual education, and self-selection of students should always be taken into consideration.

Content and Language Integrated Learning (CLIL) is an educational approach that aims to help teachers teach the content of a school subject through a second language and is used in TTO schools throughout the Netherlands. CLIL is predominantly used in Europe, but also in parts of Australia and other countries in the world (Whittaker, Llinares, & McCabe, 2011). It has been recognised as being one of the most important tools to achieve the goals set by the European Union and more specifically the European Platform in the Netherlands (EP-Nuffic, n.d.; Dalton-Puffer, 2008). Whittaker et al. obtained the written output of CLIL students for over four years and analysed them on coherence, prepositional phrases and attributive adjectives. Whittaker et al. found that students developed significantly over four

years in the control of textual resources and were able to provide more complex written output. It is also suggested that when students are in a CLIL environment, they are provided with a suitable context to develop written discourse. Llinares et al. (2012) pointed out that language used in a CLIL classroom is very different from language used in a language classroom and often requires specific grammar and vocabulary.

**Science in two languages**

Science subjects in schools have long been considered mostly empirical and practical topics. However, many students struggle with the language, words and terms of science subjects (Wellington & Osborne, 2001, p. 3). Since students are believed to struggle with language when learning science subjects, it raises the question whether students in bilingual education might encounter even more problems when they are taught the science subjects in their L2 (Airey, 2010; Wellington & Osborne, 2001). On the other hand, it is suggested by Lemke (1990) that students learn science best by understanding the language of science in such a way that they can use it themselves to describe scientific phenomena. Therefore, it could also be the case that students in bilingual schools benefit from the extra emphasis on reading, writing and speaking during the years they are in bilingual education. Taking these things into consideration, it is interesting to find out what influence bilingual education has on students' learning about science.

Airey (2010) raised the question if "students taught in L2 learn their subject as well as if they were taught in L1?" (p. 35). He stated that when the language in which the content is taught changes, the already existing communication problems might be accentuated. He also argued that students taught in their L2 might not be able to access their prior knowledge on the topic discussed in class. To test this, he looked at three groups of undergraduate Physics students in Sweden. Two of the groups were taught one class in English and one in their L1, Swedish. The other group had one class that was both taught in English and in Swedish. Airey

(2010) found that although a few students were not able to adjust to the change of language, the majority of students were unaffected by it. When graded on a four-point scale on correctness from a physics perspective, strong agreement between the ratings in both languages can be seen, with a maximum of a point difference. He argues that the few students who were not able to adapt, might be able to do so over time but otherwise would have to leave this type of education. What is interesting is how this study translates to bilingual education and if students in Dutch TTO schools would also be just as able to explain science content in their L2 as regular students in their L1.

**The language of science**

In the natural science subjects, language is mainly used to explain, but also to describe and analyse scientific phenomena and it is used when writing, speaking, listening and reading in science classes (Dale & Tanner, 2012; Wellington & Osborne, 2001). Throughout primary and secondary school, students learn about the way they can express themselves more academically during reasoning, problem-solving and evaluating complex subject matter. For example, Mercer, Dawes, Wegerif, and Sams (2004) found that when students were encouraged to talk and reason about science, it "leads to better learning and conceptual understanding in science" (p. 374). Not only do the natural science subjects deal with specific grammar (passive in the simple past), there are also subject-specific and general academic words to be learned. The subject-specific words differ per subject, but mostly also per topic within that subject.

Llinares et al. (2012) pointed out that there are three basic types of text in science lessons: procedures, reports and explanations. Many biology lessons also contain additional types of texts, for example newspaper articles and food labels. From these texts, the students need to obtain, analyse and understand information in order to answer the questions correctly. In addition to the input being provided in a foreign language, the output also asks for high

linguistic ability of the students, especially when students are asked to respond in an essay-like manner. They are then not only marked on their subject-specific knowledge, but also on the way they have constructed their arguments.

**Aim and research question**

TTO students have been instructed in English for biology since the beginning of secondary school. Since this is their L2, it might mean that when the students need to answer a test questions using a lot of words, they are held back because of language issues. This project aims to investigate whether the CLIL students are just as capable of answering biology test questions in English (their L2) as regular students are in Dutch (L1). The main research question of this research is:

> What is the difference in level of ability to answer biology questions between CLIL students answering in L2 and non-CLIL students answering in L1?

The 'ability to answer biology questions' will be operationalised in different steps described in the next section.

A significant influence of bilingual education on the students' performance could have implications for the implementation of bilingual education for biology. Since all students need to do their final exams in their first language (Dutch, their L1), it is important to make sure there are no negative effects for the comprehension of biology content for students that have done bilingual education for a number of years. This project intends to add to the evaluation of bilingual education in the natural science subjects.

## Method

The aim of this research calls for a quasi-experimental study design for which the type of education is the independent variable and the ability to answer biology questions the

dependent variable. In this study, the answers of both CLIL and non-CLIL students on a biology test will be compared and analysed.

**Participants**

The participating school is a large school of about 2000 students which offers secondary education at HAVO, VWO and Gymnasium level. It offers a bilingual stream to both VWO and Gymnasium students. The school meets the quality standards set by the European Platform and is therefore officially recognised as a senior TTO school.

For this research project, two groups were compared, CLIL students and non-CLIL students. The group of CLIL students are represented by 16 VWO students from a third year class who are in the bilingual stream, and included 10 females and 6 males, ranging from 13 to 15 years old. The CLIL group is represented by only 16 students since the class originally consists of both VWO and Gymnasium students, but only the VWO students received biology lessons in the third year. Therefore, the CLIL groups consists of only about half of that class, 16 students in total. They are compared to 49 VWO students from two classes from the same school who are in the Dutch stream, of which 21 were female and 28 male, also ranging from 13 to 15 year old. All three classes were taught by the same biology teacher, a female with 16 years of experience in teaching.

Third year students have been chosen for this study for two reasons. First, during the first two years of secondary education the students are still developing in English proficiency and they are still getting familiar to the CLIL method. Studies show that a ceiling effect of English proficiency is reached during the third year of bilingual education (Verspoor et al., 2015). Since this project tries to investigate their biology comprehension when taught and assessed in English, third year students seem most appropriate. Second, after the third year all subjects are taught in Dutch and not in English, so selecting students from higher years seemed inappropriate.

Since this research compares two groups, TTO and VWO, scholastic aptitude of both groups have to be determined. During the last year of primary school, approximately 160,000 students in the Netherlands participate in an end test (Cito Netherlands, n.d.). This test consists of multiple choice questions in different subjects and is administered by Centraal Instituut voor Toets Ontwikkeling (CITO). The overall score for this test is used by primary schools as an indicator for the stream level in secondary school that is most appropriate for each student since it is considered to give a good overall view of scholastic aptitude. The moment the Cito tests are distributed to the students is the last moment they have not yet received bilingual education. Therefore, this is the best moment to compare the initial levels of scholastic aptitude for both groups.

The Cito scores of both groups were compared using an independent-samples *t* test. The scores of the CLIL group ($M$ = 547.1, $SD$ = 2.09) were significantly higher than the non-CLIL group ($M$ = 543.6, $SD$ = 4.38); $t(63) = 2.975$, $p = 0.004$, $d = 1.02$. This was to be expected since one of the conditions from the school for students entering the bilingual stream was a Cito score of 545 or higher, although individual cases with lower scores will be considered as well.

**Materials and procedure**

To investigate the influence of bilingual education on the ability to answer biology questions, a test was constructed for both groups. To compare the answers of the CLIL and non-CLIL students, the tests only differed in language to ensure internal validity. The test language was the same as the language in which the students were instructed. The topic of the test was behaviour and both groups finished the chapter on behaviour and made a major test on this topic one week prior to the lesson that the test of the current study was given. Therefore, the students were expected to be able to understand the biological content of the text. Since the analysis of the answers focused on the use of language, the tests included

essay-like answers. For the test, a text was used from ‘Bio-Aktueel’, a magazine developed by teachers containing news articles about biology. All texts in this magazine are taken from newspapers and other magazines. The texts are accompanied by questions about the texts and correct answers are given as well. The text ‘Koolmees is een sociaal dier’ was taken from the May 2015 edition together with the questions and both were translated into English (see Appendix 1) by the researcher. This translation was then checked by a native English speaker on correctness of language.

The test consisted of nine questions in total, for which a maximum of 19 points could be obtained. The students received a mark for the test that counted one time for their overall mark for biology. The questions were in such a manner that they ask for biological terms, but also explanations on biological processes. Some of the questions were altered to match the level of the students, since the original questions were aimed at fourth year students. In addition, some extra guidance was given to indicate in how many words the students should answer, to ensure enough data to compare both groups.

The teacher was instructed to teach the two different groups as she would have usually done in terms of teaching methods. In addition, she was asked to use English 100 percent of the time in the CLIL class while teaching this biology module. The lessons prior to the test were observed in two classes (one CLIL, one non-CLIL) by the researcher to ensure treatment fidelity. Notes and video recordings were made during the lessons that were observed. Analysis of these recordings showed no major differences that could harm the internal validity. The teacher used the same teaching techniques in both lessons and used English the majority of the time in the CLIL class.

**Data Analysis**

The answers of the students were transcribed verbatim by the researcher. The test answers were analysed per student on the following aspects:

1. Test scores. A marking scheme was made by the researcher in cooperation with the teacher (see Appendix 2), by modifying the marking scheme distributed by Bio-Aktueel. The answers of the students were marked according to this marking scheme and the total number of points out of 19 the students received for their answers was analysed. The answers of the students were not only marked by the researcher but by a second rater as well and there was moderate agreement between the two raters.

2. The time needed for answering the whole test. Since the CLIL students will answer in their L2, it is probable that these students need more time to read the text and answer the questions than the non-CLIL students reading and answering in their L1. To test this, the students wrote down the time it took to make the test.

3. The number of words the students used to answer the whole test. Since the test asks the students questions about their understanding of the text, longer answers could indicate a deeper understanding.

4. The number of different subject-specific words that were used and the percentage of these subject-specific words that is used correctly. Since Lemke (1990) points out that using scientific language is important for understanding, the use of more subject-specific words in their answer could indicate a deeper understanding of the text. To test this, the researcher first made a list of all the subject-specific words used in the answers of students. Then, this list was checked by a fellow biology teacher and shortened to make the final list of subject-specific words (Table 1). Finally, the number of subject-specific words was counted per student and in accordance to the marking scheme the percentage of correct ones was calculated.

The performance of the groups was compared using SPSS. Tests were done to see if the groups differed significantly on the variables and effect sizes were calculated to investigate the influence of bilingual education on the ability to answer biology questions.

Table 1 *Subject-specific word list*

| Dutch words | English words |
|---|---|
| Aanpassen | Adaptation |
| Concurrentie | Competition |
| Evolutie | Evolution |
| Fourageergedrag | Foraging behaviour |
| Jager verzamelaar vaardigheden | Gathering/hunting skills |
| Imitatie | Imitation |
| Inzicht | Insight |
| Inprenten | Imprinting |
| Klassiek conditioneren | Classical conditioning |
| Leefomstandigheden | Living conditions |
| Nakomelingen/nageslacht | Offspring |
| Operant conditioneren | Operant conditioning |
| Overlevingskans | Chance of survival |
| Populatie | Population |
| Prikkel (sleutel/neutraal) | Stimulus |
| Rangorde | Pecking order |
| Reflex | Reflex |
| Skinnerbox | Skinner box |
| Soort | Species |
| Survival of the fittest | Survival of the fittest |
| Trial and error | Trial and error |
| Uitgestorven | Extinct |

## Results

There was a significant difference in the test scores for CLIL ($M$ = 8.3, $SD$ = 2.72) and non-CLIL students ($M$ = 10.3, $SD$ = 2.95); $t(63) = -2.343$, $p = 0.022$, $d = 0.70$. However, the differences found for the time the students needed to finish the test, the number of words used in the answers, the number of subject-specific words that were used and the percentage of correctly used subject-specific words, as shown in Table 2, were not statistically significant. Effect sizes were also very small, suggesting that these other differences do not have practical

Table 2. *Differences between CLIL and non-CLIL groups*

| | CLIL ($n$ = 16) | | Non-CLIL ($n$ = 49) | | | | Cohen's |
|---|---|---|---|---|---|---|---|
| Variable | *M* | *SD* | *M* | *SD* | *t*(63) | *p* | *d* |
| Cito scores | 547.1 | 2.09 | 543.6 | 4.38 | 2.975 | 0.004 | 1.02 |
| Test scores | 8.3 | 2.72 | 10.3 | 2.95 | -2.343 | 0.022 | 0.70 |
| Time | 21.7 | 4.33 | 22.3 | 4.82 | -0.456 | 0.650 | 0.13 |
| Total number of words | 145.7 | 28.6 | 143.3 | 33.0 | 0.260 | 0.795 | 0.08 |
| Subject-specific words | 4.94 | 1.77 | 6.00 | 2.88 | -1.385 | 0.171 | 0.04 |
| Percentage correctly used | 70.2 | 21.8 | 76.4 | 24.8 | -0.895 | 0.374 | 0.27 |

significance either. In addition, the relationship between the test scores and the other variables was determined, but no strong or moderate correlations were found.

## Discussion

Since a statistically and practically significant difference was found between the CLIL and non-CLIL students in terms of test scores, it can be concluded that we found a difference in ability to answer these biology questions for these CLIL students. The difference in test scores between the CLIL and non-CLIL students cannot be explained by the initial levels of scholastic aptitude of the students. The Cito scores of the CLIL students were significantly higher than of the non-CLIL students, suggesting they should score the same or even higher on their test. Furthermore, none of the other variables that were researched can give an explanation for these differences, since none were found to be significant or to be correlated with the test scores.

When studying the answers of all students, some additional observations can be made. One of the most noteworthy observations was that one of the CLIL students did not understand what norms are. She thought norms were some sort of an animal and started copying the birds:

> Norms that were close by the great tits and that saw what they were doing and that each great tit chose to do the same, began to copy them to increase their chance of survival. So when one Norm started acting the same as the great tits, the others followed him because it became the preferrance. [sic]

This CLIL student clearly had trouble understanding the word *norm*, and since this word is the same in Dutch, it would be interesting to see if she would have understood this term if she would have read the entire text in her L1, Dutch. This is a clear example of a CLIL student not being able to understand a vital part of the text because of one word not being understood, and it raises the question if more students had trouble understanding vital parts of the text, that were not so apparent in the answers.

The additional findings of the current study suggest some students were not able to understand the English text fully. Other studies indicate that students in bilingual education perform better when they have read a text in their L1 than in their L2. For example, Gablasova (2012) found that students in bilingual education who read a history or geography text in their L2, were less able to reproduce the meaning of the technical words than those who have read the text in their L1. De Goede (2015) did the same study in the Netherlands and found L2 readers from the third year performed better than L1 readers, but for fourth year students the results were reversed. It would be interesting to see if CLIL students perform better on the test from the current study if they would have read the text in their first language. Such results could show that the underlying cause of the lower test scores from the CLIL students in the current study might really be caused by the language of the text. Further research could therefore be done by comparing CLIL students reading the text in L1 and in L2.

Another difference between the two groups is the number of unanswered questions, when only taking into account the blank spaces left by students who did not need the

maximum number of minutes to finish the test. Nine blank spaces for the 16 CLIL students and ten blank spaces for the 49 non-CLIL students seems like a substantial difference and might indicate the CLIL students were unable to formulate an answer for some of the questions. Whether this shows the CLIL students did not understand some of the questions, remains uncertain. Although the mark for this test counted for the overall mark for biology, it could still be that some students were just too inattentive to answer. Interestingly, the number of blank spaces also suggests that the CLIL students used more words in the rest of their answers, since the number of words in total did not differ significantly between the groups.

Another aspect to consider is the way the Cito scores are calculated. Although they give a pretty good overall image of the scholastic aptitude of students in this research, the grades for the different sections, including reading comprehension, are not known. Therefore, following the method of Verspoor et al. (2015), a third group of Gymnasium students could be added in future research to add a group that is more comparable in motivation and self-selection.

In general it can be stated that for this research, although a significant difference was found, no conclusions can be drawn about the causes of the lower results the CLIL students displayed. Although the deviations in test scores were big enough to find significant differences with only 16 CLIL students, more data might give more qualitative examples to explain these differences. Hence, more in-depth research and qualitative analysis of a larger group is needed to find possible causes for these results.

This study aimed to contribute to the knowledge and to take another step in fully understanding the influence of bilingual education on students' performances when using a great deal of text in a test. Although this study was carried out on a small scale, and with only one teacher on one school, it does raise a concern for learning biology in Dutch bilingual education. It is necessary to carry out more research with CLIL students on different schools,

to see if these results can be generalised. In addition, it needs to be studied how CLIL students perform in the other sciences and mathematics, especially on topics that require a great deal of reading. If studies show that CLIL students have trouble with texts in the other sciences as well, this could have implications for schools when considering which subjects will be taught in English during the first years of bilingual education. Since the results indicate the CLIL students have trouble understanding the text and formulating answers in their L2 for biology, this has implications for teachers as well as students. For teachers teaching in bilingual education, it might be wise not to use too much text in a test if the goal is to test for comprehension. Similarly, Llinares et al. (2012) mentioned that teachers should be careful when using language-testing materials in formative and summative assessments. They state "the danger is that, by doing so, we may be in fact assessing a non-academic competency" (p. 283). For students opting for bilingual education it is important to determine if the possible negative effects the stream has on other subjects can be compensated by their motivation and ability.

Although the results of this study points out that there is a concern for learning biology in Dutch bilingual education, another note has to be made. Many other studies show the positive effects of bilingual education on students' performances in the languages, which was one of the goals when starting TTO schools in the Netherlands in the early nineties. And these positive effects might in fact still outweigh the possible negative effects it has on other subjects, as long as the negative effects do not become too substantial.

**Acknowledgements**

First of all, I would like to thank Arthur Bakker for all the time he has put into supervising me. His notes and comments since the beginning of this project have been invaluable and the encouragement has kept me focussed while finishing it. I would also like to thank Rick de Graaff for his advice on procedures and for sharing his knowledge on CLIL practices and TTO schools. Lastly, the feedback on early drafts I received from my brother Mark Mulder, fellow student Lysanne Smit and my boyfriend Gijs van Swaaij have helped me tremendously and I could not have done it without their support.

## References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English 1. *Educational Research and Evaluation, 12*(1), 75–93.

Airey, J. (2010). The Ability of Students to Explain Science Concepts in Two Languages. *Hermes*, 45(August), 35–50.

Baetens Beardsmore, H. (1993). *European models of bilingual education* (Vol. 92). Multilingual Matters.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System, 39*(4), 523–532.

Cito Netherlands. (n.d.). Retrieved October 26, 2015. http://www.cito.com/about_cito/cito-offices/cito_netherlands

Dale, L., & Tanner, R. (2012). *CLIL Activities with CD-ROM: A Resource for Subject and Language Teachers*. Cambridge University Press.

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

De Goede, L. (2015). Content learning in the L1 and in the L2 : A comparative study of CLIL students in grade 9 and grade 10 in Dutch secondary schools.

EP-Nuffic Scholen. (n.d.). Retrieved December 11, 2015, from https://www.epnuffic.nl/voortgezet-onderwijs/talenonderwijs/tweetalig-onderwijs/scholen

Gablasova, D. (2012). *Learning and expressing technical vocabulary through the medium of L1 and L2 by Slovak-English bilingual high-school students* (Doctoral dissertation, ResearchSpace@ Auckland).

Huijbregtse, I. (2001). Onderwijs in twee talen. *Levende Talen Tijdschrift*, *2*(1), 11-20.

Koolmees is een sociaal dier. (2015, May). Bio-aktueel, 21(65), 2827-2828

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

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

Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in european education: Key findings from the andalusian bilingual sections evaluation project. *Applied Linguistics, 31*(3), 418–442.

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3) 359–377.

Rumlich, D. (2013). Students' general English proficiency prior to CLIL: Empirical evidence for substantial differences between prospective CLIL and non-CLIL students in Germany. *Content and language integrated learning (CLIL) in Europe: Research perspectives on policy and practice*, 181-201.

Verspoor, M., de Bot, K., & Xu, X. (2015). The effects of English bilingual education in the Netherlands. *Journal of Immersion and Content-Based Language Education, 3*(1), 4–27.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education.* McGraw-Hill Education (UK).

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research, 15*(3), 343-362.

## Appendix 1

A great tit brings a worm to his nest box. © KINA

# The great tit is a social animal

Great tits copy the behaviour of other animals from the same species and have social norms. Therefore, traditions arise that will be permanent. An international team of biologists stated in Nature that this social behaviour was previously only seen in primates and humans.

(Translated) Pepijn Van Der Gulden 4 december 2014, 04:00

Researchers observed how wild great tits learned a new way to feed themselves. The birds learn from each other and rather choose the accepted method than a different method that gives them the same amount of food, but is unpopular in the group. Thereby, the great tits conform to the norms of the group.

Bird researcher Christiaan Both, Professor of Groningen University, is pleased with the results. 'The research was already discussed a lot. We knew that birds learn from each other, but the fact they also pass on norms, is a new discovery. A local culture arises.'

To introduce new eating behaviour, the researchers built a feeding cage. In the cage they put mealworms, a delicacy for great tits, that could be reached by opening a red or blue door with their beak. A few great tits that were trained to open the cage, but only by using one of the two doors. The trained birds were then put back in various great tit populations in a forest near Oxford, where the cages were also positioned. With small transmitters, researchers tracked which great tits ate the mealworms and via which doors.

### Norms of the group

A great tit is building his nest. © ANP

After twenty days, three quarters of the great tits found their way to the mealworms, more than in the populations where no trained great tits were put. The majority of the animals chose for the door that the trained birds were taught to open.

The great tits conformed to the norms of the group for blue or red. The great tits who were able to open the feeding cage in both ways, preferred de colour that was dominant in the group. Great tits that moved into a group somewhere else in the forest also copied the accepted method. Even when the feeding cages were put back after being absent for a whole year , the preferred colour of the group stayed the same. This is special since only two-fifths of the animals were in the original experiment.

## Competition

The advantage of the group behaviour can possibly be explained evolutionary, says Both. ‘If everyone uses a successful method, it is smart for other to copy it. Although it could also go wrong, when animals have to compete for the same food.

Norms seem likely to also be present for other species that live in stable groups, Both presumes. Innovation in behaviour by learning from others might explain fluctuations in population numbers. ‘Animals like storks and beech martens were almost inexistent a few decades ago, but are currently present again, without a clear reason. Maybe the living conditions were improved, but there is also a possibility the animals adapted their behaviour to increase their chance to survive.’

# Appendix 2

## The great tit is a social animal

Topic: Behaviour

**Questions**: *(total: 19 points)*

1 *Great tits copy the behaviour of other animals from the same species and have social norms.*
Give the **biological term** for the learning process in which behaviour is copied.
Imitation (2p)

2 *To introduce new eating behaviour, the researchers built a feeding cage. In the cage they put mealworms, a delicacy for great tits, that could be reached by opening a red or blue door with their beak. A few great tits that were trained to open the cage, but only by using one of the two doors.*
a Give the **biological term** for the learning process in which the new eating behaviour is trained. **Explain your answer.**
Operant conditioning, (1p)
Classical conditioneren (0,5p)
The bird has to do something (open a door) / achieve something (1p)
To receive a reward (a worm) (1p)

b Give the **biological term** for the feeding cage as used under experimental conditions.
A Skinner box / learning machine (1p)

3 *'The research was already discussed a lot. We knew that birds learn from each other, but the fact they also pass on norms, is a new discovery. A local culture arises.'*
**Explain** what the researchers observed to conclude that the great tits pass on norms, just like humans and primates. **Use 40-60 words to do so.**
Great tits who could open the door in both ways, would prefer the colour which was dominant in their group (1p)
When great tits moved to other groups, they would copy the dominant method (1p)
The researchers observed that the great tits would conform to the group (1p)

4 *After twenty days, three quarters of the great tits found their way to the mealworms, more than in the populations where no trained great tits were put. The majority of the animals chose for the door that the trained birds were taught to open.*
Formulate **in one sentence** a **conclusion** using these observations.
The great tits can copy the trained behaviour of other great tits (2p)
Copy norms / without word behaviour / without example (1p)

5 *The advantage of the group behaviour can possibly be explained evolutionary.*
a **Explain** what the evolutionary advantage is of this type of behaviour for the great tits.
Other animals found a method that is successful (1p)
Copying this behaviour increases your chance of survival (1p)

b **Explain** what the possible disadvantage is of this type of behaviour for the great tits.
The great tits use the same method to get to the food (1p)
There will be competition / fights for the same food (1p)
A bad method will lead to a decreased chance for survival (0,5p)

6 *Innovation in behaviour by learning from others might explain fluctuations in population numbers.*
a **Explain** this speculation..
The successful method will first lead to a greater population (1p)
Competition will then lead to less population growth (1p)
Or:
If they all behave the same, they will either all be successful or all die (2p)

b **Give** 2 factors that can cause fluctuations in population numbers.
Food, predators, diseases, living conditions, behaviour, natural disasters, competition, climate changes (2p)