Academic year 2015-2016

First examination period

The Development of Cognitive Control and Working Memory During Second Language Acquisition: a Longitudinal Study

Thesis submitted for the degree of Master of Psychology,

option Theoretical and Experimental Psychology

Promotor: Dr. Eva Van Assche

Copromotor: Dr. Evy Woumans

01101494

Sofie Ameloot

Acknowledgements

My thesis would not have been possible without the support and guidance of my supervisor dr. Eva Van Assche and co-supervisor dr. Evy Woumans. Thank you for supporting this research idea and being enthusiastic about designing an experiment suitable for children. The longitudinal study, and the amount of participants resulted in a challenge that I could only complete successfully thanks to your help. You have been my (co-)supervisors on several projects now, and I loved every minute of it.

Dr. Emmanuel Keuleurs, thank you very much for helping me with the statistical side of this study. When I was overwhelmed with the amount of data and was not sure how to analyse it, you took the time to explain and teach me the needed statistics.

I would also like to extend my gratitude to the schools, the parents, and the children. The parents gave me the opportunity to learn, by trusting me to test their child. The kindness of the principals and teachers made testing at the school a very rewarding experience. The children, who were the amazing superstars of this study, made from every test a game with its own story ("the flower is fighting with the heart") and told me everything about their dinosaur collection, which led to some great memories.

My fiancé also deserves a special thanks for being my tower of strength, and for proofreading my thesis several times. Arno, your humour helped me to put my stress into perspective and to relax. Even though I talked too much about "my thesis children", you kept listening and giving me advice. You gave me the courage to keep trying, and to pursue my dreams.

I would also like to thank my friends and family that came with me to Wallonia to help me. Ruben, I know the small chairs were terrible for your back, but you remained the awesome, energetic cousin I have always known. Delphine, Elyse, Leentje, Laura and Dieter, you got really early out of bed on your free days to spend a day or more trying to keep the children concentrated on hearts, flowers and puzzles. Thank you very much for being so enthusiastic about "Jaques, le petit poison" over and over again!

Last but not least, my family gave me, as always, unconditional moral support. Steven, thank you so much for revising my French translations (since it took me some time to get used to French again, I know I gave you a huge job). Oma and opa, you are always there for me, and no words can describe how grateful I am for your support and love. You are the best grandparents a granddaughter can wish for. Papa, I know I talked a lot about my research. Thank you for indulging me. Katrien, even though you were miles away while I was doing the first part of my thesis, you were always there for me when I needed some motivation.

Sofie

Ghent, May 15, 2016

Abstract

Bilingualism is thought to result in advantageous effects, like better developed cognitive control (CC; e.g. Costa & Sebastian-Galles, 2014) and better evolved working memory (WM; Morales, Calvo, & Bialystok, 2013). Further research has shown that the reverse effect is also possible; CC predicts artificial language learning (Kapa & Colombo, 2014). However, an artificial language is not as complex as a normal spoken language. This study aimed to investigate the effect of IQ, CC and WM on language acquisition by examining children receiving immersion education. At the first time point (T1), we tested 60 French children starting in a Dutch immersion education program the subsequent school year. At the second time point (T2), approximately a year later, 43 children were tested again. Only 35 participants were included in the analysis. They were administered a range of tests: an intelligence test (Raven Progressive Matrices), language tests (Peabody Picture Vocabulary Test in French at T1 and T2 and Dutch at T2), CC measurements (DCCS and flanker task) and a WM task. Reaction times and overall interference effects were measured, as well as accuracy on language tests and on an intelligence test. IQ scores were significantly higher at T2 and French scores were marginally significant higher. Using linear regression modelling, we found that intelligence, CC and WM predicted the scores on the Dutch vocabulary tests. Looking into the progression rate of the native language, we found that SES and CC were a main predictor for native language development. Implications for future research are discussed.

Keywords: bilingualism, immersion education, working memory, cognitive control, language acquisition, longitudinal, children

Abstract

Tweetaligheid wordt geassocieerd met voordelen, zoals beter ontwikkelde cognitieve controle (CC; bv. Costa & Sebastian-Galles, 2014) en een beter ontwikkeld werkgeheugen (WG; Morales, Calvo, & Bialystok, 2013). Verder onderzoek toonde aan dat het omgekeerde effect ook mogelijk is; CC voorspelt artificiële taal acquisitie (Kapa & Colombo, 2014). Echter, een artificiële taal is niet zo complex als een normaal gesproken taal. Dit onderzoekspaper had als doel het effect van IQ, CC en WG op taalacquisitie verder te onderzoeken door kinderen te testen die deelnamen aan immersieonderwijs. Op T1 testten we 60 kinderen die het volgende schooljaar zouden starten aan een Nederlands immersieprogramma. Op T2, ongeveer een jaar later, testten we 43 kinderen opnieuw. In totaal werden 35 participanten opgenomen in de analyse. Een batterij aan testen werd afgenomen: een intelligentie test (Raven Progressive Matrices), taaltesten (Peabody Picture Vocabulary test in Frans op T1 en T2, en in Nederlands op T2 voor de immersiekinderen), CC-metingen (DCCS en flanker taak) en een WG-taak. Reactietijden en algemene interferentie-effecten werden gemeten, evenals accuraatheid op taaltesten en op een intelligentietest. IQ was significant homer op T1 en Franse scores warn marginal significant hoger. Op basis van een lineair regressie model, vonden we dat IQ, CC en WG de taalkennis van het Nederlands op T2 voorspelden. De vooruitgang in de moedertaal tussen de twee testmomenten werd significant beïnvloed door SES and CC. Implicaties voor verder onderzoek worden besproken.

Trefwoorden: tweetaligheid, immersie onderwijs, werkgeheugen, cognitieve controle, taalverwerving, longitudinaal, kinderen

Table of Contents

The Development of Cognitive Control and Working Memory During Second Language Acquisition: a Longitudinal Study

In 1982, it was estimated that approximately half of the world knows a second language (Grosjean, 1982). This resulted in a renewed increase in bilingualism research. Some researchers found that bilingualism leads to advantageous effects (e.g. Morales et al., 2013), while others are contradicting this (e.g. Paap & Greenberg, 2013; Hilchey & Klein, 2012). Interestingly, the majority of studies has looked into the effects of bilingualism, and not into variables that could influence second language acquisition. Therefore, this study investigated language, cognitive control, working memory and intelligence before and after children came in contact with a second language. Identifying these variables are of great importance for a better understanding of (second) language acquisition, and for offering better acquisition programs and support.

Before explaining the findings of this study, we will first look into the research on bilingualism that has been done the past century. Bilingualism has led to different opposing views, which we will discuss briefly. Then, we will go over the natural development of cognitive control. Bilingualism has been linked several times to an advantageous effect on cognitive control, but cognitive control is no static skill. It evolves rapidly during childhood, and continues to develop until it deteriorates in old age. The link between bilingualism and cognitive control has been investigated thoroughly, and is therefore discussed more deeply in the next part. Theoretical views and research findings are put forward and weighted against each other. Then, the influence of bilingualism on WM is discussed, since cognitive control and working memory are strongly related, and working memory is an important variable when learning a new language. After the research on the advantageous effects of bilingualism, the critique on these findings is discussed. A solution to address both sides, namely a longitudinal design, is put forward and implemented in this study. Finally, the current proposal is put forward. This study tested kindergarteners to investigate the variables predicting language acquisition, and how language acquisition influences these variables.

Research on Bilingualism in the Past

Bilinguals are often defined as "those people who need and use two (or more) languages in everyday lives" (Grosjean, 1982, p. vii). The influences of bilingualism, for instance the social implications of knowing two languages, have been investigated for almost a century and the views on the effects of bilingualism have varied. The first studies on monolingual and bilingual children concluded that bilingualism led to disadvantageous effects. Decroly (1926) tested Walloon and Flemish children of different ages on IQ. Since the Walloon children, who were monolingual, scored higher, he concluded that monolingual children had more advanced "school" intelligence. It should be noted that the IQ test was in French, which was the Flemish children's second language. However, using non-verbal IQ tests, Arsenian (1937) found the same effect of bilingualism on IQ; monolinguals scored higher than bilinguals. Apart from IQ, studies also found that bilingual children knew less vocabulary (Grabo, 1931) and monolingual children had better writing skills and grammar (Saer, 1923).

In order to increase the quality of research on bilingualism, Arsenian (1937) suggested five conditions that should be stated in definite and objective terms for every bilingual. The first was the degree of bilingualism, which means that proficiency in both languages is tested. For the second condition, the differences between the languages needed to be taken into account. Arsenian reasoned that more similar languages would be more easily learned, and would have a more overlapping culture. The age of acquisition and the method of learning also needed to be recorded. Learning a language by playing in the schoolyard with a bilingual friend, or learning a language in class would result in different types of language skills. Finally, religious, national and political attitudes towards the second language had to be taken into account since it affected the willingness to study a second language. While other conditions are now deemed more and more important to take into account (e.g. socio-economic status (SES) and IQ), two of the five conditions are still used in research nowadays, namely language proficiency and age of acquisition.

The negative effects of bilingualism described in these studies have later been attributed to the lack of control for SES, which was often lower in the bilingual group

(McCarthy, 1930) and for IQ, which is associated with low SES (Fischbein, 1980). Apart from methodological issues, the former prevalent and negative view on bilingualism (e.g. “It may be assumed that children with high bilingualism will be more retarded in their school progress than children who have a low degree of bilingualism”, Arsenian, 1937, p. 115) has also been attributed to historical events at that time window (Fitzgerald, 1993). At the beginning of the 20th century, World War I encouraged a feeling of nationalism. As a consequence, bilinguals were often seen as outsiders and had a negative connotation. After World War II, this negativity subsided to a lesser degree (Fitzgerald, 1993).

The first positive effects for bilingualism were found by Davies and Hughes (1927) and by Stark (1940), who both concluded that bilinguals were superior to monolinguals in intelligence. However, the former did not measure the degree of bilingualism of the participants, and the latter did not take age, gender and SES into account. Null effects were also found when measuring the difference on IQ between bilinguals and monolinguals (e.g. Darsie, 1926). In 1962, the negative findings were countered for the first time reliably by Peal and Lambert. They too stated five variables that needed to be controlled in bilingual research: age, gender, SES, degree of bilingualism and the used tests. Taking these variables into account, the bilinguals performed better on both verbal and nonverbal intelligence tests. One of the explanations for these positive findings was that bilinguals had greater mental flexibility, since bilinguals appeared to have a more diverse set of mental abilities. The diversity was attributed to experience in switching between languages (Peal & Lambert, 1962).

Eventually, no consensus was reached and research on bilingualism faded into the background until it revived again when a study reported that bilingual children had more metalinguistic awareness. Bialystok (1988) tested two skill components of metalinguistic awareness, namely grammaticality judgment and form-meaning selection. For the grammaticality judgment task, the children were asked if the offered words were existing words and sentences. For form-meaning selection, the participant had to define a word. The conclusion was that bilingual children performed better on these tasks, since they already had more experience from having two linguistic systems. Bilinguals have at least two terms for one semantic representation, compared to the monolingual's one linguistic system. A few years later, Bialystok (1992) connected metalinguistic awareness to cognitive control (CC). CC (also known as executive functioning and executive control) is essential for adaptation of behaviour towards a certain goal and for processing information. Current research is now focussing on the advantageous effects of bilingualism, such as a better development of cognitive control.

Development of Cognitive Control

Cognitive functions, like inhibition and working memory, play an important role in everyday life. The development of the CC system, which supports processes like working memory, is one of the most essential processes in childhood (Diamond, 2002). It is known that CC develops rapidly while growing up, especially between the ages of three and six (Best & Miller, 2010).

Zelazo (2015) recently developed a new framework to understand the development of CC, by using the iterative reprocessing model (e.g. Cunningham & Zelazo, 2007). Based on Miyake et al. (2000), Zelazo (2015) defined CC as three skills: cognitive flexibility, working memory, and inhibitory control. Cognitive flexibility is the skill to switch successfully between two tasks, working memory is essential for holding, processing and manipulating information, and inhibitory control is the ability to ignore irrelevant stimuli and focus on the relevant stimuli (e.g. ignoring the incoming text message on your phone in order to focus on completing a paper). Zelazo (2015) put forward that it were not CC skills that develop during childhood, but rather the ability to reflect upon information. Self-reflection was seen as an essential step in perceiving a stimulus and acting upon it. Young children are not yet able to reflect as deeply and as quickly on responses as adults do, therefore they will act more impulsively. See Figure 1 for an overview of the model.

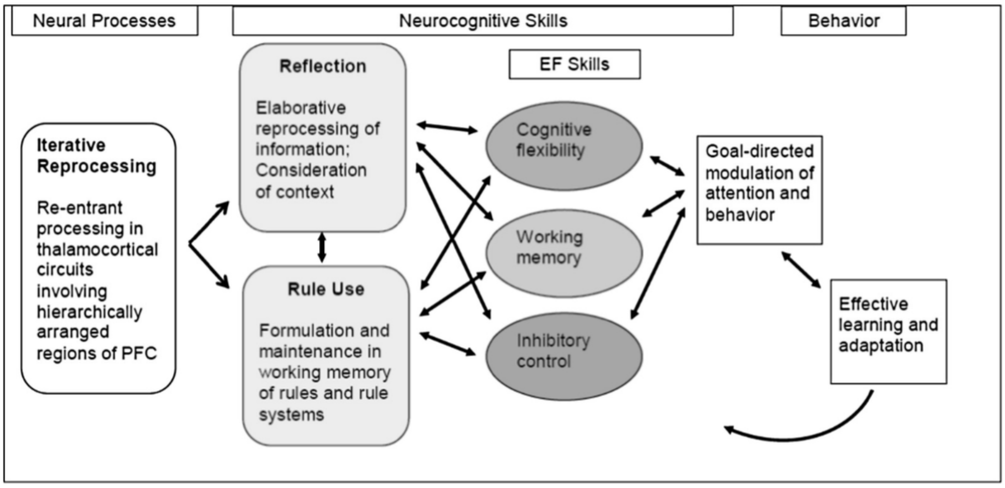


Figure 1. The development of CC (also named Executive Functioning, EF) according to the iterative reprocessing model. Figure copied from Zelazo (2015).

One of the most well-known experiments testing CC is the Stanford marshmallow experiment. Here, children sit at a table with a marshmallow on a plate. The child is told that if it can wait 15 minutes, it will get two marshmallows. If it does not wait, it will only get the one marshmallow on the plate (Mischel, Ebbesen, & Raskoff Zeiss, 1972). This experiment is a very clear example of CC because children have to adapt their behaviour (i.e. waiting instead of eating) to reach a certain goal (i.e. getting two marshmallows). One of the important factors that predicts the successful attempt to wait is age. The older the child, the better it it able to wait. It was later shown that the children who are successful in this experiment, obtain higher SAT scores (i.e. a test in the United States of America that every student needs to partake in order to go to university), are more cognitively and socially competent, and are better at coping with stress and frustration (Mischel, Shoda, & Rodriguez, 1989).

Since age was an important predictor for success, other studies have looked into the changes in CC at a young age. Cragg (2016) compared 7-, 10- and 20-year old participants on their skill in response inhibition using a flanker task. In this task, participants have to react to the direction of the central stimulus (< or >) which is surrounded by two flankers on each side, leading to congruent (<<<<< or >>>>) or incongruent (<<><< or >><>>) trials (see *method* section for a more extensive description). Cragg (2016) found that the youngest group had significantly less response inhibition than the 10-year old children, while the two oldest groups did not perform differently. Carlson (2005) tested 602 children ranging in age from two to six years old on several CC tests. He reported an age difference in performance on CC tests, therefore also attesting for the development of CC during this age period. Therefore, it can be assumed that the largest changes in CC occur in young children, implying that this time period is crucial for development of CC.

Bilingualism and Cognitive Control

Some studies have shown that bilingual children develop their cognitive control system earlier than their monolingual peers (e.g. Bialystok, 2001). This has been attributed to the fact that bilinguals have to control the languages they know. Green's (1998) model of inhibitory control can be used to explain these beneficial effects. This theory states that because both languages of bilinguals are always activated (Marian, Spivey, & Hirsh, 2003; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009), bilinguals experience a constant need of attentional control and language inhibition, which eventually leads to better general cognitive control (i.e. also outside the linguistic domain), called the 'bilingual advantage'. For example, when speaking in Dutch about a *fiets* (which means *bike* in Dutch), the English translation *bike* needs to be inhibited. Hence, CC is necessary to suppress the unneeded language, and to select the correct language.

To study bilingualism, children are an interesting group to study. Compared to monolingual children, bilingual children receive approximately twice the amount of different verbal input, but still reach the same linguistic milestones (De Houwer, Bornstein, & Putnick, 2013), thereby attesting that bilingualism has no immediate negative effect on linguistic development. Moreover, research showed that bilingualism had an advantageous effect on CC even in young children, which is investigated thoroughly (e.g. Costa & Sebastian-Galles, 2014; for a review, see Bialystok, 2009).

When we look at the bilingual advantage, the strongest evidence seems to concern young children and ageing adults, suggesting that bilingualism mainly benefits cognitive development and cognitive decline (Bialystok, Craik, Klein, & Viswanathan, 2004; Schweizer, Ware, Fischer, Craik, & Bialystok, 2011; Struys, Mohades, Bosh, & van den Noort, 2015; Woumans et al., 2015; Woumans, Surmont, Struys, & Duyck, in press). Regarding cognitive development, beneficial effects are reported in children from birth to age six (e.g. Crivello et al., 2016; Martin-Rhee & Bialystok, 2008; Morales et al., 2013). For example, the difference between monolingual and bilingual upbringing has been found in preverbal infants of 7-months-old (Kovács & Mehler, 2009a). Monolingual and bilingual infants were taught that responding to a certain verbal cue would lead to a reward, namely seeing a puppet. The children learned that the cue predicted the location of the puppet, which resulted in directing their gaze to that location after the cue, and before the puppet appeared. After the habituation phase, new cues were introduced that predicted the opposite location of the puppet. Only bilingual infants were able to learn that the cue now predicted the other location of the puppet. In a second study, Kovács and Mehler (2009b) investigated this effect further with 12-month-old children. These children were also taught that a cue predicted the location of the reward. Only the infants who came in contact with a second language on a daily basis were able to learn that two different cues could predict the same location. The researchers concluded that this advocated for bilinguals being more flexible learners.

The advantage in CC as a result of bilingualism has been found in various contexts. Crivello et al. (2016) tested bilingual and monolingual toddlers on language and CC in a longitudinal design. They found that the rate of increased vocabulary growth predicted the better performance on conflict tasks, like the shape Stroop task. In a shape Stroop task, the participants have to name the shape they see (e.g. circle or square), while a word is printed in letters over the shape. The word can be congruent (e.g. *circle* printed on a circle) or incongruent (e.g. *square* printed on a circle) with the shape. Since bilinguals performed better on this task, Crivello et al. (2016) concluded that the exposure to a second language leads to CC advantages. The same effect was found by Struys et al. (2015). They tested children who were bilingual from birth and children enrolled in immersion education (starting at the age of three) on CC and verbal fluency in their second language (L2). Immersion education offers children education in two languages, where both languages are used as a medium of instruction. Therefore, immersion education is a good setting for natural contact with an L2. Often, L2 is educated in an interactive way (see Appendix A for an example). In the study by Struys et al. (2015), both groups had the same level of language proficiency in L1 and L2. While there was no difference on verbal fluency, there was a difference between both groups in CC. Learning a second language did not lead to a disadvantage in L1 verbal fluency, but early bilingualism led to advantages on CC for bilingual children compared to children who learned an L2 later in life.

Even though Struys et al. (2015) found a difference between bilingual children and L2 learners on CC, previous research showed that language acquisition in immersion education can be compared to becoming bilingual and also resulted in the positive outcomes associated with bilingualism (Hermanto et al, 2012). Poarch and van Hell (2012) studied monolinguals, bilinguals, L2 learners, and trilinguals with respect to CC. The bilingual and trilingual children showed significantly better performance on CC, while the L2 learners had a numerically -but not significantly- better performance compared to monolinguals. These results were interpreted as a possible emerging advantage of L2 learners. Furthermore, no significant disadvantages were found between L2 learners, bilinguals, and trilinguals. In the same line, Carlson and Meltzoff (2008) compared monolinguals, bilinguals, and children receiving immersion education using 10 measurements of CC. However, no significant difference was found for the immersion children compared to the monolinguals, but this could be due to the short period of immersion education. These children came in contact with a second language for only 6 months. Yet, the bilingual group differed significantly from the monolingual group on several tasks, in particular on the conflict CC tasks.

Studies with balanced bilinguals (i.e. bilinguals that use both languages an equal amount of time, and have learned both languages as a child) showed that cognitive and phonological processing abilities are linked to L1 and L2 lexical development

(Gathercole & Baddeley, 1989; Segalowitz & Hulstijn, 2005).

The link between CC and second language acquisition has also been found in the opposite direction. Indeed, it has been shown that CC predicts the success in learning a second language. Nicolay and Poncelet (2013a) aimed to determine to what extent these abilities were linked to L2 acquisition through immersion education. They studied the cognitive abilities of 5-year-old kindergarteners enrolled in English immersion classes in a longitudinal study during three years. Once a year, the participants were administered a speech perception task, phonological awareness task, phonological short term memory task and several tasks measuring attentional and executive skills. All tested abilities, except phonological awareness and response inhibition, appeared to be associated with the beginning of L2 acquisition in the immersion school context. Nicolay and Poncelet (2013a) concluded that CC predicted the overall success in L2 acquisition of bilinguals compared to monolingual control groups. However, these participants had already started with the immersion program at the first test moment (time point 1; T1), so no clear baseline was used in the analysis, the test scores (e.g. IQ test and tests for the native language) were not standardised, and different second language tests were used between time points. Therefore, more research is needed with standardised tests and a baseline.

The predicting effect of CC, when measuring a baseline, on artificial language acquisition has also been found. Kapa and Colombo (2014) presented 4-5 year old children with three different CC tests. The first, the Attentional Network Test (ANT; Fan, Wu, Fossella, & Posner, 2001), is a measurement of response inhibition. This task uses a cue to alert the participant, then a cue orients attention towards the upcoming stimuli, and finally arrows are shown. The participant has to indicate the direction of the central arrow, which can be congruent or incongruent with the surrounding arrows (see Figure 2). This task is very similar to the flanker task. In the second task, a Simon task, participants were asked to press left when seeing a red cue, and right when seeing a green cue. This cue was presented on the left or right side of the screen, thus resulting in congruent and incongruent responses (see Figure 2). The third task, DCCS (Dimensional Change Card Sorting task; Zelazo, 2006), is a sorting task where cards first had to be sorted by one dimension (e.g. colour) and then sorted by another dimension (e.g. form; see *method* section for a more extensive description). This switch in sorting rules made the task more difficult for younger children. WM was assessed with a digit span task, meaning that the participant had to repeat a series of digits. This series increased in length when the participant answers correctly. Knowledge of English was tested with the Peabody Picture Vocabulary Task – 4. During this task, the participant had to indicate which of the four presented pictures represents the orally presented word (see *method* section for a more extensive description).

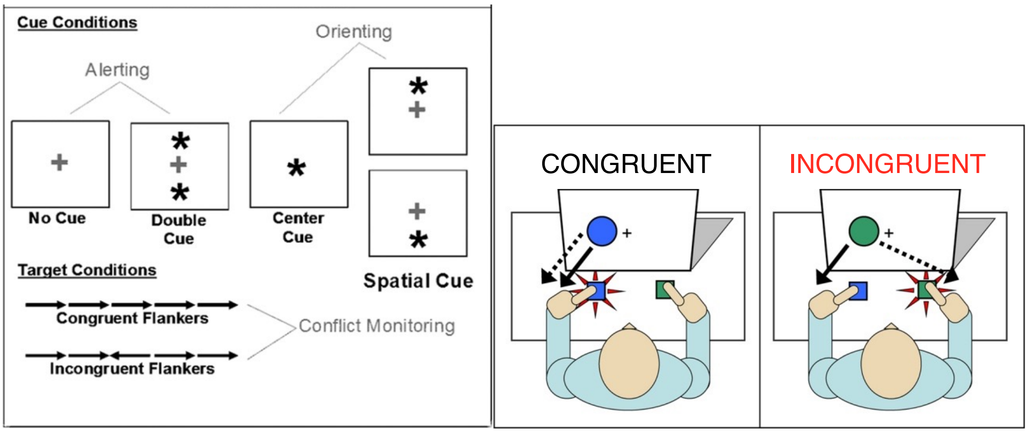


Figure 2. Left: ANT task. This tasks uses different cue conditions to alert and orient the participants. Different target conditions lead to congruent and incongruent flanker. This task measures conflict monitoring. Right: Simon task. This task uses congruent and incongruent stimuli to measure response inhibition.

Following these baseline tests, participants learned an artificial language via training videos and picture books for a total duration of 180 minutes spread over two days. After the artificial language acquisition phase, the children were assessed on their success of acquisition. As hypothesised, WM significantly predicted the outcome of artificial language acquisition. Participants' English knowledge was not a significant predictor for language acquisition. Controlling for English vocabulary scores and WM, CC was also a significant predictor. However, only the measurements of DCCS significantly predicted the outcomes, and performance on the ANT was a marginally significant predictor. Kapa and Colombo (2015) concluded that CC predicted artificial language acquisition.

Bilingualism and Working Memory

As CC is incorporated in all working-memory models (e.g. Baddeley, 1992) and CC is needed to update working memory (WM; e.g. Garon, Bryson, & Smith, 2008), the influence of bilingualism on WM has also been investigated. Morales et al. (2013) tested 5- and 6-year-old children - monolingual and bilingual - with tasks that manipulated WM demands. The first task used a Simon-like design (see *method* section), thereby manipulating conflict resolution, while the second task manipulated other CC aspects. On both tasks, bilinguals responded faster and more accurately. Bilinguals outperformed monolinguals especially on trials where more CC was needed. Kaushanskaya, Cross, and Buac (2014) tested children participating in immersion education on word learning, CC, short-term WM, and verbal WM. Children participating in immersion education scored higher on verbal WM and word learning, thereby attesting that bilingualism also positively affects WM.

The influence of bilingualism on WM was also tested with children differing in SES. Blom, Küntay, Messer, Verhagen, and Leseman (2014) used visuospatial and verbal WM tests with monolingual Dutch children, and bilingual Dutch-Turkish children. The bilinguals had a lower SES compared to the monolinguals. An advantage was found for the bilingual children, on one visuospatial WM task and on the verbal WM task. The researchers concluded that bilingualism improved WM regardless of the low SES background. This effect was previously also reported by Carlson and Meltzoff (2008) and Engel de Abreu et al. (2012).

Critique on the bilingual advantage

Reading the previous paragraphs, it would be logical to deduce that there is a clear link between speaking a second language and enhanced cognitive control. However, there is currently a strong debate with researchers who claim that bilingual advantages are actually insubstantial (e.g. Paap & Greenberg, 2013; Antón et al., 2014). This led to a special issue of the journal *Cortex*, devoted to this discussion. Paap, Johnson, and Sawi (2015) advocated that cognitive advantages as a result of bilingualism do not exist, or are limited to exceptional and undetermined circumstances. They argued that most positive findings were likely to result from the fact that different groups of monolinguals and bilinguals are typically compared on CC, but that these groups are not always appropriately matched on all relevant variables (such as IQ or SES). As Woumans and Duyck (2015) pointed out in a reaction to Paap et al. (2015), the sole possibility to exclude confounding variables is to use longitudinal studies where bilingualism becomes a variable over time. Such a rare longitudinal design was used by Woumans et al. (in press). They tested 5-year-old children before (T1) and after a year (T2) of monolingual or immersion education on IQ. All children were matched on IQ, SES and L1 knowledge at T1. At T2, the children that followed immersion education scored higher on IQ compared to the children following standard education. Immersion education led to a significant increase in IQ scores compared to standard education. However, no effects on CC were found.

The effects on WM are also not reliably found. For example, Engel de Abreu (2011) followed 6- to 8 year old children during three years, and tested them repeatedly on language, fluid intelligence and working memory. Fluid intelligence and WM did not differ significantly between bilinguals and monolinguals, while monolinguals scored higher on the language tests. Engel de Abreu (2011) concluded that a possible explanation for the lack of effects on WM tasks is that bilingualism trains certain aspects of CC. WM, one of the three aspects, is not trained by bilingualism. In a reaction to these statements, Calvo, Ibanez, and Garcia (2016) wrote an opinion paper stating that it might be possible that WM as a whole was not trained by bilingualism, but that there was evidence for enhanced WM performance in bilinguals. When WM demands were high in the task, bilinguals surpassed monolinguals (e.g. Blom et al., 2014; Morales et al., 2013).

Furthermore, other studies did not find any beneficial effects of bilingualism in children over the age of six (e.g. Antón et al., 2014; Martin-Rhee & Bialystok, 2008).

Abdelgafar and Mouawad (2015) tested 7- to 10-year-old children on a battery of CC tests. For most of the tests, no significant differences between monolinguals and bilinguals were found. Duñabeitia et al. (2014) tested a large sample of children (*n* = 504) between the ages of 8 and 13 on a different CC tests. No significant differences were found between monolinguals and bilinguals for all age groups.

The current proposal

The controversy around the effects of bilingualism leads to new questions regarding the substantiality of the advantageous effect of bilingualism and how CC and bilingualism are linked; more research on bilingualism is necessary to better understand the link between CC and bilingualism. Also, previous research (e.g. Carlson & Meltzoff, 2008; Nicolay & Poncelet 2013a) has neglected the natural development of CC and WM, and the influence it could have on learning a second language, since these variables were not measured before children came into contact with a second language. Kapa and Colombo (2013) already showed that CC and WM predicted success in learning an artificial language, but artificial and natural languages differ in complexity.

The aim of this study was to investigate how CC and WM determines the success of second language acquisition, and how second language acquisition influences CC and WM compared to monolingual children. However, since not enough monolingual children were recruited, the second aim could not be addressed in this study. By using a longitudinal design instead of matched designs, confounding variables could be excluded from the analysis. We conducted one experiment in which CC and WM were measured twice: before L2 acquisition and one year later, when the participants had been immersed in their L2 for almost one school year. Firstly, we expected that CC and WM would improve over time (e.g. Best & Miller, 2010). Secondly, we anticipated that the scores on CC and WM (at T1 and the progress between T1 and T2) would predict the language proficiency on Dutch at T2: the faster the reaction times and the lower the interference effect, the higher the vocabulary scores (e.g. Kapa & Colombo, 2014). Thirdly, we hypothesised that IQ, CC and WM would also influence language improvement of the mother tongue (e.g. McClelland, et al., 2007). Initially, we also wanted to test monolinguals longitudinally, so we could compare their scores and progress to bilinguals. However, of the seventy-one schools that were contacted, only one responded, resulting in too few participants.

Method

For this experiment, seventy-one schools with French education and thirty-one schools with an immersion program in the regions Brussels and Walloon Brabant were contacted by email (see appendix B for the information letter for the immersion schools, Appendix C for the French schools). Each letter was addressed personally to the principle of the school. If no answer was received after a week, a reminder was sent. When the principle agreed, further arrangements were made to prepare for the first test point. If the school required permission from the mayor's office, they were also contacted (see appendix D). Seven schools with an immersion education program and only one school with a traditional French education program confirmed. Since one immersion school offered an English programme, while the others schooled their students in Dutch, we tested only the French-Dutch children to ensure a group as homogenous as possible.

All participants were tested on CC using two tests, a flanker task (Eriksen & Eriksen, 1974) and dimensional change card sorting task (DCCS; Zelazo, 2006), on WM with one test (Morales et al., 2013), on IQ with one test (Raven Progressive Matrices; Raven, 2000) and on language (Peabody Picture Vocabulary test). At T1 and T2, they all completed the French language tests, at T2 the immersion children also performed a Dutch test. Each child was tested individually. The total duration for the test battery was between 30 and 45 minutes per child.

Participants

At the end of second kindergarten (May - June 2015), we started to study 74 children who had attended only French-speaking kindergarten (36 male, 38 female, M = 58.53 months, SD = 3.46). Of these 74 children, 13 children were raised with two languages at home. All children were tested so we could do additional comparisons between monolinguals, bilinguals and L2 learners, but this analysis is not included in this study. In September 2015, 60 children started with immersion education and 14 with standard French education. All children were recruited from six different schools, which were all located in the same French part of Belgium. The main difference between schools offering the immersion education and schools offering the standard education, is the language in which skills and competences are taught. In immersion education, Dutch is used in an interactive way to teach the children approximately 50% of the time. The children are taught for instance, poems, specific vocabulary related to a task (see Appendix A) and songs.

All children were tested at the end of the school year, since their immersion education would start in the next school year. Before the experiment started, all parents were contacted through the schools. Information letters, questionnaires, and informed consents had to be filled out (see Appendix E). The description of the hypothesis in the letter was kept vague, so the parents would not know we would make comparisons between educational programs. The questionnaire included questions about the participant's and parents' linguistic background and SES. The questionnaire also inquired after learning disorders, problems with language development, comprehension or sight problems. No problems were indicated. Not taking the bilingual group into account, all parents were monolingual and none of the children were exposed to another language that French.

At the second time point, March - April 2016, we retested 42 of the 60 children that were enrolled in an immersion program (22 male, 20 female, M = 68.12 months, SD = 2.44) and 13 children enrolled in standard education (6 male, 7 female, M = 70.46 months, SD = 2.44). All participants were in their third year of kindergarten. Seventeen participants of the immersion program were not tested when there were measurement errors at T1 (n = 2), because they were sick at home (n = 5), changed schools (n = 2), the demographics deviated (n = 2), or because the school did not respond in time (n = 7). One student of the standard program was not tested due to sickness. This study included only the participants who did not come into contact with a second language before the immersion program, were enrolled in the immersion program since September, and completed all tests (n = 35), so six children were excluded since they were bilingual, and one child did not complete all the tests. These children are from now on referred to as bilinguals. See table 1 for the demographics of these participants. If the parents and school consented, the children received a reward (stickers and a stamp) after the experiments were completed.

Materials

Échelle de Vocabulaire en Images Peabody. The Échelle de Vocabulaire en Images Peabody (EVIP) is a French translation of the Peabody Picture Vocabulary Task – Revised. This norm-referenced language assessment can be used for participants between 2.5 and 18 years old to measure receptive vocabulary. The test exists of 170 items, but only 25 to 50 items needed to be administered to determine the score of the participant. Each item consisted of four black and white pictures, presented on a card in multiple choice format. Participants were required to choose the picture that best depicted the word that was read aloud by the experiment leader. The test administration was not timed, but lasted approximately 15 minutes. In the analyses, the number of correct items will be the dependent variable. See Figure 3 for an example.

Peabody Picture Vocabulary Task III – Dutch. The Peabody Picture Vocabulary Task III – Dutch (PPVT) is similar to EVIP. This receptive vocabulary test is norm-referenced for participants between 2.3 and 90 years old. The test can also be used to measure the knowledge of Dutch as a second language. It consists of 204 items, each comprising of 4 pictures. For each item, our participant had to choose the correct picture for the verbally presented word. Once more, the administration of the test had no time limit, but lasted between 10 and 15 minutes. The number of correct answers will again be the dependent variable in the analyses. Since the participants just started learning Dutch, test administration started with item 1 instead of the age-based start item. See Figure 3 for an example.

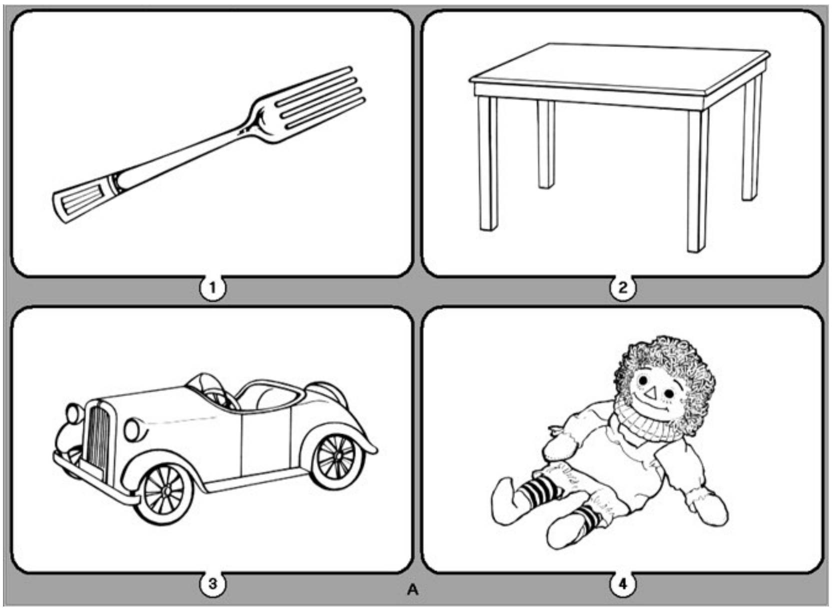


Figure 3. Example of an item of the Peabody Picture Vocabulary Test. When the experimenter says “table”, the child is supposed to point at picture 2 for a correct answer. Any other picture is considered as a error.

Dimensional Change Card Sorting task. The Dimensional Change Card Sorting task is a measurement of cognitive flexibility. Our participants were presented with two target pictures that varied according to two dimensions, namely colour and shape. The participant had to sort a series of bivalent cards according to one dimension, in this case colour, in the pre-switch trials. On post-switch trials, participants were asked to sort the cards according to the other dimension, in this case shape. Pre-switch and post-switch trials existed each of eight stimuli, and each stimulus (red rabbit, blue rabbit, red boat or blue boat) occurs two times. In the border trials, the sorting rule was randomised. If a border was presented around the stimuli, the participant had to sort based on shape. The lack of border meant the relevant dimension was colour. This switch occurred randomly eight times. The border trial consisted of 16 presented stimuli. The task was programmed in TScope5 (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006), based on the experiment and employing the stimuli of Zelazo

(2006). No time limit was set for the response. The instructions were given at the beginning of a trial and the relevant dimension was repeated for each stimulus. The administration of the test lasted for approximately five minutes. The number of correct responses and the reaction times were the dependent variables. See Figure 4 for a visualisation.

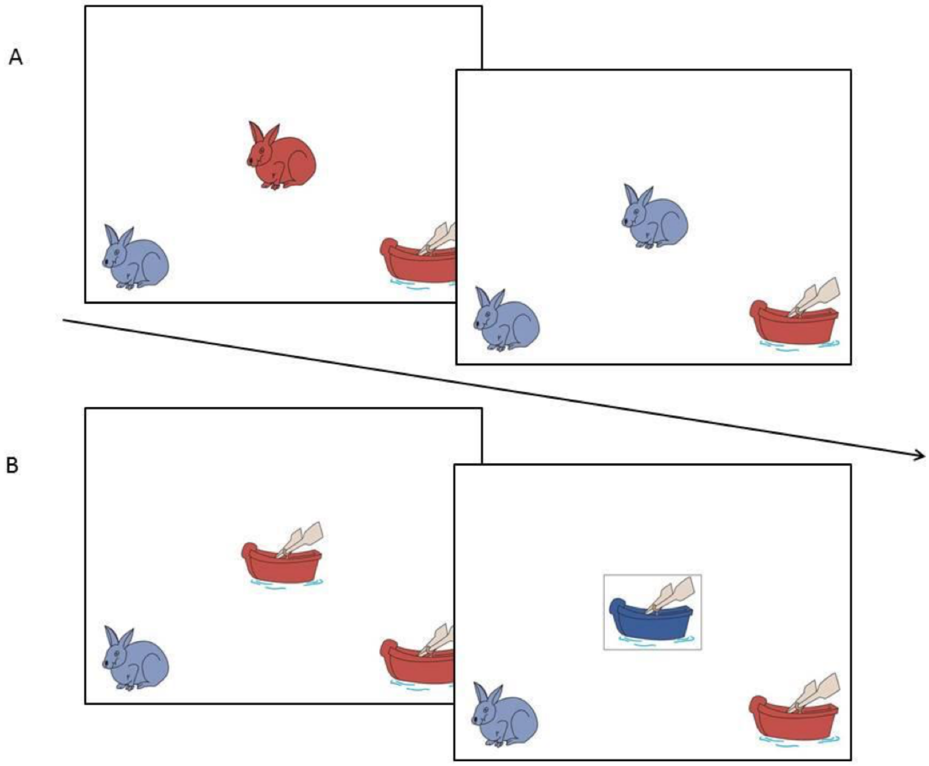


Figure 4. DCCS task. (A) procedure for the pre- and post-switch trials. (B) procedure for border trials.

Flanker task. To measure cognitive inhibition, a flanker task was used. The task was based on the experiment developed by Eriksen and Eriksen (1974), and was programmed in TScope5 (Stevens et al., 2006). In this task, a central stimulus, < or >, was surrounded by four more ‘flankers’ (two on each side), which were congruent or incongruent with the central stimulus. This resulted in four possible presentations of stimuli: <<<<<, >>>>>, <<><< and >><>>. To make the task suitable for children, the arrowheads were replaced by images of a fish. The subjects were told that the middle fish was named ‘Jacques’ and that they had to indicate the direction in which Jacques was swimming by pushing the corresponding button. These buttons (Q and M on an AZERTY keyboard) were labeled with stickers with a fish swimming to the left (Q) and right (M). The trial started with a fixation cross that remained on the screen for 500 ms, followed by the stimuli. The stimuli remained on screen until the participant responded or for a maximum of 3,000 ms. The inter-trial interval was 750 ms. Participants first completed a practice block of 10 trials, where feedback was given if they responded too slowly or incorrectly (“wrong” or “too slow” in red). This way, the experiment leader was able to follow their progress and give verbal feedback. Then the percentage of errors was given (again for the experimenter’s convenience) followed by the instructions. If the experimenter deemed it necessary to complete the practice block again (e.g. when the participant did not pay attention), the task was started over. This only happened once. After the second practice block, the participant continued. The experimental block consisted of 68 trials. At 34 trials, the instructions reappeared and an optional break was suggested. All children continued the task without a break. See Figure 5 for a visualisation.

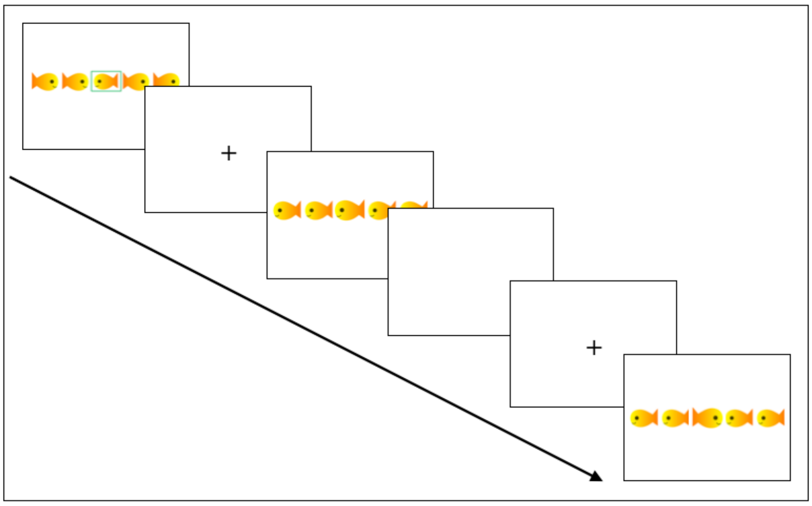


Figure 5. Flanker task. Each block starts with the instructions, where the target stimuli is outlined in green.

Working-Memory task. This task was based on the task used by Morales et al. (2013). The picture task was programmed in Tscope5 (Stevens et al., 2006) and was presented on a laptop with a 15-inch monitor. The task consisted of four blocks of 40 trials and this resulted in a total of 160 trials. Each block stood for one level of the 2 x 2 design: (2 stimuli vs. 4 stimuli) x (central presentation vs. side presentation). The first and second block consisted each of two different stimuli, while the third and fourth block consisted of four different stimuli. In the first and third block, the stimuli were presented in the middle of the screen, whereas they were presented randomly left or right of the screen in the second and fourth block. The first and third block started with a practice blocks of four trials. The second and fourth block had a practice block of eight trials. Each trial started with a fixation cross for 500 ms, followed by the stimulus that remained on the screen for a maximum of 3,000 ms. This was followed by an inter-trial interval of 500 ms. Instruction screens appeared before the start of a new block. The target stimuli were shown on a coloured background which responded to the response key (e.g. when the heart was shown on a purple background, the participant had to respond by pressing the purple key). The left and right Shift keys were indicated by a yellow and purple sticker. These colours were chosen since they were different from the colours of the stimuli. Response mapping was counterbalanced across all participants. See Figure 6 for a visualisation.

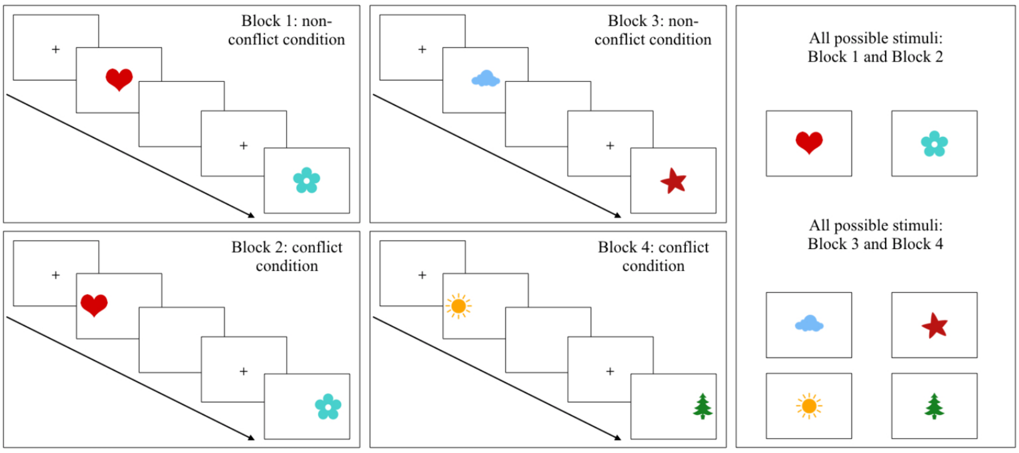


Figure 6. Working memory task. Left: procedure for the conflict and non-conflict conditions of block 1 and 2. Centre: procedure for the conflict and non-conflict conditions of block 3 and 4. Right: all possible stimuli.

Raven Coloured Progressive Matrices. To measure fluid intelligence, Raven Coloured Progressive Matrices was administered (Raven, 2000). Importantly, this test is a non-verbal measurement of fluid intelligence. Therefore it can be used to measure general cognitive development, independently of linguistic development. This test is suitable to test children of five years old. There are three different sets: A, AB and B, each consisting of twelve geometrical patterns with one piece missing. In this study, each pattern together with its six possible completion options were presented on an A4 page. Participants were asked to complete the puzzle, by choosing one of the six presented pieces. See Figure 7 for an example. Percentile scores were calculated from the raw scores according to the manual (Raven, Court, & Raven, 1998). These percentiles were used for further analysis. The manual offers percentile scores for children between 48 and 120 months old, for every six months. Extrapolation was used to calculate the scores for ages in between according to the equations in the manual. A percentile score of 50 is equal to the mean score of the population, namely 100.

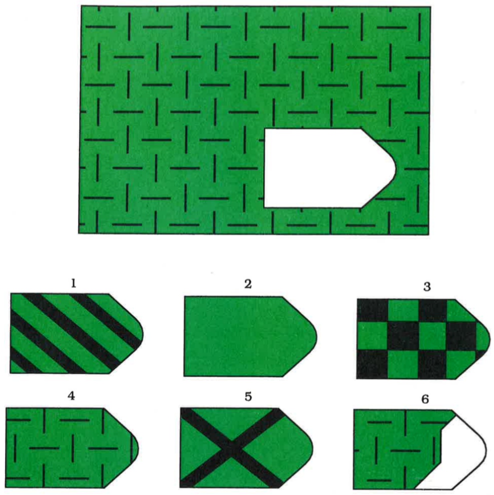


Figure 7. Example of a geometrical pattern with the six pieces.

Results

Background measures

Only the children that completed all tasks and were not bilingual from birth, were included in the analysis. This resulted in 35 children from the immersion schools, and 9 children from the French school. Since not enough monolingual children were tested, no analysis for this group was computed. All participants were tested twice on each test, except on Dutch, with approximately 9 to 10 months in between. Scores of the French test and IQ test were standardised. The presented scores of the Dutch test were not standardised because all children came in contact with Dutch for the same amount of time, which was independent of their exact age. Therefore, we opted for raw scores. The children participating in immersion education came in contact with Dutch at school for approximately eight months. See Table 1 for the demographics and scores of all participants.

Table 1. Demographics and scores.

	T1	T2
N	35	35
Age (months)	58.8 (3.4)	68.4 (3.4)
Gender [a]	F = 18, M = 17	F = 18, M = 17
Mother's education[b]	low = 2, high = 33	low = 2, high = 33
Father's education[b]	low = 10, high = 25	low = 10, high = 25
French	106.9 (13.3)	112.8 (15.6)
Dutch	/	33.8 (15.6)
IQ	5.1 (1.8)	6.8 (1.7)
flanker (RT)	1.53 (.30)	1.25 (.89)
flanker (ACC)	.67 (.17)	.87 (.13)
WM block 1 (RT)	1.21 (.26)	.98 (.17)
WM block 1 (ACC)	.82 (.16)	.90 (.11)
WM block 2 (RT)	1.26 (.25)	1.07 (.15)
WM block 2 (ACC)	.81 (.11)	.87 (.09)
WM block 3 (RT)	1.29 (.27)	1.09 (.19)
WM block 3 (ACC)	.79 (.15)	.90 (.10)
WM block 4 (RT)	1.28 (.39)	1.13 (.15)
WM block 4 (ACC)	.78 (.13)	.84 (.11)
DCCS (RT)	3.62 (1.98)	2.20 (.48)
DCCS (ACC)	.75 (.13)	.70 (.23)

Note. All numbers are the mean, with the standard deviation between brackets. [a] F = female, M = male. [b] Parents could indicate four options, which were afterwards coded as followed: low = primary school or secondary school as highest diploma, high = university or university college as highest diploma. RT = reaction times, ACC = accuracy scores.

Progress on CC, WM, IQ and French

To make sure the group was as homogenous as possible, the following analysis were computed on the children who were enrolled in immersion education and did not came in contact with a second language before the start of the immersion program. Mean RT and accuracy rates were calculated for each participant. Outliers were not included in the analysis.

Flanker task. Using repeated measures, we analysed the progress of CC and WM. The flanker task was analysed with a 2 x 2 design (Time point x Congruency), where Time Point is T1 and T2 and Congruency are the reaction times on congruent and incongruent trials. A significant main effect was found for Time Point ($F(1,34) = 28.468$, $p < .001$) and for Congruency ($F(1,34) = 30.607$, $p < .001$). Children were faster at T2 ($M = 1.25$, $SD = .29$) than at T1 ($M = 1.53$, $SD = .30$) and they were faster on congruent trials ($M = 1.32$, $SD = .25$) compared to incongruent trials ($M = 1.47$, $SD = .28$). The interaction effect between Time Point and Congruency was marginally significant ($F(1,34) = 4.020$, $p =.053$), thus the congruency effect was smaller at T2 ($M = 0.12$, $SD = .17$) than at T1 ($M = .07$, $SD = .09$). The same 2 x 2 design was used to analyse accuracy: (T1 and T2) x (ACC Congruent and ACC Incongruent). This resulted in a main effect for Time Point ($F(1,34) = 48.919$, $p < .001$) and for Congruency ($F(1,34) = 24.535$, $p < .001$). The participant was more accurate at T2 ($M = .87$, $SD = .13$) than at T1 ($M = .67$, $SD = .13$), and at congruent trials ($M = .83$, $SD = .11$) than at incongruent trials ($M = .71$, $SD = .17$). The interaction between Time Point and Congruency was also significant ($F(1,34) = 10.458$, $p = .003$). At T2, the participants made fewer errors on congruent and incongruent trials than at T1 (T2: $M = .86$, $SD = .28$; T1: $M = .69$, $SD = .24$).

Dimensional Change Card Sorting task. Regarding DCCS, two paired t-tests were used to look into differences in reaction times and accuracy. Accuracy did not differ significantly between T1 and T2 (T1: $M = .75$, $SD = .13$, T2: $M = .70$, $SD = .23$; $t(34) = 1.048$, $p = .302$), whereas the reaction times were significantly faster at T2 (T1: $M = 3.62$, $SD = 1.98$, T2: $M = 2.19$, $SD = .48$; $t(34) = 4.483$, $p < .001$). Participants were faster at T2, and were as accurate at T2 compared to T1.

Working Memory task. WM reaction times were analysed using two repeated measures, like Morales et al. (2013). The first was a 2 x 2 design, (Time Point x WM Load) which analysed the first and second time point, and the reaction times on the low (Block 1) and high (Block 3) working memory load. This resulted in a main effect of Time Point ($F(1,34) = 23.557$, $p < .001$) and of WM Load ($F(1,34) = 17,126$, $p < .001$). The participants were faster at T2 ($M = 1.07$, $SD = .13$) than at T1 ($M = 1.26$, $SD = .24$), and they were faster at low WM load ($M = 1.09$, $SD = .17$) than at high WM load ($M = 1.19$, $SD = .18$). The interaction between Time Point and WM Load was not significant ($F(1,34) = .308$, $p = .583$), thus the participants were not faster at T2 for the high or low WM load than at T1. The second design, a 2 x 2 x 2 (Time Point x WM Load x Congruency), tested T1 and T2, a low and high WM load, and congruent and incongruent trials in Block 2 and Block 4. There was a significant main effect for Time Point ($F(1,34) = 14.627$, $p = .001$) and for Congruency ($F(1,34) = 10.360$, $p = .003$). The children were faster at T2 ($M = 1.10$, $SD = .13$) than at T1 ($M = 1.27$, $SD = .28$), and were faster for congruent trials ($M = 1.17$, $SD = .17$) than at incongruent trials ($M = 1.21$, $SD = .19$). No significant interaction effects were present, thus they were as fast at T2 for congruent and incongruent trials, and for high and low WM load.

Accuracy of the WM task was also analysed. The first 2 x 2 design (Time Point x WM Load), included T1 and T2, and a low and high working memory load. There was a significant main effect for Time Point ($F(1,34) = 12.409$, $p = .001$), thus participants were more accurate at T2 ($M = .90$, $SD = .09$) than at T1 ($M = .80$, $SD = .14$). No significant main effect of WM Load ($F(1,34) = 1.055$, $p = .312$) nor a significant interaction effect ($F(1,34) = 1.109$, $p = .300$) was found. Children performed as accurately on low WM blocks as on high WM blocks, and there was no difference in accuracy at T2 compared to T1 for low and high WM load. The second analysis, a 2 x 2 x 2 design (Time Point x WM Load x Congruency), looked into the first and second time point, high and low WM load and the congruent and incongruent trials. This resulted in a significant main effect for Time Point ($F(1,34) = 9.804$, $p = .004$), for WM Load ($F(1,34) = 4.644$, $p = .038$) and for Congruency ($F(1,34) = 8.322$, $p = .007$). Thus, children were more accurate at T2 ($M = .87$, $SD = .09$) than at T1 ($M = .80$, $SD = .10$), at congruent trials (*M* = 1.17, *SD* = .17) than at incongruent trials (*M* = 1.21, *SD* = .19), and at low WM load (*M* = .85, *SD* = .08) than at high WM load (*M* = .81, *SD* = .09). The interaction between WM Load and Congruency was also significant ($F(1,34) = 5.372$, $p = .027$). Children made more errors on incongruent trials when the WM load was high.

Raven Progressive Matrices. IQ was analysed using a paired t-test. IQ scores on T1 and T2 differed significantly ($t(34) = -5.388$, $p < .001$). Thus, IQ was higher at T2 (*M* = 6.8, *SD* = 1.7) than at T1 (*M* = 5.1, *SD* = 1.8).

Échelle de Vocabulaire en Images Peabody. Development on French language scores was measured with a paired t-test. Participants' scores differed marginally significant between T1 and T2 ($t(34) = -2.024$, $p = .051$). Scores were higher at T2 (*M* = 112.8, *SD* = 15.6) compared to T1 (*M* = 106.9, *SD* = 13.3).

Influence of IQ, CC, and WM on second language acquisition and first language progress

This analysis included all measurements of the 35 children following immersion education. Since the education level of the mother had only two measurements of low education, we opted to use only the education level of the father as indication of SES. These scores were grouped in low SES (primary school and secondary school) and high SES (university and university college). There were many variables, so we used a Principal Component Analysis (PCA) first to reduce the amount of variables. Then, we used linear regression modelling to investigate which variables predicted Dutch language acquisition and the French vocabulary progress.

A correlation matrix and a correlation scatter plot were computed to explore the data. All correlations were below .70 (see Table 2), indicating that a PCA can be used for all variables. The correlation scatter plot indicated that the accuracy scores of the WM task were computed towards the high end of their range (see Figure 8 and 9 for the scatter matrix of the T1 scores and progression scores, respectively), so we applied a logit transformation on these four variables. These transformed variables were used in the following PCA.

Table 2. Correlation matrix between all variables measures at T1.

	IQ	French	flanker (prop)	flanker ACC	WM BL 1 RT	WM BL 2 prop	WM BL 3 RT	WM BL 4 prop	WM BL 1 ACC	WM BL 2 ACC	WM BL 3 ACC	WM BL 4 ACC	DCCS RT	DCCS ACC
IQ	1													
French	.218	1												
flanker (prop)	.1	.132	1											
flanker ACC	-.225	.275	.079	1										
WM BL 1 RT	-.083	-.007	-.15	.099	1									
WM BL 2 prop	.232	-.059	.106	-.039	.141	1								
WM BL 3 RT	-.269	-.092	-.114	.004	.692**	.109	1							
WM BL 4 prop	-.044	.192	-.124	.251	.349*	-.101	.537**	1						
WM BL 1 ACC	.085	.168	.107	-.122	-.327	-.253	-.03	.128	1					
WM BL 2 ACC	-.119	.304	-.019	.288	.116	-.287	.151	.361	.286	1				
WM BL 3 ACC	.318	.353*	-.088	-.21	-.187	-.056	-.098	.085	.645**	.309	1			
WM BL 4 ACC	.262	.264	-.236	-.059	.103	.189	.033	.231	.08	.369	.625**	1		
DCCS RT	-.135	-.088	-.276	-.003	.335*	.319	.33	-.057	-.128	.189	-.191	.064	1	
DCCS ACC	.105	.011	-.242	.17	.125	-.1	-.127	.011	-.183	-.264	-.08	-.178	-.262	1

Note. prop = proportion of RT. * $p < .05$ ** $p < .01$

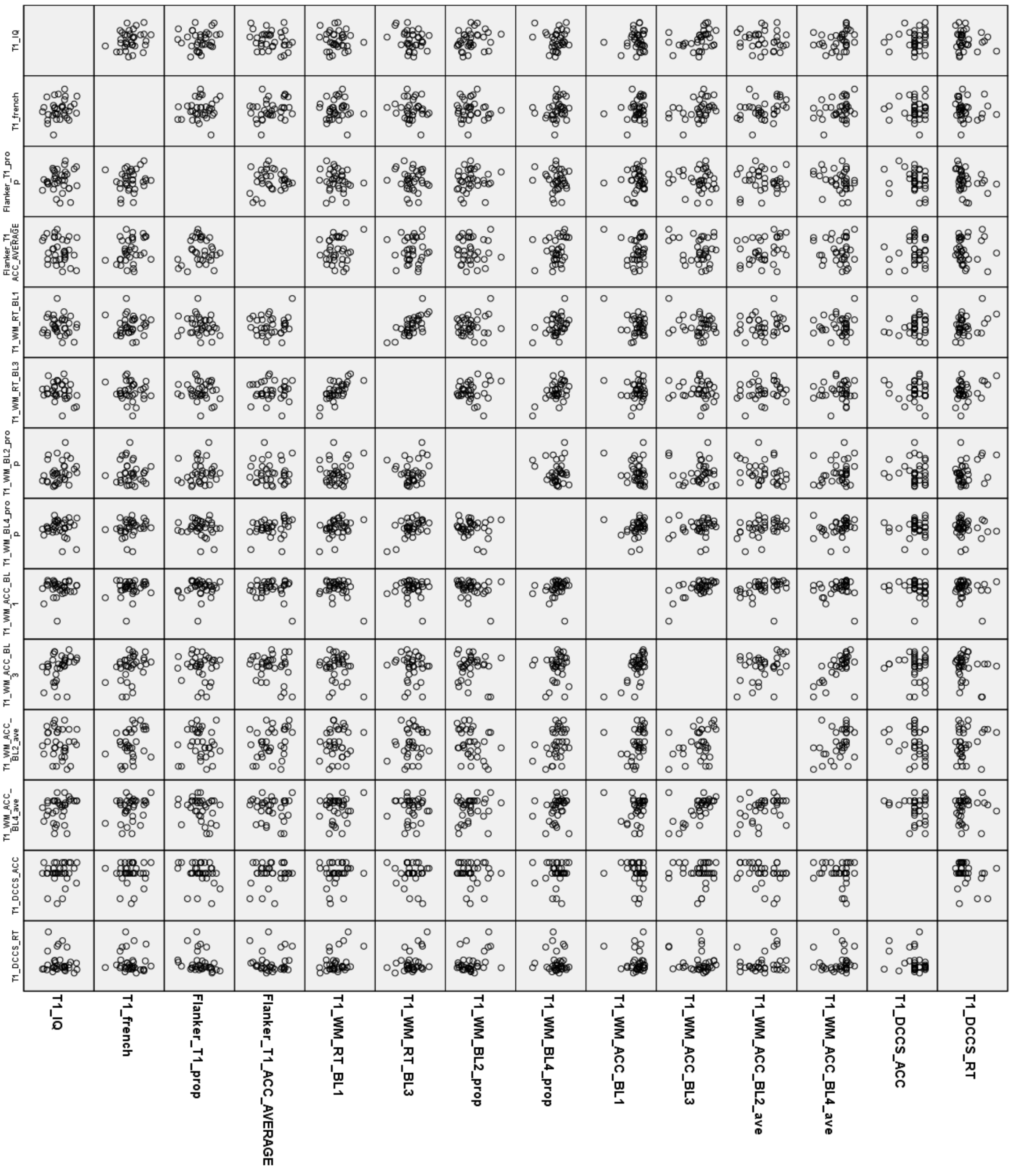


Figure 8. Scatter matrix of all variables at T1.

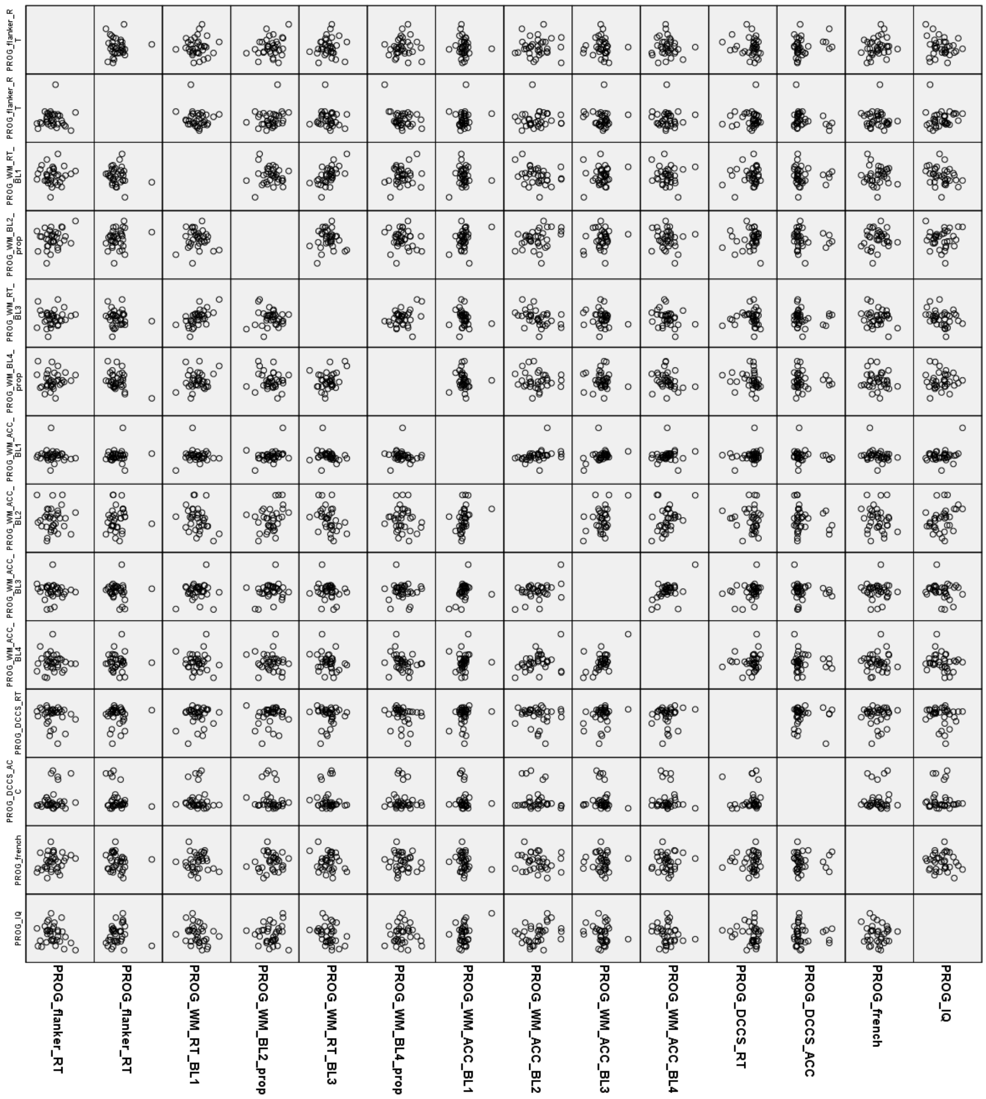


Figure 9. Scatter matrix of all progression variables.

PCA was used to reduce the amount of variables that explained the same variance in the data. Initially, we included all variables from T1 and the items that measured the progress the participant made on that test. However, this led to low validity within the components ($\alpha < .3$), so we opted to do a PCA twice, once for T1 variables and once for the progression variables.

The first PCA was conducted on 15 variables. Initially, we used an eigenvalue ≥1 in combination with a scree plot to determine the number of factors. We used direct oblique rotation (oblimin) because we expected that the underlying factors could be related. This allowed a maximum amount of non-orthogonality (Delta = 0). We used factor loadings that were higher than .4, resulting in 6 components that explained 80.34% of the variance. However, Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .45, which indicated a diffusion in the pattern of correlations and implied that PCA could be inappropriate. Kaiser (1974) advocates for a KMO measure of at least .5 to be acceptable. Thus, we used an anti-image correlation matrix to select adequate variables. An anti-image correlation matrix consists of measures of sampling adequacy on the diagonal, and the negatives of the partial correlation on the off-diagonal. Like the KMO measure, all diagonal elements should be greater than .5 to be an adequate sample. As Field (2009) suggested, we excluded the variables that had a low anti-image correlation.

A new PCA was done with the six resulting variables, all from the WM task, using oblique rotation (oblimin). Based on the eigenvalues ≥1 and the scree plot, the PCA resulted in 2 components that explained 68.49% of the variance. KMO measure had a sampling adequacy of .57, and Bartlett's test of sphericity was significant (χ^2 (15) = 61.035, $p < .001$) meaning that the correlations between the variables are significantly different from 0. The pattern matrix and structure matrix were used to interpret the variance of the components. Component 1, which explained 38.02% of the variance, was composed out of the RT of block 1, 3 and 4. Component 2, explaining 30.48% of the total variance, existed out of the ACC scores of block 2, 3 and 4. The internal consistency of each component was assessed with Cronbach's α, which was .77 for component 1 and .70 for the second component, which is around the needed consistency of .7 ("we're looking for values in the range of .7 to .8 (or thereabouts)" Field, 2009, p. 679). New regression variables were computed based on these two components.

The second PCA was also conducted on 15 variables, which were computed by combining the scores of T1 and T2. For RT, we used the following formula: RT_{T1}- RT_{T2}, where a negative score indicated participants were slower at T2. For the ACC scores, we calculated proportions by the following formula: ACC_{T1}/ACC_{T2}. The same method was used as with the previous PCA. First, all variables were entered and we used an eigenvalue ≥1 and a scree plot to assess the number of factors. Oblique rotation (oblimin) was used, since we expected that the variables were correlated. Delta = 0, so this allowed a maximum amount of non-orthogonality. Factor loadings >.4 were used. Six components explained 72.46 % of the total variance, and KMO measure was .52, which is moderate. Bartlett's test of sphericity was significant (χ^2 (15) = 76.21, p = .003) meaning that the correlations between the variables are significantly different from 0. However, the anti-image correlations showed that some variables did not have an adequate sampling adequacy (< .5), and were therefore removed (Field, 2009). This resulted in the inclusion of 11 variables, which loaded on 4 components explaining 70.55% of the total variance. KMO measure was .62, which was considered as mediocre. Bartlett's test of sphericity was significant (χ^2 (36) = 70.809, p < .001).

Since one component consisted of only two variables, which is less than the required three variables, and one component had a low internal consistency (α = .23), the PCA was computed again with 2 components, which explained 56.85% of the total variance. KMO measure was .61, which was considered as mediocre. Bartlett's test of sphericity was significant (χ^2(21) = 53.576, p < .001). The pattern matrix and structure matrix were used to interpret the variance of the components. The internal consistency of the components was assessed with Cronbach's α, and variables were removed if it was necessary to improve α. Component 1 consisted of all the ACC progression scores of the WM task and explained 34.89% of the total variance. Cronbach's α = .69. Component 2 included RT progress scores of block 1, 3 and 4 of the WM task. This component explained 21.96% of the total variance, and Cronbach's α was .65. New regression variables were computed based on these two components.

Linear regression modelling was used to assess the influence of CC, WM, IQ, SES and first language on second language acquisition. All variables were entered at the first stage (see table 2 for an overview), with Dutch as the outcome measure. Using the backwards stepping method, an initial model with all the variables was first computed.

The computer then calculated which variable was the least best predictor for the outcome measure, by calculating a single correlation with the outcome measure. This predictor was removed from the model, and then the computer calculated the next variable that was not a good predictor. These steps were repeated until the best fit was found. After each removal, an ANOVA was computed to assess whether the model was a significant fit for the data overall. Durbin-Watson measure, a measure for serial correlations between the errors, was 2.29. This test statistic should be between 1 and 3 (Field, 2009). See Table 3 for an overview of the variables per model, and the associated *p*-value. The predictors in model 6, 7 and 8 showed a marginally significant effect. Model 9 to 12 included all the predictors that led to a significant fit for predicting the outcome measure Dutch.

Table 3. Variables entered per model and associated *p*-values.

	Dutch			French		
Model	Variables entered	*F*	*p*	Variables entered	*F*	*p*
1	**WM**: PROG_RT, PROG_ACC, T1_RT T1_ACC **flanker:** T1_RT, T1_ACC, PROG_RT PROG_ACC **DCCS:** T1_RT, PROG_ACC T1_ACC PROG_RT **IQ:** PROG **French**: T1, PROG **SES**	$F(16,34) = 1.147$	.387	**WM**: PROG_RT T1_ACC T1_RT PROG_ACC **flanker:** T1_RT PROG_RT T1_ACC PROG_ACC **DCCS:** PROG_RT PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**	$F(14,34) = .689$	.760

	Dutch			French		
Model	Variables entered	*F*	*p*	Variables entered	*F*	*p*
2	**WM**: PROG_RT, PROG_ACC, T1_RT T1_ACC **flanker:** T1_RT, T1_ACC, PROG_RT PROG_ACC **DCCS:** T1_RT, PROG_ACC T1_ACC **IQ:** PROG **French**: T1, PROG **SES**	$F(15,34) = 1.290$	.296	**WM**: PROG_RT T1_ACC T1_RT **flanker:** T1_RT PROG_RT T1_ACC PROG_ACC **DCCS:** PROG_RT PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**	$F(13,34) = .779$	.673
3	**WM**: PROG_RT, PROG_ACC, T1_RT T1_ACC **flanker:** T1_RT, T1_ACC, PROG_RT PROG_ACC **DCCS:** T1_RT, PROG_ACC T1_ACC **IQ:** PROG **French**: T1, PROG	$F(14,34) = 1.447$	.219	**WM**: PROG_RT T1_ACC T1_RT **flanker:** T1_RT PROG_RT T1_ACC PROG_ACC **DCCS:** PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**	$F(12,34) = .883$	.575
4	**WM**: PROG_RT, PROG_ACC, T1_RT	$F(13,34) = 1.603$	.162	**WM**: PROG_RT T1_ACC T1_RT	$F(11,34) = 1.006$	.471

	Dutch			French		
Model	Variables entered	*F*	*p*	Variables entered	*F*	*p*
	flanker: T1_RT, T1_ACC, PROG_RT PROG_ACC **DCCS:** T1_RT, PROG_ACC T1_ACC **IQ:** PROG **French**: T1, PROG			**flanker:** T1_RT PROG_RT T1_ACC **DCCS:** PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**		
5	**WM**: PROG_RT, PROG_ACC, T1_RT **flanker:** T1_RT, T1_ACC, PROG_RT PROG_ACC **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: T1, PROG	$F(12,34) = 1.782$	.116	**WM**: T1_ACC T1_RT **flanker:** T1_RT PROG_RT T1_ACC **DCCS:** PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**	$F(10,34) = 1.148$	.370
6	**WM**: PROG_RT, PROG_ACC, T1_RT **flanker:** T1_RT, T1_ACC, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: T1, PROG	$F(11,34) = 1.96$	.084	**WM**: T1_ACC T1_RT **flanker:** PROG_RT T1_ACC **DCCS:** PROG_ACC T1_ACC T1_RT **IQ:** PROG **SES**	$F(9,34) = 1.316$	.278

	Dutch			French		
Model	Variables entered	*F*	*p*	Variables entered	*F*	*p*
7	**WM**: PROG_ACC, T1_RT **flanker:** T1_RT, T1_ACC, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: T1, PROG	*F*(10,34) = 2.088	.068	**WM**: T1_ACC T1_RT **flanker:** PROG_RT T1_ACC **DCCS:** PROG_ACC T1_ACC T1_RT **SES**	*F*(8,34) = 1.511	.201
8	**WM**: T1_RT **flanker:** T1_RT, T1_ACC, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: T1, PROG	*F*(9,34) = 2.276	.051	**WM**: T1_ACC T1_RT **flanker:** PROG_RT T1_ACC **DCCS:** PROG_ACC T1_RT **SES**	*F*(7,34) = 1.733	.143
9	**WM**: T1_RT **flanker:** T1_RT, T1_ACC, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: PROG	*F*(8,34) = 2.51	.036	**WM**: T1_RT **flanker:** T1_ACC PROG_RT **DCCS:** PROG_ACC T1_RT **SES**	*F*(6,34) = 2.028	.095
10	**WM**: T1_RT	*F*(7,34) = 2.833	.024	**WM**: T1_RT	*F*(5,34) = 2.442	.058

	Dutch			French		
Model	Variables entered	*F*	*p*	Variables entered	*F*	*p*
	flanker: T1_RT, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG **French**: PROG			**flanker:** T1_ACC PROG_RT **DCCS:** PROG_ACC **SES**		
11	**WM**: T1_RT **flanker:** T1_RT, PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG	$F(6,34) = 2.967$	.023	**WM**: T1_RT **flanker:** T1_ACC PROG_RT **SES**	$F(4,34) = 2.677$	.051
12	**WM**: T1_RT **flanker:** PROG_RT **DCCS:** T1_RT, PROG_ACC **IQ:** PROG	$F(5,34) = 3.292$	.018	**flanker:** T1_ACC PROG_RT **SES**	$F(3,34) = 3.108$	.041
13				**flanker:** PROG_RT **SES**	$F(2,34) = 3.539$	.041

Note. PROG = progression, RT = reaction times, ACC = accuracy, T1 = first time point

The same method and variables as the previous analysis were used to assess the influence of IQ, CC, and WM on the rate of learning of the mother tongue of the participants. Progress French, which was the subtraction between the standardised scores at T2 and T1, was now used as outcome measure. Since the score at T1 is already present in the outcome measure, it was now removed as predictor. The backward stepping method led to composing 13 models, where one variable at the time was removed. Durbin-Watson measure was 1.70. Model 9 to 11 were marginally significant (p = .095, p = .058 and p = .051, respectively). Model 12 and 13 were significant (p = .041 for both). The significant predictors in model 13 were SES and the accuracy in the flanker task at T1. Model 12 also included the progression rate on the RTs in the flanker task.

Discussion

The effects of bilingualism have been up for debate for decades. In the past, conclusions based on research have varied between disadvantageous effects on IQ and cognition (e.g. Arsenian, 1937), advantageous effects (e.g. Morales et al., 2013) and null effects (e.g. Paap & Greenberg, 2012). Thus, the link between cognition and bilingualism is not clear at the moment. However, no studies measured cognitive control (CC) and working memory (WM) before the acquisition of a second language (e.g. Nicolay & Poncelet, 2013b), so research has not taken the baseline into account. Therefore, it is not known how CC and WM develop and influence language acquisition. The only indication that CC and WM could be very important for learning a new language, is the study by Kapa and Colombo (2014). They found that children who scored high on CC, learned an artificial language better. However, since an artificial language cannot be compared to a real language, the possible advantageous influence of CC and WM on second language acquisition needed further investigation. As indicated by Woumans and Duyck (2015), the only possibility to draw reliable conclusions is to conduct a longitudinal study where bilingualism becomes a variable over time.

Therefore, in the present study, we set up a longitudinal design which followed the development of CC, WM, and IQ of children during second language acquisition.

CC was measured with a flanker task and the DCCS, WM with a task varying in WM load and IQ with the Raven Progressive Matrices. Children enrolled in immersion education (i.e. an education program where non-language courses are taught in a language which is not the mother tongue) were tested twice; one test moment before the start of the immersion program and one test moment after having contact with a second language for seven to eight months on language, CC, WM and IQ.

Progress on CC, WM, IQ, and French

The progress of the participants on these measurements was assessed. As expected, children were faster and made fewer mistakes on the flanker task at T2. The congruency effect, which is the difference between RT on incongruent and congruent trials, was significantly smaller. The amount of errors on congruent and incongruent trials was also diminished at T2. Thus, the participants were better at inhibiting responses after one year, which can be attributed to the normal development of CC (Best & Miller, 2010). The second CC task, DCCS, also resulted in a significant effect. Participants reacted faster, which could be interpreted as being able to implement a rule faster after at the age of 6 compared to the age of 5. However, the lack of significant effect regarding the accuracy scores at T1 and T2 implies that, even though the participant could apply a rule faster, they made the same amount of errors. Research has shown that three-year-old children can apply the first rule correctly, but perseverate during the next block. Thus, they have difficulty switching. At the age of five, children are able to switch in the next block (Zelazo, 2006). We used a variant of the DCCS that can also be used with older children, between 7 and 10 (Karinski, 2015; Zelazo, 2006), since we had to make sure no ceiling effect was reached at T1. Thus, it is possible that this task was too difficult to see any effect on the accuracy scores, while the RT do have an effect. Initially, we would have been able to compare these scores to the monolingual group to investigate if the bilinguals were more successful (e.g. Martin-Rhee & Bialystok, 2008) and how both groups developed, but due to the recruitment difficulties, such a comparison was not possible.

Regarding WM, results showed that participants were faster after one year, and were slower when the WM load was high at T1 and at T2. The non-significant interaction effect implied that time did not influence how fast they reacted on low- or high WM load. After one year, participants were also more accurate overall, more accurate on low WM load blocks than on high WM load blocks, and more accurate on congruent trials compared to incongruent trials. The interaction effect between WM load and congruency could be interpreted as participants making more mistakes on incongruent trials in the high WM load block.

Overall, the progress on these tasks was expected because the largest development of CC occurs between the age of 3 and 6 (Best & Miller, 2010). These effects were also found by Cragg (2016), who tested children of different ages. The largest difference was found between the two youngest age-groups.

IQ differed significantly between both time points, meaning that the children's IQ incremented from the 50th percentile (which is the mean score in the population) to the 68th percentile. The 68th percentile should be interpreted as there only being 32% of children of this age category that score higher or the same. Raven Progressive Matrices is an age normed test, so the increase in IQ is very likely real and not due to an artefact. Furthermore, this effect was also found by Woumans et al. (in press), where bilingual children scored higher on Raven Progressive Matrices after one year of immersion education compared to monolinguals who participated in a regular education program. At baseline, both groups were matched on SES, age, IQ and CC. When the children were retested approximately nine months later, the bilingual group scored significantly higher on IQ (T1 = 50.61, *SD* = 24.6; T2 = 76.04, *SD* = 20.6). The monolingual children did not perform better (T1 = 50.39, *SD* = 27.9; T2 = 58.54, *SD* = 30.2). Since the participants were matched at T1 for IQ, in contrast to other studies (e.g. Nicolay & Poncelet, 2015), they concluded there may be a bilingual advantage for non-verbal intelligence. We found the advantageous effect when the participants were enrolled for approximately 7 to 8 months, which is shorter than the timespan of Woumans et al. (in press). Thus, it is very likely the positive effect on IQ occurs at the primary stages of second language acquisition. Since the effect of Woumans et al. (in press) was higher than ours, it is possible that the IQ scores will be even higher when measured after more than 9 months.

Participants also scored significantly higher on French vocabulary after one year. The average percentile at T1 was 68, which increased to the 80th percentile at T2. The test scores were age-normed, therefore it should be expected that the children scored the same one year later. Peabody Picture Vocabulary test is a widely used test, and is tested extensively on validity and on reliability, as indicated by the manual. Therefore, it is likely that these scores reflect a real increase in French vocabulary, which cannot be attributed to age, and that the participants had a more extensive vocabulary after one year of contact with a second language. To investigate if this advantageous effect is only present for vocabulary, more research with different language tests (e.g. tests on fluency) and a comparison with monolinguals is needed.

Influence of IQ, CC and WM on second language acquisition and progress first language

The goal of this study, was to investigate if children's ability to learn language can be predicted by their cognitive skills and WM. As hypothesised, we found that five variables predicted the scores for Dutch vocabulary significantly; progress on IQ, progress on accuracy of the DCCS and progress on RT of the flanker task. The RT at T1 on the flanker task and the RT at T1 on the WM task were also significant predictors. Four out of five variables had a positive *Beta* coefficient (i.e. a standardised coefficient that indicates how much the outcome will change if a one standard deviation change occurs in the predictor). Thus, the higher the increase in IQ, the more language was acquired. Also, the more progress the participant made in accuracy, or the faster the child reacted at T2 compared to T1, or the faster it was at T1 on the flanker, the more Dutch it knew. Since the scores of the WM task were transformed and combined into variables, no conclusions can be drawn from this *Beta*.

These findings indicate that the relation between bilingualism, CC and WM is not as straightforward as previously assumed. On the one hand, previous research has found that bilingualism leads to better evolved WM (e.g. Morales et al., 2013) and CC (e.g. Barac & Bialystok, 2011), however these studies did not take the initial CC and WM performance (before the onset of bilingualism) into account. On the other hand, there was the study by Kapa and Colombo (2013), which was to our knowledge the first to reverse the design and investigate the influence of CC and WM on second language acquisition. However, the second language that needed to be acquired was an artificial language, which was composed of arbitrary words but was still based on English grammar, the native language of the participants. Nevertheless, the authors found that artificial language acquisition was predicted by the participants' scores on DCCS and Attention Network Test. Apart from the more complex language, our study also tested for IQ, which is an important variable to take into account. Additionally, the timespan of this study was ten months - eight months in daily contact with the second language - and should therefore yield more reliable effects than two one-and-a-half hour sessions spread over two days. Our findings indicate that all three components of CC (response inhibition, switching and updating WM) are predictors of language acquisition when the language is more complex.

The finding that the combination of these three components is important for language acquisition also supports the research by Miyake et al. (2000), in which they concluded that CC is composed of these three components. Zelazo (2015) also used these three components in his iterative reprocessing model, to explain development of CC in life.

The same method of analysis as with second language acquisition was used to investigate the effect on first language growth. In contrast to second language acquisition, CC and WM were no longer the main predictors in these models. SES and one measurement of CC, response inhibition (RT at T1 of the flanker), predicted the increase in vocabulary scores. SES had a positive relation with the outcome, thus the higher the education of the parents, the more vocabulary the child knew after seven to eight months of immersion education. The RT of the flanker task had a negative relation with the outcome. However, the use of a flanker task to measure the inhibitory control of an individual is limited (Salthouse, 2010). Interestingly, Kapa and Colombo also found a negative relation between inhibitory control and artificial language acquisition. More research with different tests is needed to further explore this effect, but we can conclude that inhibitory control has an effect on first and second language acquisition. It should be noted that SES was no significant predictor for second language acquisition, which could imply that IQ, CC and WM are mainly important in the first stages of language acquisition, while SES influences the later stages of language acquisition.

This study had some limitations. Firstly, the amount of participants should ideally be higher. However, due to fall-out, we lost approximately 30% of the participants. Secondly, since we tested very young children, we were limited to choosing an extensive test or using a battery of short tests. Four-year-old children had difficulty concentrating for more than 30 minutes, 45 minutes if breaks are given. Therefore, we could not test them as extensively as we would have liked. Ideally, we would have added a second vocabulary measurement for Dutch and French, a second WM task and another IQ test. Future research could focus on testing children of this age extensively on one component, such as a full intelligence test.

Conclusion

Belgium, as a trilingual country, offers an ideal context to become bilingual or trilingual. The standard education program includes courses on three to four different languages by the time the student is 18 years old. These languages are often Dutch, French and English, with German or Spanish as a fourth language. In Flanders, classes in French as a second language start at the age of 10. Immersion education, an education method that is very popular in Wallonia, teaches a second language in a natural setting starting in kindergarten. Not only do the children have more experience with a second language by the age of 10, compared to a standard education setting, our findings indicate that learning a second language at this age leads to a higher IQ and higher vocabulary scores in their native language. We also found that IQ, CC, and WM are important predictors in the first stages of second language acquisition. Future research could investigate if language acquisition could be supported by offering cognitive training programs. Even if the effects of the training or not lasting, it could help children with the first stages of language acquisition. Our findings also indicate that SES and CC were important factors in the later stages of language acquisition. Children who had a higher SES at home, scored higher on vocabulary of their native language. Since low SES is often correlated with less vocal interaction with the parents (e.g. reading bedtime stories; Hale, Berger, LeBourgeois, & Brooks-Gunn, 2009), future research could, for example, focus on an effective method to offer these children more vocal interaction, to support the development of their vocabulary.

This study therefore offers some practical implications for language learning. The SES at home needs to be taken into account when teaching children. Children with lower SES might need more support, e.g. more vocal interaction, from the educational staff to reach the same goals as a child with high SES. Therefore, we suggest adapting the teaching style to each individual, taking background information into account. Primary schools and kindergartens in Flanders already offer extra teaching periods for children coming from a low SES home ("Decreet Basisonderwijs", 2015, article 131 - 135). It might help the children's first (and second) language acquisition to focus on cognitive control as well as more contact (speaking, writing and listening) with the native language during these classes. Cognitive development is very important in the whole lifespan. Supporting these processes by offering challenging games or teaching a second language to children, might help them immediately and later in life. We cannot guarantee that the positive effect of bilingualism on IQ is long-lasting - more longitudinal research is needed - but this effect is very relevant for future educational programs.

References

Abdelgafar, G. M., & Moawad, R. A. (2015). Executive function differences between bilingual Arabic–english and monolingual Arabic children. *Journal of Psycholinguistic Research*, *44*(5), 651-667. doi: 10.1007/s10936-014-9309-3

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., & Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in psychology*, *5*(398), 1-12. doi: 10.1027/1618-3169/a000243

Arsenian, S. (1937). Bilingualism and mental development. *The Teachers College Record*, *39*(3), 242-243.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559. doi: 10.1126/science.1736359

Barac, R., & Bialystok, E. (2011). Cognitive development of bilingual children. *Language Teaching*, *44*(1), 36–54. doi: 10.1017/S0261444810000339

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development*, *81*(6), 1641-1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental psychology*, *24*(4), 560. doi: 10.1037/0012-1649.24.4.560

Bialystok, E. (1992). Selective attention in cognitive processing: The bilingual edge. *Advances in Psychology*, *83*, 501-513. doi: 0.1016/s0166-4115(08)61513-7

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, *21*, 169-181. doi: 10.1017/s0267190501000101

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. Psychology and Aging, 19(2), 290–303. doi:10.1037/0882-7974.19.2.290

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and cognition*, *12*(01), 3-11. doi: 10.1017/s1366728908003477

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children.

Journal of experimental child psychology, *128*, 105-119. doi: 10.1016/j.jecp.2014.06.007

Calvo, N., Ibáñez, A., & García, A. M. (2016). The impact of bilingualism on working memory: A null effect on the whole may not be so on the parts. *Frontiers in psychology*, *7*. doi: 10.3389/fpsyg.2016.00265

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*(2), 595-616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*(2), 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

Costa, A., & Sebastian-Galles, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, *15*(5), 336–345. doi: 10.1038/nrn3709

Cragg, L. (2016). The Development of Stimulus and Response Interference Control in Midchildhood. *Developmental Psychology, 52*(2), 242-252. doi: 10.1037/dev0000074

Crivello, C., Kuzyk, O., Rodrigues, M., Friend, M., Zesiger, P., & Poulin-Dubois, D. (2016). The effects of bilingual growth on toddlers' executive function. Journal of Experimental Child Psychology, 141, 121–132. doi:10.1016/j.jecp.2015.08.004

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. Trends in Cognitive Sciences, 11(3), 97–104. doi:10.1016/j.tics.2006.12.005

Darsie, M. L. (1926). Mental capacity of American-born Japanese children. *Comparative Psychology Monographs*.

Davies, M., & Hughes, A. G. (1927). An Investigation into the Comparative Intelligence and Attainments of Jewish and non-Jewish school children. *British Journal of Psychology. General Section*, *18*(2), 134-146. doi:10.1111/j.2044-8295.1927.tb00457.x

Decreet Basisonderwijs (2015, november 23), retrieved May 15, 2016 from http:// data-onderwijs.vlaanderen.be/edulex/document.aspx?docid=12254

Decroly, O. (1926). IV. Essai d'application du test de Ballard dans les écoles belges. *L'année psychologique*, *27*(1), 57-93. doi: 10.3406/psy.1926.6313

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual-monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*(6), 1189–1211. doi: 10.1017/ S0142716412000744

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In D.Stuss & R.Knight (Eds.), *Principles of frontal lobe functioning* (pp. 466–503). New York: Oxford University Press. doi: 10.1093/acprof:oso 9780195134971.003.0029

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited. *Experimental Psychology*. doi: 10.1027/1618-3169/a000243

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory*, *19*(5), 529-537. doi: 10.1080/09658211.2011.590504

Engel de Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism enriches the poor enhanced cognitive control in low-income minority children. *Psychological science*, *23*(11), 1364-1371. doi: 10.1177/0956797612443836

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. doi: 10.3758/bf03203267

Fan, J., Wu, Y., Fossella, J. A., & Posner, M. I. (2001). Assessing the heritability of attentional networks. *BMC neuroscience*, *2*(1), 1.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications: London. doi: 10.1002/bjs.7040

Fishbein, S. (1980). IQ and social class. *Intelligence, 4*, 51-63. doi: 10.1016/0160-2896(80)90006-9

Fitzgerald, J. (1993). Views on bilingualism in the United States: A selective historical review. *Bilingual Research Journal*, *17*(1-2), 35-56. doi: 10.1080/15235882.1993.10162647

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological bulletin*, *134*(1), 31. doi: 10.1037/0033-2909.134.1.31

Gathercole, S., & Baddeley, A. (1989). Evaluation of the Role of Phonological Stm in the Development of Vocabulary in Children - a Longitudinal-Study. *Journal of Memory and Language*, *28*(2), 200–213. doi: 10.1016/0749-596X(89)90044-2

Grabo. R. P. (1931). *A study of comparative vocabularies of junior high schoolpupils from English and Italian speaking homes. Bulletin No. 13.* Washington, D.C.: U. S. Office of Education.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81. doi: 10.1017/s1366728998000133

Grosjean, F. (1982). Life with two languages: An introduction to bilingualism. Harvard University Press.

Hale, L., Berger, L. M., LeBourgeois, M. K., & Brooks-Gunn, J. (2009). Social and demographic predictors of preschoolers' bedtime routines. *Journal of developmental and behavioral pediatrics: JDBP*, *30*(5), 394. doi: 10.1097/dbp.0b013e3181ba0e64

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual? *International Journal of Bilingual Education and Bilingualism*, *15*(2), 131–145. doi: 10.1080/13670050.2011.652591

Hilchey, M., & Klein, R. (2012). Cognitive Control and Bilingualism. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale*, *66*(4), 312–312.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. doi: 10.1007/bf02291575

Kapa, L. L., & Colombo, J. (2014). Executive function predicts artificial language learning. *Journal of Memory and Language*, *76*, 237–252. doi: 10.1016/j.jml. 2014.07.004

Karasinski, C. (2015). Language ability, executive functioning and behaviour in school-age children. *International Journal of Language & Communication Disorders, 50*(2), 144–150. doi:10.1111/1460-6984.12104

Kaushanskaya, M., Gross, M., & Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science*, *17*(4), 564–583. doi: 10.1111/desc.12142

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560. doi: 10.1073/pnas.0811323106

Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, *325*(5940), 611-612. doi: 10.1126/science.1173947

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, *86*(1), 70–82. doi: 10.1016/S0093-934X(02)00535-7

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: language and cognition*, *11*(01), 81-93. doi: 10.1017/s1366728907003227

McCarthy. D. A . (1930). *The language development of the pre-school child.* Minneapolis: University of Minnesota Press.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. Developmental Psychology, 43, 947–959. doi: 10.1037/0012-1649.43.4.947

Mischel, W., Ebbesen, E. B., & Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of personality and social* *psychology, 21*(2), 204. doi: 10.1037/h0032198

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science, 244*(4907), 933-938. doi: 10.1126/science.2658056

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. doi: 10.1006/cogp.1999.0734

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. doi: 10.1016/j.jecp.2012.09.002

Nicolay, A.-C., & Poncelet, M. (2013a). Cognitive abilities underlying second-language vocabulary acquisition in an early second-language immersion education context: A longitudinal study. *Journal of Experimental Child Psychology*, *115*(4), 655–671. doi: 10.1016/j.jecp.2013.04.002

Nicolay, A.-C., & Poncelet, M. (2013b). Cognitive advantage in children enrolled in a second-language immersion elementary school program for three years. *Bilingualism-Language and Cognition*, *16*(3), 597–607. doi: 10.1017/S1366728912000375

Nicolay, A.-C. & Poncelet, M. (2015). Cognitive benefits in children enrolled in an early bilingual immersion school: A follow up study. *Bilingualism: Language and Cognition, 18* (4), 789-795. doi: 10.1017/s1366728914000868

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences?. *Journal of Cognitive Psychology*, *26*(6), 615-639. doi: 10.1080/20445911.2014.944914

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: general and applied*, *76*(27), 1. doi: 10.1037/h0093840

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *113*(4), 535–551. doi: 10.1016/j.jecp.2012.06.013

Raven, J. (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology*, *41*(1), 1–48. doi: 10.1006/cogp.1999.0735

Raven, J. C., Court, J. H., & Raven, J. (1998). *Progressive Coloured Matrices*. Oxford: Oxford Psychologists Press.

Saer. D. J. (1923). The effects of bilingualism on intelligence. *British Journal of Psychology, 14,* 25-38. doi: 10.1111/j.2044-8295.1923.tb00110.x

Salthouse, T. (2010). Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. Brain and Cognition, 73, 51–61.

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I., & Bialystok, E. (2011). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex*, *48*(8), 991-996. doi:10.1016/j.cortex. 2011.04.009

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in Bilingualism and Second Language Learning. In J. F. Kroll & A. M. B. de (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). New York, NY, US: Oxford University Press.

Stark, W. A. (1940). The effect of bilingualism on general intelligence: an investigation carried out in certain Dublin primary schools. *British Journal of Educational Psychology*.

Stevens, M., Lammertyn, J., Verbruggen, F., & Vandierendonck, A. (2006). Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*, *38*(2), 280–286. doi: 10.3758/BF03192779

Struys, E., Mohades, G., Bosch, P., & van den Noort, M. (2015). Cognitive Control in Bilingual Children Disentangling the Effects of Second-Language Proficiency and Onset Age of Acquisition. *Swiss Journal of Psychology, 74*(2), 65-73. doi:

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does Bilingualism Change Native-Language Reading? Cognate Effects in a Sentence Context. *Psychological Science*, *20*(8), 923–927. doi: 10.1111/j.1467-9280.2009.02389.x

Woumans, E., & Duyck, W. (2015). The bilingual advantage debate: moving toward different methods for verifying its existence. *Cortex*, *73*, 356-357. doi: 10.1016/j.cortex.2015.07.012

Woumans, E., Santens, P., Sieben, A., Versijpt, J., Stevens, M., & Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and Cognition*, *18*(03), 568-574. doi: 10.1017/s136672891400087x

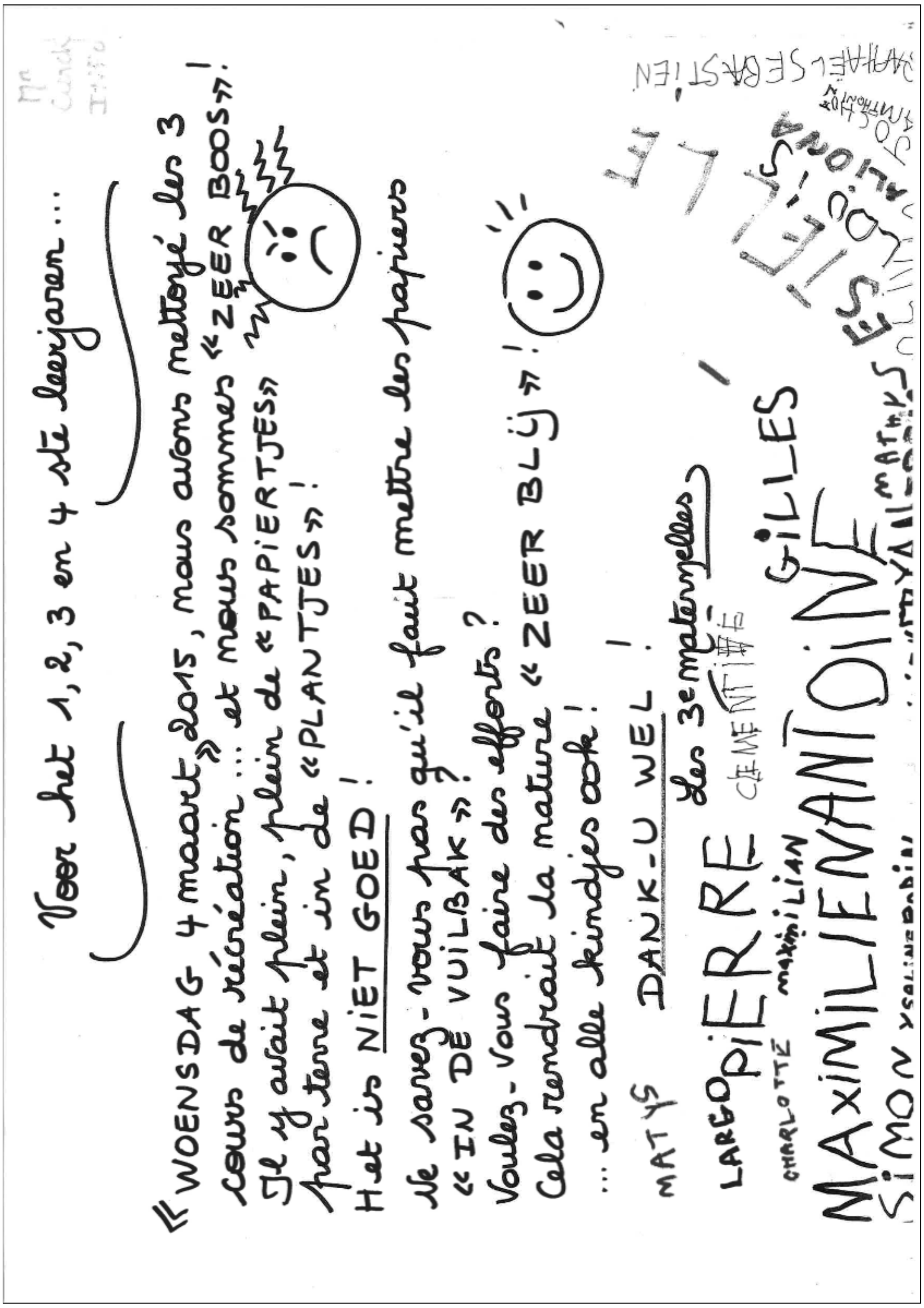
Woumans, E., Surmont, J., Struys, E., & Duyck, W. (in press). The longitudinal effect of bilingual immersion schooling on cognitive control and intelligence. *Language Learning.*

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review*, *38*, 55-68. doi: 10.1016/j.dr.2015.07.001

Appendix A

Example of a teaching method in an immersion school



Appendix B

Information letter for the immersion schools

UNIVERSITEIT GENT

Gand, le 17 Novembre 2014

Madame,
Monsieur,

Comme étudiante Psychologie Théorique et Expérimentale à l'université de Gand, je suis à la recherche d'une école qui voudrait m 'aider avec ma thèse de fin d'études.

Ils existent déjà plusieurs écoles qui offrent la possibilité de participer au programme d'immersion. L'avantage le plus important de ce programme est la compétence de parler un seconde langue. Selon beaucoup de chercheurs, cela n'est pas le seul avantage : des personnes bilingues sont meilleurs dans des fonctions de genre exécutive. Mais il n'est pas clair si ces avantages sont à l'origine ou bien les conséquences du bilinguisme; je voudrais ainsi trouver plus d'information sur ce processus. L'objet de ma thèse est « l'étude de l'influence de l'enseignement par immersion sur les compétences cognitives et le développement linguistique ». Trouver la réponse sur cette question n'est pas seulement important pour la recherche même, mais aussi pour les écoles. Plus de clarté sur les effets et les causes du bilinguisme permet d'offrir plus de soutien aux élèves et leur avenir.

Pour ma thèse, je voudrais examiner les enfants de la deuxième maternelle pendant le printemps de 2015, avant qu'ils entrent en contact avec une seconde langue. La prochaine prise d'essai serait l'année suivante au moment où les enfants auraient eu pendant presqu' une année déjà contact avec une deuxième langue, et ce pendant le printemps de 2016. Pour ma recherche, je voudrais examiner deux classes qui participent au programme d'immersion et deux classes qui suivent l'éducation normal. L'enquête prendra environ une demi-heure par personne, mais comme cela se ferait en groupe il sera possible d'examiner plusieurs personnes simultanément, donc la longueur totale sera minimisée autant que possible. Les tests consistent en un test QI bref, un test vocabulaire et quelques petits tests par ordinateur.
Si l'école et les parents seraient d'accord, les élèves recevront une surprise par après.
N'hésitez pas à me contacter pour davantage d'informations.

Je me réjouis de la possibilité d'une future collaboration.

Veuillez recevoir mes plus cordiales salutations,

Sofie Ameloot

Adresse courriel: sofie.ameloot@ugent.be
numéro de telephone: 0474/88.96.44

Appendix C

Information letter for the French schools

UNIVERSITEIT GENT

Gand, le 6 mars 2015

Madame,
Monsieur,

Comme étudiante Psychologie Théorique et Expérimentale à l'université de Gand, je suis à la recherche d'une école francophone qui voudrait m 'aider avec ma thèse de fin d'études.

Ils existent déjà plusieurs recherches qui sont sur le sujet d'acquisition de langue. Mais il n'est pas clair si il y a des facteurs qui aident à l'acquisition. Je voudrais ainsi trouver plus d'information sur ce processus. L'objet de ma thèse est « l'étude de l'influence de l'enseignement sur les compétences cognitives et le développement linguistique ». Trouver la réponse sur cette question n'est pas seulement important pour la recherche même, mais aussi pour les écoles. Plus de clarté sur les effets de acquisition permet d'offrir plus de soutien aux élèves et leur avenir.

Pour ma thèse, je voudrais examiner les enfants de la deuxième maternelle pendant le printemps de 2015. La prochaine prise d'essai serait l'année suivante, pendant le printemps de 2016. Pour ma recherche, je voudrais examiner des classes qui suivent l'éducation normal. L'enquête prendra environ 45 minutes par personne, mais comme cela se ferait en groupe il sera possible d'examiner plusieurs personnes simultanément, donc la longueur totale sera minimisée autant que possible. Les tests consistent en un test QI bref, un test vocabulaire et quelques petits tests par ordinateur.
Si l'école et les parents seraient d'accord, les élèves recevront une surprise par après.
N'hésitez pas à me contacter pour davantage d'informations.

Je me réjouis de la possibilité d'une future collaboration.

Veuillez recevoir mes plus cordiales salutations,

Sofie Ameloot

Adresse courriel: sofie.ameloot@ugent.be
numéro de telephone: 0474/88.96.44

Appendix E:

Information letter, informed consent and questionnaire for the parents.

UNIVERSITEIT GENT

SUJET: ETUDE CONCERNANT L'INFLUENCE DE L'ENSEIGNEMENT PAR IMMERSION SUR LES COMPETENCES COGNITIVES ET LE DEVELOPPEMENT LINGUISTIQUE

Gand, 07/12/2014

Chers parents,

L'objectif de notre étude est d'examiner le développement linguistique et cognitif des enfants. Pour cela, nous avons besoin de votre aide. Par la présente, nous vous prions de nous informer si vous autorisez votre enfant à participer à cette étude. Elle aura lieu à l'école de votre enfant et les tests prendront maximum une heure pendant les heures de cours.

<u>Qu'attendons-nous de l'école et des enfants?</u>
Cette étude est composée de deux périodes de test : la première en avril 2015/ mai 2015 et la seconde une année plus tard, en 2016. Les enfants devront accomplir deux exercices simples sur l'ordinateur, un test vocabulaire. Ceci se passera à l'école sous la supervision du chercheur principal de ce projet (Sofie Ameloot).

<u>Bon à savoir</u>

- La participation se fait sur base volontaire et vous pouvez décider à tout moment d'arrêter la collaboration.
- La participation est totalement anonyme. L'identité de votre enfant ne sera divulguée à aucun moment.
- La participation à cette étude ne comprend aucun risque.
- La participation à cette étude n'a aucune conséquence financière.

Si vous donnez votre accord, veuillez remplir le questionnaire et le formulaire de consentement ci-joints et faites-les parvenir à l'école.
Si vous avez des questions concernant la participation de votre enfant à cette étude, n'hésitez pas à nous contacter :

Sofie Ameloot (Chercheur principal)
Faculté Psychologie Expérimentale
Henri Dunantlaan 2, 9000 Gent
sofie.ameloot@ugent.be
GSM : 0474 88 96 44

Eva Van Assche (Promotor)
eva.vanassche@ugent.be
Evy Woumans (Co-promotor)
evy.woumans@ugent.be
GSM : 0474 54 83 46

Nous vous remercions d'avance pour votre collaboration à ce projet.

Cordialement

Sofie Ameloot	Eva Van Assche	Evy Woumans
Universiteit Gent	Universiteit Gent	Universiteit Gent

UNIVERSITEIT GENT

Formulaire de consentement

Je soussigné, ………………………………………… (nom) déclare que:

(1) j'autorise mon enfant à participer à cette étude

(2) je donne l'autorisation aux chercheurs de consulter les données de mon enfant et de les conserver de façon anonyme

(3) je sais que mon enfant peut interrompre sa collaboration à tout moment

(4) je sais que je peux m'informer de l'évolution de cette étude

(5) je sais que je peux consulter les résultats de cette étude et obtenir une synthèse

Lu et approuvé le …………………………. (date) à ………………………………...…….. (lieu),

Signature

Nom de l'enfant: ………………………………………………………………………………

Date de naissance de l'enfant: ……………………………………………………………………

UNIVERSITEIT GENT

Questionnaire pour les parents

Information sur l'enfant

Prénom: .... Nom: ....

Sexe: M / F Date de naissance: ....

Lieu de naissance: .... Nationalité: ....

Composition de la famille

Noms des parents: ....

....

Cohabitant / marié / divorcé / célibataire (encercler la bonne réponse)

---

1. Avez-vous l'impression que le développement de la langue de votre enfant est différent de celui d'autres enfants du même âge ? Oui / Non / Je ne sais pas

   Si oui, veuillez expliquer : ....

   ....

   ....

2. Avez-vous l'impression que la compréhension de la langue chez votre enfant est différente de celui d'autres enfants du même âge ? Oui / Non / Je ne sais pas

   Si oui, veuillez expliquer : ....

   ....

   ....

3. Votre enfant, a-t-il des problèmes de vue ? Oui / Non / Je ne sais pas

   Si oui, le(s)quel(s) : ....

   ....

   ....

4. Votre enfant, est-il déjà exposé à une autre langue que Français ? Oui / Non / Je ne sais pas

   Si oui, le(s)quel(s) : ....

   ....

   ....

UNIVERSITEIT GENT

Information supplémentaires au sujet des parents

	MERE	PERE
profession		
langue maternelle		
langue(s) parlée(s) avec l'enfant		
plus haut diplôme obtenu (encerclez la bonne réponse)	-Enseignement primaire -Enseignement secondaire inférieur -Enseignement secondaire supérieur -Enseignement supérieur/université	-Enseignement primaire -Enseignement secondaire inférieur -Enseignement secondaire supérieur -Enseignement supérieur/université