

Marlisa Hommel 3774945
English Language and Culture: Education and Communication
MA Thesis, Utrecht University
Supervisor: René Kager
Second reader: Sharon Unsworth

August 2012

**English in Dutch primary schools:
The sooner the better?**

Table of contents

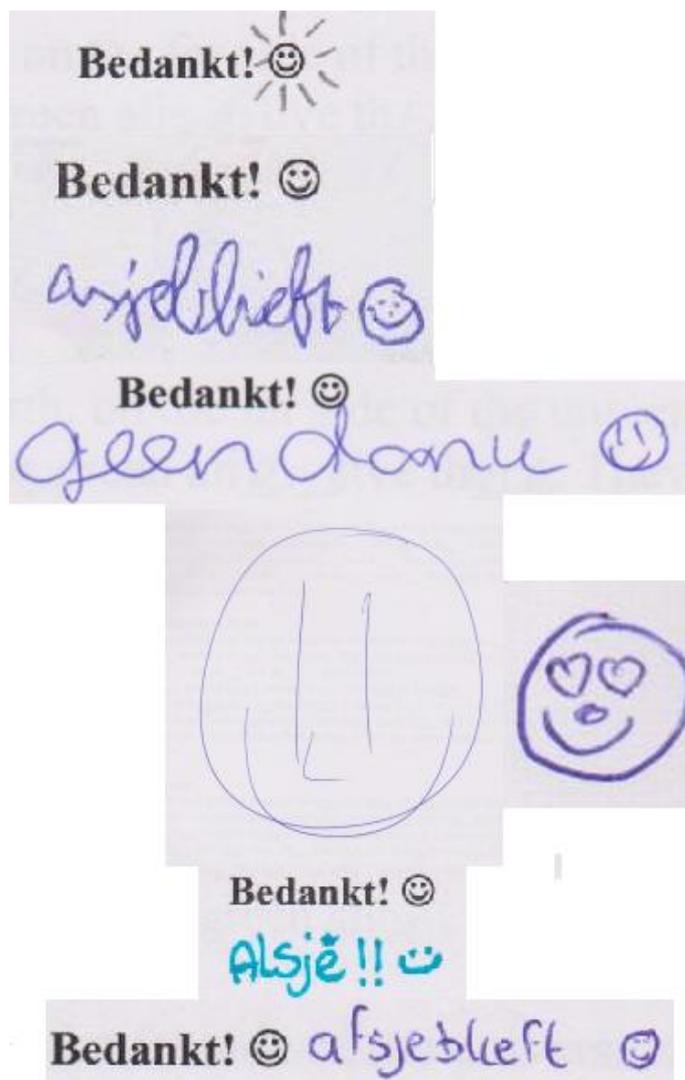
Preface	2
0 Abstract	3
1 Introduction: EFL in the Netherlands	4
2 Literature review: Factors affecting L2	8
2.0 Overview	8
2.1 Systematical, qualitative and quantitative input	8
2.1.1 Quantity	10
2.1.2 Age Of Learning (AOL)	11
2.2 Previous research on English in Dutch primary schools	16
2.3 Summary	20
3 Method	21
3.1 Participants	21
3.2 Testing materials: The C-test and Survey	21
3.3 Procedure	24
4 Results	26
5 Conclusion and Discussion	30
6 Appendices	33
Appendix 1: The questionnaires	33
Appendix 2: The C-Tests	35
Appendix 3: C-test by educated native speaker	39
Appendix 4: The Results	40
7 Bibliography	43

Preface

The topic of this thesis, English in Dutch (primary) schools, is a topic that I have been interested in for a while already. In the Netherlands, English is seen as an important foreign language, since it is a small country with international ambitions. Unlike Germany or France, where English TV programmes are usually dubbed, Dutch children can watch many English TV programmes with Dutch subtitles. This abundance of English on television shows the dominance of the language in this domain (e.g. Hollywood) and might also help to create a positive attitude towards the language.

Starting to write this thesis, I initially wanted to look at the effect of introducing English early as well as the effect of native input on English language proficiency, and a possible interaction effect of these factors. Unfortunately, this proved to be impossible in the limited time span of this thesis (around four months) and the factor of native input has been left out. Nevertheless, I still think that this thesis is a contribution to the academic world, because the results presented here contradict results from recent other MA theses.

I would like to thank my supervisor for his honest feedback, my mother for her optimism and Vadim Biryukov for being my listening ear. Last but not least, I would like to thank the participating schools for their openness and their pupils for participating in the experiment. I could not help but put some of their creativeness on this page. ☺



0 Abstract

This study looks at the influence of Age Of Learning (AOL) or amount of input (quantity) on Dutch pupils' proficiency of English. On the one hand, studies in informal settings have shown that starting earlier (AOL) as well as massive exposure to a language (quantity) cause higher language proficiency. On the other hand, studies in formal settings have shown that the amount of input is not enough for AOL or quantity to cause a significant effect on language proficiency. By comparing six primary schools (of which three start early with English instruction and of which three start late), this study wants to find out whether the amount of input that early starters receive at school is enough to reach a higher level of English proficiency than the late starters. This study is different from previous studies conducted on the same topic because the instrument used to test the subjects' language overall proficiency, the C-test, has a high validity and reliability. The results show that the early starters significantly outperform the late starters, which is in agreement with the current government policy, that promote an early start of instruction in foreign languages in primary schools.

1 Introduction: EFL in the Netherlands

Much research has been done in the field of Second Language Acquisition (SLA). The field has many aspects. For example, it has different disciplinary perspectives such as the linguistic, sociolinguistic, psychological and socio-psychological perspective (Saville-Troike, 2006). The field generally makes distinctions between formal and informal learning, and between a Second Language (L2) and a Foreign Language (FL). Formal language learning takes place in classrooms, whereas informal learning happens without instruction. A second language is a language spoken in the environment of learners that speak a different language at home, for example English in Zambia. The L2 is usually acquired in an informal way, but can also be taught via formal instruction. A foreign language is a language usually learned at school (formal instruction) and the language is not needed to fulfil immediate communicative needs, for example English in the Netherlands. This study, however, will not always distinguish these two as an FL is also a second language for a language learner. Therefore, an L2 can also refer to a learner's FL.

Why some learners of an L2 are more successful than others is also a big issue in the field of SLA. Language learning depends on many factors like Age Of Learning (AOL), Individual Differences (IDs) and quality and quantity of exposure. Concerning AOL, adults have an initial head start when learning an L2 because of cognitive advantages and better learning strategies. Examples of IDs are motivation, learning strategies, aptitude and attitude (Dornyei, 2005). Many factors go hand in hand. For example, it has been reported that starting early (= the factor AOL) positively influences a learner's motivation and attitude towards the L2 (Dornyei, 2005). The quality of exposure may be defined as the type of input of the language, i.e. whether the input is native(-like) or not. A native speaker as teacher sounds ideal but is controversial, because even though the input is of high quality, this might not automatically be in the best interest of the child. Since the regular (non-native) teacher also had to learn the same foreign language, this teacher has a higher empathic involvement (de Bot & Herder, 2005: 25). In addition, the regular teacher "shares the same background as the students" and "knows the cultural context (...) of the classroom (...) rather than just modelling it on the target community" (Seidlhofer, 1996: 70). It is also difficult to find enough native teachers as "the vast majority of EFL teachers (...) worldwide are non-native speakers of English" (Seidlhofer, 1996: 63). Moreover, due to the globalisation of English, "the ratio of non-natives to natives (is) steadily growing" (Árva & Medgyes, 2000: 356).

To avoid confusion, I will give a brief explanation of the Dutch education system. In literature, there is quite some divergence in the translation of the Dutch education system into English. The Dutch education system is slightly different from foreign countries. Dutch children start in 'groep 1' which is preschool in some countries, learn how to read and write in 'groep 3', which is grade 1 in other countries. Some studies literally translate 'groep 1' into 'group 1' or 'grade 1', which might cause some confusion. In this study, grade 1 will refer to Dutch 'groep 3' (6 or 7 year-olds). The Dutch education system generally distinguishes two types of English in primary schools. The most common type of English is called 'Engels in het Basisonderwijs' (EIBO) and generally starts in grade 5. The other -less common- type is called 'Vroeg VreemdeTalenOnderwijs' (VVTO) which generally starts in preschool (4 to 6 year-olds) and receives subsidies from the European Platform, an organisation that works together with the Dutch Ministry of Education and European Commission to stimulate internationalisation of Dutch education. An extensive survey conducted on behalf of the Dutch Ministry of Education with more than 500 respondents showed that a majority of schools (66%) start with English in grade 5 and 6 (regular EIBO). Around 15% start in preschool (VVTO), and the other 19% start somewhere between grade 1 and 4 (but usually not in grade 1 and 2 because the focus then is learning how to read and write) (Thijs et al, 2011). Regarding quantity, English is being taught around 45 minutes a week. These 45 minutes are not entirely in English, however; more than half of the teachers (56%) mainly use Dutch during English lessons (Thijs et al, 2011). Regarding quality, the survey showed that at a vast majority of schools (86%) the English lessons are taught by the regular teacher. Some schools work together with a specialist teacher, and only 1% of the schools has a specialised or (near)native speaker that prepares all English classes (Thijs et al, 2011). A case study conducted by Van Holten (2011) shows that there is a noticeable difference (regarding input, interaction and feedback) between the regular teachers and the specialist teacher. The regular teachers have a CEFR level between A2 and B1, whereas the specialist teacher has a CEFR level between B2 and C1 (Holten, 2011: 28). Most teachers (78%) estimate their own English at B1 CEFR-level or higher, but some (22%) admit that they do not reach this level. The B1 level is the minimal level that teachers need to possess in order to properly teach English (Thijs et al, 2011).

Ever since 1986, English has been a compulsory subject in Dutch primary education. The purpose of English in primary school is its focus on communication and meaning instead of its focus on form, i.e. it is more important for the pupils to get the message across than to be grammatically correct. The important areas of English in primary school are speaking and

listening (=oral skills) and vocabulary. Even though officially compulsory, English is still not taught in every primary school. Personally I have never received any English education when attending primary school, and 3 out of the 13 (23%) schools in Amsterdam that I contacted this year openly admitted that they do not teach English at all. If this is already the case in the capital city of the Netherlands, one might assume that this number will be even higher in less international cities and towns. This failure to meet the requirements is tolerated in the sense that the Dutch school inspection does not actively check whether English is actually being taught, and there are no sanctions against schools that do not teach English (de Bot & Herder, 2005). Another requirement or guideline that is not met is the amount of hours that should be spent on teaching English. The guideline is to teach 80-100 hours of English in grade 5 and 6 (10 to 12 year-olds) (de Bot & Herder, 2005: 8). Unfortunately, barely any school reaches this recommendation and the actual average of English that is taught is half the required time (around 30-50 hours) (de Bot & Herder, 2005).

The initial aim of this study was to look at two of the many before-mentioned factors that influence L2 learning: AOL/quantity and quality of exposure. Unfortunately, due to the limited number of native speakers in Dutch primary schools (only 1%), looking at the effect of quality of input seemed practically unattainable. This is unfortunate because little is known about the interaction effect of quantity and quality. For example, maybe starting early alone does not result in a higher L2 proficiency, but starting early with a native speaker does. The research question of this study is: is the quantity of input that VVTO children receive at school enough to reach a higher level of English competence than the EIBO children, i.e. does AOL/quantity of exposure affect the level of English proficiency? Or do the EIBO children catch up in the two years they receive English because they are older (and therefore learn faster) and because older (VVTO and EIBO) learners have a higher exposure to English outside the classroom (such as games and TV)? Or, in other words, can EIBO children make up for the six years they missed out on receiving English language instruction? By testing subjects of the same age (in grade 6), cognitive advantages by older subjects are ruled out (i.e. older subjects have better test-taking strategies which might be a confounding factor), but unfortunately AOL and quantity of input cannot be clearly distinguished from one another. Therefore, if VVTO subject score higher, this might not be due to the factor age but to the factor of quantity (or a combination of these). The subjects were tested in grade 6 because this is the last year of primary school, and in secondary school these two different groups are usually put together in the same class which might be another confounding factor.

The next chapter, chapter two, provides a literature review of international studies on how different factors have an influence on language proficiency as well as a summary of Dutch studies on English in Dutch primary schools. Chapter three explains the method used for this experiment and chapter four provides the results of the experiment. The last chapter offers a conclusion and a discussion with suggestions for further research.

2 Literature review: Factors affecting L2

2.0 Overview

The first part of this chapter provides a literature review on factors that have an influence on language proficiency. It starts with the importance of systematical, qualitative and quantitative input. It then goes into more detail regarding quantitative input and Age Of Learning, since these are the two factors important in this study. The second part of this chapter focuses specifically on the situation in the Netherlands. It starts with concerns regarding the current situation of English instruction in Dutch primary schools and then some previous experiments are mentioned where the proficiency of subjects attending VVTO schools was compared to subjects attending EIBO schools.

2.1 Systematical, qualitative and quantitative input

Many researchers report that young L2 learners become more proficient in the target language than older learners (cf. Birdsong and Molis, 2001; DeKeyser, 2000; Flege et al, 1995; Johnson and Newport, 1989; Larson-Hall, 2008; Muñoz, 2008; Singleton & Ryan, 2004). Studies in informal settings have consistently shown that younger learners reach a higher ultimate attainment (Muñoz, 2008). Younger learners are reported to develop a higher metalinguistic awareness, achieve a better pronunciation, have a lower affective filter (i.e. they are more open towards learning a foreign language) and they clearly have more time to master a foreign language (de Bot & Herder, 2005: 23).

Regardless of the advantages, however, concerns have been raised about whether learning an L2 at an early age does not harm the first language development of a child. Fortunately, there is ample evidence that shows that learning a foreign language at an early age does not affect the first language development. For example, a study with 88 Dutch preschool children showed that English VVTO does not have a negative effect on Dutch language development (Goorhuis-Brouwer & de Bot, 2005). A recent outcome of the 'Foreign Languages in Primary school Project' (FLiPP), an ongoing project on VVTO English in Dutch primary schools, showed that English instruction to Dutch toddlers does not (harmfully) affect the language development of Dutch (Europees Platform, 2012).

Do these positive findings imply that it is better to start teaching foreign languages as early as possible? It is important to conduct more formal classroom research because it is

these findings that “are most relevant to decisions concerning the time and timing of second language instruction” (Muñoz, 2008: 591). Research has generally shown that starting early is better, but only when it is taught *systematically* with high *quality* (native-like input) and *quantity* (immersion) (Muñoz, 2008). It is important for English to be taught *systematically*, because language attrition occurs easily with young children. If toddlers start with learning a foreign language, they should continue at least till the age of eleven or twelve (de Bot & Herder, 2005: 24). A well-known term in the field of SLA is Krashen’s Input Hypothesis. It entails that L2s are learned by receiving *systematic* input+1, which is comprehensible language input “a little beyond” the learner’s current level of competence (Krashen, 1981). Formal environments can provide comprehensible input, but if this type of input would be enough, formal language learners would be able to reach a native-like competence in the L2, which is generally not the case.

Language input should also be of high *quality*, and Native English Speaking Teachers (NESTs) are good sources of this type of input. For pronunciation purposes, for example, it might be better to have a native speaker as a teacher. Larson-Hall (2008) looked at early starters that reported to have NESTs and compared them to early starters with non-NESTs. The early starters with NESTs did not score significantly higher on a phonemic discrimination task, and the researcher did not find any clear profits from having a NEST (Larson-Hall, 2008). This is probably due to the fact that even if native(-like) input is available, the amount of input will still not be high enough to trigger implicit language learning (Muñoz, 2008).

The amount of input, or *quantity*, is important “since length of exposure to an L2 is recognised as an important predictor of L2 success” (Singleton & Ryan, 2004: 223). In order for acquisition (as opposed to conscious learning) to take place, learners have to be exposed to a massive amount of input as well as input from a native speaker (Muñoz, 2008). In regular schools, unfortunately, “children are not provided with the massive amounts of input” required to trigger the implicit learning mechanisms (Muñoz, 2008: 586). There is an important distinction to be made between the situation in the Netherlands, where English is a Foreign language (EFL) compared to countries where English is a Second language (ESL) because in EFL countries there is little opportunity for the pupils to learn the language outside the language classroom and the main source of language learning is inside the classroom (Muñoz, 2008). This makes the purpose of English in Dutch primary schools questionable; do pupils learn much by receiving English input from a Dutch native speaker for less than an hour a week?

2.1.1 *Quantity*

Input is a necessary condition for both L1 and L2 learning (Saville-Troike, 2006). The language learner should be exposed to great amounts of input to acquire an L2, and this goes for learners of all ages. Contrary to what some may believe, younger learners do not possess a “‘magical’ ability to learn a second language quickly and natively in a situation of minimal input” (Larson-Hall, 2008: 58). Research has shown that immersion programmes or bilingual education, where the language learner receives a higher amount of input, cause a higher L2 achievement than regular school programmes (Genesee, 1998). This only goes for immersion programmes that use the target language at least 50% of the time, because programmes with less than 50% seem to be ineffective (May & Hill, 2005). Thus, the more exposure, or quantity, the pupil gets, the higher the achievement in L2 becomes, but only if the exposure is beyond the 50% threshold. This 50% is not limited to the language classroom, but encompasses all school subjects. Studies that showed the benefit of bilingual education are for example Canadian studies, where English pupils who followed the regular school programme were compared with English pupils that followed a French immersion programme. The studies indicated that the latter proved to outperform the former in French language tests (Swain, 1974 and Barik & Swain, 1975). A recent study conducted in the Netherlands also yielded positive results for bilingual schools (Verspoor et al, 2010). Previous research in the Netherlands already showed that in grade 9, pupils that followed bilingual education scored significantly higher on an English test than pupils from regular schools (Verspoor et al, 2010). The recent longitudinal study looked at five different schools with 500 pupils that did or did not follow the bilingual programme. The subjects were followed from grade 7 to grade 9 and were tested for vocabulary knowledge by the use of an EFL test and were tested for writing skills by the use of a writing assignment (a short story of around 150 words). English was used more than 95% of the instruction time, which exceeds the 50% threshold. The results were consistent; pupils participating in the bilingual programme performed better in both the vocabulary and writing test over the period of three years (Verspoor et al, 2010). In the Netherlands, bilingual education exists since 1989 and has been growing in popularity ever since. By now, there are over 100 schools that offer English-Dutch bilingual programmes (Europees Platform, 2012).

2.1.2 Age Of Learning (AOL)

AOL and quantity of input are correlated factors, in the sense that age only plays a significant role in second language acquisition when the learner is provided with enough input, which is the case in informal settings or immersion programmes where the target language is spoken as much as possible ('classroom language') (Larson-Hall, 2008). Studies in informal settings ('migrant studies') have shown that age significantly affects rate of learning (i.e. younger learners are initially slower) and ultimate attainment (i.e. younger learners reach a higher a proficiency) (Muñoz, 2008). The influence of the factor age on ultimate attainment has been regarded as proof for the existence of a Critical or Sensitive Period in second language learning. It has been claimed that after a certain age, learners will never be able to reach a native-like proficiency in an L2 due to maturational constraints. However, if there is indeed such a thing as a Critical Period (CP) for a second language, it is unclear when (at what age) the Critical Period would take place. It is claimed that the CP would be around puberty (Saville-Troike, 2006).

After residing in the L2 speaking country for about five years, younger learners start to consistently outperform older learners (Singleton & Ryan, 2004). The factor age also affects certain aspects of the language faculty differently, "phonology being affected at a much younger age than basic morphosyntax" (Larson-Hall, 2008: 40). For example, Flege et al (1995) conducted an experiment where the pronunciation of 240 native speakers of Italian with different AOL was examined. The subjects resided in Canada for around 32 years. The subjects were recorded and a group of native speakers of English assessed their production of stops and fricatives. The subjects were divided into 10 groups, ranging from AOL 3 to AOL 21. The subject's AOL and the assessed native-like score was consistent; the lower the AOL the higher native-like score (Flege et al, 1995). Another example of a migrant study is one of DeKeyser (2000), who performed a grammaticality judgment task on 57 native speakers of Hungarian who migrated to the US. The migrants were divided in a group of Early Arrivals (before the age of 17) and Late Arrivals (17 years of age or over). The subjects resided in the US for around 34 years. Results showed a high negative correlation between test score and age of arrival ($r = -0.63$, $p < 0.001$) (DeKeyser, 2000: 510). Results also showed that "very few adult immigrants scored within the range of child arrivals on a grammaticality judgment test" (DeKeyser, 2000: 499). Studies conducted by Johnson and Newport (1989) (46 Korean and Chinese migrants in the US) and Birdsong and Molis (2001) showed similar results using a similar grammaticality judgment task. Birdsong and Molis (2001) performed an experiment

with 61 native speakers of Spanish who migrated to the US. The migrants were divided in a group of Early Arrivals (before the age of 17) and Late Arrivals (17 years of age or over). The subjects resided in the US for around 10 years. It is unclear whether the Early Arrivals scored significantly higher than the Late Arrivals, but no correlation was found between test score and Early Arrivals, whereas a strong negative correlation was found between Late Arrivals and test score ($r = -0.69$, $p < 0.0001$) (Birdsong and Molis, 2001: 240). It is interesting that the researchers mention that this is some evidence against the existence of a Critical Period in second language learning, since Late Arrivals should not show a decline in test score when age of learning increases (because the brain of a 17 year-old should be similar to, for example, that of a 27 year-old) (Birdsong and Molis, 2001).

However, the formal language context is different from the informal environment and research conducted in naturalistic contexts cannot be generalised to the formal language context. Similar to informal settings, formal settings have shown that older learners are faster learners (Muñoz, 2008). But contrary to the general belief, “no evidence has yet been found that the proficiency of younger *formal* learners eventually becomes higher than that of older *formal* learners after the same amount of instruction” (Muñoz, 2008: 581). Even in the area of phonology, younger starters have *not* proven to outperform late starters (Singleton & Ryan, 2004). Therefore, it is important to keep in mind that “an early start is not synonymous with better long-term performance” (Miralpeix, 2006: 91). In other words, starting to teach a language at an earlier age does not automatically mean that the younger learners will reach a higher level of proficiency. Even if younger learners perform better than older learners, this can be due to longer instruction time instead of an earlier start. Unfortunately, many classroom policies have accepted the idea that an earlier start will yield better long-term results in the foreign language on the assumption that “younger = better” (Singleton & Ryan, 2004: 198). On the other hand, even though there is no strong evidence that learners profit from starting L2 instruction earlier, other valid reasons for starting early might be the “advantage of starting some subjects early in the context of the modern crowded curriculum” or “the educational merits of early contact with another culture” (Singleton & Ryan, 2004: 224).

As mentioned before, previous research on AOL in informal settings shows no advantages of younger learners and “if there are advantages, they tend to disappear over time” (Cenoz, 2003: 78). Cenoz (2003) conducted a research in Spain where the general proficiency of English as a third language was tested. The subjects received a similar amount of instruction (600 hours) at different ages (4, 8 and 11). In all tests (e.g. vocabulary, grammar

and fluency) the later starters performed significantly better (Cenoz, 2003). Possible explanations for this are the cognitive maturation of the later starters or the type of input the later starters received (i.e. more complex) (Cenoz, 2003). Younger learners did however seem significantly more positive towards and motivated for learning English. Even though positive attitude and motivation is linked to language proficiency, this link does not show in this study. The link is probably too weak to have a big influence (Cenoz, 2003).

Other research on AOL was the longitudinal BAF (Barcelona Age Factor) project. The project looked at the age factor on English language acquisition as a third language and the data was collected over a time-span of six years. The study compared subjects that started learning English at different ages but that received the same amount of language instruction at three different times (after 200, 416 and 726 hours of instruction) (Cenoz, 2007: 628). The project looked at different areas, such as writing, oral fluency and vocabulary acquisition, and all the results indicate that late learners acquire a language faster and that “late starters outperform early starters” (Cenoz, 2007: 628). An example of a study that was part of the BAF project is the experiment conducted by Miralpeix (2006), who researched whether an early start with English in a Spanish/Catalan context has any lasting positive effects regarding productive vocabulary. The subjects received the same amount of formal instruction in English and they were both tested at the ages 16 or 17. The first group started at age 8, the other at age 11. Results showed that if significant differences between the groups were found (as in for example the cloze test and holding an interview), the late starters actually significantly outperformed the early starters. This indicates that -at least concerning productive vocabulary- late starters learn a language more efficiently than early starters (Miralpeix, 2006). The downside of the BAF project is that it did not take into consideration the cognitive advantages and better “test-taking strategies” that older learners have over younger subjects (Cenoz, 2007: 628). Thus, late starters might not outperform early starters when they are of similar age. However, this project does provide evidence that it is not necessarily better to teach a foreign language at an earlier age.

The problem with Cenoz’s experiment (2003) and the BAF project is that it looked at English as a *third* language. An experiment that looked at EFL as a *second* language is the older study of Burstall et al (1974). This is a longitudinal study that covers a time-span of ten years (1964-1974). The researchers looked into French taught in primary schools in Britain and Wales. French as a primary school subject was just introduced in 1964, and the British Ministry of Education wanted to see the possible benefits or drawbacks of introducing this foreign language to younger pupils. The study mainly looked at the attitudes of earlier starters

towards the language and consists of many experiments. The subjects were selected on the basis of age, i.e. they had to be 8 to 9 years olds. In 1964, 1965 and 1968 around 6,000 subjects that met the age requirement participated in the study. Besides the subjects that received French education in primary school, there were also control groups that started French in secondary school. The study initially wanted to advise the Ministry of Education on whether it is feasible and advisable to teach French in primary schools, but the study turned out to have more of an exploratory nature due to practical issues like staffing problems (not enough trained teachers of French). The answer to the research question “is any substantial gain in mastery achieved by beginning to learn French at the age of eight” is: “unequivocally negative” (Burstall et al, 1974: 243). No gains for earlier starters were found. For example, a control group that was two years older than the subjects at the time of testing (15 vs. 13 year-olds) but who studied French for an equivalent period of time, scored significantly higher on a test battery (including listening, reading, writing and speaking) (Burstall et al, 1974). Another example is an experiment where 16 year-olds were tested and those who started learning French at the age of eleven scored similar to the earlier starters. The younger starters did show more positive attitudes towards the French language, but this does not reflect in a “higher level of achievement” in French (Burstall et al, 1974: 244).

Another older study that looked at EFL as a second language is that of Oller and Nagato (1974) who looked at the long-term effect of early English at a Japanese school. By means of a cloze test they compared students who learned English from grade 1 to pupils that started in grade 7. The subjects were tested at 3 moments. In grade 7 (12 or 13 year-olds) the early starters scored significantly better ($t(28.72) = 1.100$, $p < 0.001$), in grade 9 they still performed significantly better ($t(5.56) = 1.78$, $p < 0.025$), but in grade 11 no significant difference was found ($t(1.61) = 1.46$). The researchers concluded that teaching English at an earlier age does not have a lasting positive effect (Oller and Nagato, 1974). A confounding factor, however, is that the subjects were integrated in the same class from grade 8 on. Another possible confounding factor is that the subjects were all girls, so the experiment did not contain a random sample representing the whole Japanese population.

Similar to Oller and Nagato (1974), Larson-Hall (2008) also looked at the long-term effect of early English in Japan. As many as 200 Japanese students took part in the experiment. The early starters started learning English when they were around 9 years old, and the late starters started at age 12 or 13. The subjects were tested at the average age of 19.4 years, and conducted an oral grammaticality judgement test and a phonemic discrimination task. No significant difference between groups on the GJT test was found. There was however

a significant difference between the groups for the phonemic discrimination task in favour of the early starters. The effect size is nevertheless small because it explains only 3% of the variance in the scores. The results were similar when quantity of input was controlled for by means of a one-way ANCOVA. Thus, the study found modest evidence in favour of the early starters “even in a situation of minimal input” which is “contrary to predictions that age only plays a role in naturalistic or immersion environments” (Larson-Hall, 2008: 58). However, this study cannot distinguish between the factor age and quantity, so either AOL or amount of input (or a combination of these) has a significant effect on the score of the phonemic discrimination task. Despite the fact that the results cannot say that it is useful to start earlier because of AOL or because of amount of input, the results do touch upon “the practical question of whether (...) any advantage can be found for Japanese students studying language before it becomes mandatory” (Larson-Hall, 2008: 43). Despite the fact that the researcher found only a modest advantage for the early starters, Larson-Hall (2008) considers that even the smallest possible advantage for early starters is worthwhile and therefore recommends “to begin foreign language study as young as possible, with as many hours of input as are possible” (Larson-Hall, 2008: 59).

2.2 Previous research on English in Dutch primary schools

In the introduction it was mentioned that this study translates the Dutch ‘groep 3’ (6 or 7 year-olds) as grade 1. It was also mentioned that the Dutch primary education system has two types of English instruction: ‘Engels in het Basisonderwijs’ (EIBO), which starts in grade 5 and is most common (66%), and ‘Vroeg VreemdeTalenOnderwijs’ (VVTO) which starts earlier, usually in preschool. An average of 45 minutes a week is spent on English instruction, usually (86% of the time) by the regular teacher (i.e. a non-native speaker of English). This section will focus on previous research that has been conducted on the topic English in Dutch primary schools. First, a study is mentioned that focuses on the lack of a clear policy on English instruction in Dutch primary schools. Then, two studies are mentioned that conducted experiments that are similar to this study, where Dutch VVTO pupils were compared to Dutch EIBO pupils to see whether there is any noticeable difference between these groups.

Waninge (2011) looked at the policy of Early English in Dutch primary schools. Waninge found that in the majority of cases the regular school teacher teaches English instead of a specialist teacher or native speaker, and that 2 out of 20 VVTO schools have native speakers of English that provide the English lessons, which is more than the previously mentioned 1% found by a larger-scale study. Another large-scale study conducted amongst 1,778 primary schools found that in over 90% of the cases English is taught by the regular teacher (Heesters et al, 2008). Waninge’s survey showed that English was being taught for around 69 minutes a week (Waninge, 2011). The large-scale study, however, concluded that most primary schools teach English for only 45 minutes a week (Heesters et al, 2008).

Because schools are free to “determine their own goals for English as a foreign language, and within the school, teachers are free to determine their own course of action”, there seems to be little within-school and between-school agreement about how to teach English (Waninge, 2011: 36). That there is a need for a sound policy for all schools becomes clear by an anecdote from Waninge; “teachers from the same school indicated different starting grades” and two teachers “reported that within the same school, different methods for teaching English were used that differed extremely” (Waninge, 2011: 36). Another problem is that there are no official end-terms regarding the level of English the children have to reach before they go to secondary school. Concerning vocabulary, pupils should be able to hold simple conversations about everyday subjects. Pupils are expected to be familiar with common words and phrases (Heesters et al, 2008). The CEFR level pupils are *expected*, but not *required*, to reach at the end of primary school is A2 (Heesters et al, 2008). Unfortunately,

many teaching methods in primary schools “do not provide a clear overview of (CEF) proficiency levels and how these relate to the material, with the exception of material from the Anglia Network” (Chevalking, 2010: 17). One result of a lack of a clear policy in primary schools is that many (primary and secondary) school teachers are unsatisfied with the way English is currently taught (de Bot & Herder, 2005). Primary school teachers also feel that their English skills are not good enough and that the teacher education does not prepare them well enough. The survey conducted by Thijs et al (2011) showed that many teachers (59%) indicated that they have not received any English instruction during their teacher education. It should be mentioned that nowadays English is an obligatory component in the teacher education, so this number will probably decrease. The main problem that should be addressed here is that teachers that *did* receive English are not satisfied with the way it is taught (Thijs et al, 2011). This is why teachers do not want to implement VVTO before the quality of the subject is improved in EIBO first (de Bot & Herder, 2005). This issue might be solved if more attention is paid to English in the teacher education. Another result of a lack of a clear policy is that secondary schools cannot build on what is learned in primary schools, which is why English in secondary schools starts at a very basic level all over again. This issue might be solved if English were to become part of the national exam at the end of primary school (the “Citotoets”). At the moment, this exam tests pupils’ knowledge of Dutch, maths en learning skills. In fact, the Citotoets already has an existing test for English called “Me2! Engels” based mainly on A1 and some A2 CEFR level exercises (Cito, 2012).

Unfortunately, no research -at least not to the best of my knowledge- has been done on the effect of native input versus non-native input in Dutch primary schools. Considering the small amount of native teachers (somewhere between 1-10%), this might not be surprising. Fortunately, previous research on the effect of AOL or quantity in Dutch primary schools has been done. In her thesis, Gidding (2011) compares VVTO pupils to EIBO pupils to be able to say what, if any, benefits the VVTO pupils have from their longer exposure to English in class. Gidding tested VVTO and EIBO children in grade 6 (on average 12-year-olds). By testing pupils of the same age, the researcher controlled the factor age, because, as mentioned before, older children have cognitive advantages in making tests. Gidding tested the pupils’ English proficiency by means of two tests: a vocabulary test called the *Peabody Picture Vocabulary Test 4* (PPVT-4) and a grammar test called *Test for Reception of Grammar 2* (TROG). These tests have problems with validity and reliability, however. First, both tests were developed for native speakers of English, not for language learners. With the PPVT-4 test, the subject gets to see four pictures and hears a word. The subject then has to choose the

right picture corresponding to the word (Gidding, 2011: 25). The problem with this is that Dutch and English are both Germanic languages and have many cognates. Cognates are words that look similar and have similar meanings (e.g. full-‘vol’, ship-‘schip’, flag-‘vlag’, plant-‘plant’, ring-‘ring’, fruit-‘fruit’, etc.). Language typology matters; overlap in vocabulary has a facilitating effect “because school-age children can make effective use of cognates in their acquisition of L2 vocabulary” (Paradis, 2007: 396). Cognates that occur in the PPVT-4 test are for example: cat-‘kat’, apple-‘appel’, hand-‘hand’, nest-‘nest’, horn-‘hoorn’, pear-‘peer’, chin-‘kin’, neck-‘nek’, blue-‘blauw’, circle-‘cirkel’, hammer-‘hamer’, knee-‘knie’, etc (Naber, 2011).

With the TROG test, the subject also gets to see four pictures and hears a sentence. The subject then has to choose the right picture corresponding to the sentence. Again, there is the problem with language-relatedness; many Dutch and English grammatical structures are similar. The TROG test consists of twenty grammatical structures, almost all of which also exist in Dutch, for example two elements (“the flower is white”), negation (“the girl is not standing”) and X but not Y (“the house but not the tree is big”) (Gidding, 2011: 33-34). The only grammatical structure that is different is the centre-embedded sentence (“the book the pillow is on is yellow”). The grammatical structures build up in degree of difficulty (for native speakers of English). This differs somewhat for the Dutch subjects; even though the grammatical structure in Dutch is similar, the subjects had more difficulty with the reversible *above* and *below* (“the book is below the chair”) and the reversible passive (“the horse is chased by the man”) (Gidding, 2011: 33-34). It seems as though a lack of vocabulary knowledge seems to have affected the scores on the grammar test, which undermines the validity of the test. For example, upon hearing the sentence “the cow is chased by the girl”, 50% of the subjects chose a picture showing “the cow is *fed* by the girl”. Thus, Dutch subjects seemed to have understood the grammatical structure, but they simply did not yet acquire the meaning of the word “chase” (Gidding, 2011: 37).

Gidding also looked at the exposure of English outside the classroom and found that the VVTO group are exposed to English for about 7.4 hours and the EIBO group for only 4.8 hours. The average receptive vocabulary of the 12-year-old VVTO pupil is 6;7 years, the EIBO pupil’s average is 6;2 years. The PPVT-4 test provides a list with vocabulary knowledge of a native speaker deducted from the test score. Considering the limited amount of English that the Dutch subjects have been exposed to, Gidding considers the level of a 6 year-old native speaker to be a good thing. The problem however is that these children are older and have more cognitive advantages over young children in making tests, so their actual

receptive knowledge might be lower. Another issue is that *productive* vocabulary will be even lower than 6 year-old native speakers. More importantly, Gidding's t-test results show that the VVTO group did not significantly differ from the EIBO group in the PPVT-4 test scores ($t(38) = 0.77, p = 0.442$). The VVTO group did positively differ from the EIBO group in the TROG test scores ($t(33.315) = 2.10, p = 0.043$). It is hard to draw conclusions from this, because these tests not only have problems with their validity, they can also not be generalised to being the same as the overall proficiency of the pupils. Grammar and vocabulary are only components of overall language proficiency; other components are reading, writing and oral skills (Eckes & Grotjahn, 2006).

Another thesis also investigated whether VVTO and EIBO pupils differ with regards to their English vocabulary knowledge (Naber, 2011). An experiment was conducted with 79 pupils from grade 6, consisting of 42 VVTO pupils and 37 EIBO pupils. This study also made use of the Peabody Picture Vocabulary Test 4 (PPVT-4) to measure the children's vocabulary knowledge. The questionnaire that was used was developed by FLiPP. The questionnaire aimed at two issues: English exposure of the child outside the classroom and the socio-economic status (SES). Renske Naber (2011) defined socio-economic status (SES) as "the parents' education" and found no significant correlation between parents' education and vocabulary knowledge of the child. Thus, a higher SES does not lead to a higher test score (Naber, 2011). The questionnaire also showed that VVTO pupils watch more English TV and listen to more English music, whereas the EIBO pupils play more English video games (Naber, 2011: 26). When both VVTO and EIBO groups are combined, there is a positive correlation between the score on the PPVT-4 test and time spent on watching English TV programmes ($r = 0.33, p = 0.01$) and playing English video games ($r = 0.26, p = 0.05$) (Naber, 2011: 27). Naber (2011: 4) claims that "English influences outside the classroom play a *huge* role in the level of English the student reaches", which might be somewhat exaggerated since the correlations mentioned here are only medium. Similar to Gidding's findings (2011), Naber's results of the PPVT-4 test also showed no significant difference in knowledge of English vocabulary between the VVTO and EIBO group ($t = 0.21, p = 0.84$) (Naber, 2011). Therefore, it might be concluded that the way English in Dutch early education is currently taught, is not profitable for the pupils' English vocabulary knowledge. Explanations for this could be for example an inefficient teaching method, the little exposure children receive (only 45 minutes a week), the low quality of English (no input from a native speaker) or a combination of these. Even though early education might not be profitable for the pupils'

vocabulary knowledge, it might prove profitable for the overall language proficiency in English.

2.3 Summary

This chapter provided a literature review on factors that have an influence on language proficiency. It highlighted the important difference between formal and informal language learning, and mentioned quality, quantity and age as factors that influence second language learning. The chapter also mentioned some issues concerning the policy of early English language instruction in the Netherlands, for example the low quality of input (less than 10% of the teachers are specialist teachers or native speakers) the low quantity of input (only 45 minutes a week), no official end-terms that have to be met which result in little within-school and between-school agreement, widely varying teaching methods and regular teachers that are not prepared well to teach English. These problems can be solved if (1) more attention is paid to English in the teacher education and (2) the government creates official end-terms (and checking these by, for example, integrating English into the national exam the “Citotoets”). Finally, summaries of studies on the topic of English in EFL primary schools (such as Japan and the Netherlands) were provided, and they showed no (obvious) advantages of young starters over older starters. This study intends to find out whether starting earlier does or does not make a difference in the L2, but not just by looking at one language component such as grammar or vocabulary, but by looking at the overall language proficiency.

3 Method

3.1 Participants

Similar to the previously mentioned studies of Gidding (2011) and Naber (2011), this study will compare two groups of pupils: VVTO pupils in grade 6 (who received 4 to 8 years of English) and EIBO pupils in grade 6 (who received 2 years of English). The six participating primary schools are all situated in the province of Noord-Holland. The pilot took place in an EIBO school in Heemskerk, the five other participating schools are situated in Amsterdam and Diemen. One subject, an acquaintance who happened to be in grade 6 of a school in Amsterdam, also conducted the test. All EIBO school started with English in grade 5, whereas the VVTO schools had three different starting points: preschool, grade 1 and grade 2. A total of 138 subjects participated in this study, of whom 67 were excluded due to factors such as dyslexia (N= 6), native speakers (N= 9), speakers of other languages (N= 37) and subjects that came from a different primary school (N= 15). This left 71 subjects for data analysis. Of this group, 55 are EIBO pupils and 16 are VVTO pupils. The analysed group consisted overall of 39 girls and 32 boys; the EIBO group consisted of 31 girls and 24 boys, and the VVTO group consisted of 8 girls and 8 boys. The average age of all subjects was 11.7 years, 11.67 for the EIBO group (SD= 0.58) and 11.75 for the VVTO group (SD= 0.45).

3.2 Testing materials: *The C-test and Questionnaire*

Language proficiency encompasses many subcategories, such as reading, writing, grammar and vocabulary knowledge (Eckes & Grotjahn, 2006). Each of these subcategories are in itself also broad terms. For example, a learners' vocabulary knowledge can differ in size, quality ("linguistic and pragmatic characteristics") and accessibility ("organization" and retrieval) (Chapelle, 1994: 165). Language proficiency is not only a broad term, it can also be measured in different ways. As an anchor point, language competence can be assessed by testing a language learner's performance on an absolute level (compared to an adult native speaker) or on a relative level (compared to peers) (Klein-Braley, 1985). Choosing a language test that measures general language competence can be difficult, as there are many. Two of these, the cloze test and C-test, are considered "more suitable for assessing the learner's competence than other types of tests" because the actual text used to assess the examinee is irrelevant (Klein-Braley, 1985: 80). This is one of the reasons why the cloze test and C-test are commonly used tools to measure the language proficiency of the (second) language learner.

They both consist of texts with gaps that the subject has to fill in. Of these two tests, the C-test is seen as superior to the cloze test because the latter test has issues with its validity and reliability. A cloze test deletes every n th word in a text (where n is a number between 5 and 10) (Klein-Braley, 1985). Klein-Braley has proven that this systematic deletion is not equivalent to random deletion, because different deletion rates yield different results. Furthermore, cloze tests have a low correlation among different cloze tests which make them less reliable and finally they miss an anchor point, i.e. an “educated native speaker” should be able to obtain a perfect score (Klein-Braley, 1985: 83). The construct validity of the C-test is higher because it does produce random samples and native speakers can obtain “virtually perfect scores” (Klein-Braley, 1985: 84). Other advantages of the C-test are (Katona&Dörnyei, 1993: 1-2):

1. multiple diverse texts are more representative samples and give all subjects a fair chance to show their knowledge;
2. by “damaging” words instead of deleting them, a more representative sample of all the different language elements can be obtained;
3. they are less time-consuming;
4. scoring is objective because in most cases there is only one right answer;
5. students usually like doing C-tests, whereas the cloze test is often a frustrating experience.

The C-test “consists of four to six short (...) texts in the target language” where “the second half of every second word has been deleted, beginning with the second word of the second sentence; the first and last sentences are left intact”, they usually consist of a total of a hundred gaps and “only entirely correct restorations are accepted”, which means that spelling errors are regarded as incorrect (Katona&Dörnyei, 1993: 1).

The purpose of the C-test is to measure general language proficiency. Klein-Braley argues that there is “sufficient evidence to substantiate the claim that C-tests are authentic tests of the construct of general language proficiency” and that these tests “show meaningful relationships to other aspects of language knowledge and performance” (Klein-Braley, 1985: 101). More recently, Hastings (2002) argued that “the value of C-testing as a measure of global proficiency in second language has been demonstrated too many times to be open to dispute” (Eckes & Grotjahn, 2006: 292). The C-test has shown to have a high correlation with other acknowledged language tests such as the Test Of English as a Foreign Language (TOEFL) and the English Language Battery test (ELBA) (Eckes & Grotjahn, 2006).

Moreover, C-tests have been shown to correlate with aspects of language proficiency such as speaking and listening proficiency, grammar and vocabulary (Eckes & Grotjahn, 2006). Thus, many researchers would agree that there is ample evidence that the C-test reliably measures general language ability. However, not all researchers share this view, which is why Eckes & Grotjahn (2006) conducted a methodologically sound and large-scale experiment where they compared the C-test with the German TestDaF. The TestDaF is a standardised language test that measures the components “reading, listening, writing and speaking” that together represent general language proficiency (Eckes & Grotjahn, 2006: 300). As many as 843 subjects took the C-test, but ‘only’ 470 of these completed both the C-test and the TestDaF. Results showed that the C-test is positively correlated with the four components of the TestDaF and that it can be concluded that the C-test is indeed a good tool for measuring general language proficiency.

The text materials used for the C-test in this study were obtained from the website of the British Council (2012), where short stories designed especially for young language learners are found. The materials on the website were designed by teachers that work with young language learners (British Council, 2012). Besides English, the website is available in five other languages, which are probably the first languages of the target groups: Arabic, Russian, French, Spanish and Chinese. Two stories used for the C-test are A1-level, which the British Council classifies as medium level for 6 to 8 year-olds, and two stories are A2-level, which the British Council classifies as hard level for 9 to 11 year-old language learners. The stories can all be found in *Appendix 2: The C-Tests*.

The C-test in this experiment consists of four texts with at least 24 blanks each and contained 109 blanks in total. The participants got five minutes per text and “no details about the deletion principle were provided” (Eckes & Grotjahn, 2006: 303). To be able to measure the beginning learner’s proficiency, familiar subjects that contained high frequency words were chosen. When looking at the frequency of the words in these stories (Cobb, 2012), there does not seem to be any difference in frequency between A1 and A2-level texts: Both A1 and A2-texts contain between 82% and 86% of the 2,000 most frequent words in the English language, of which a vast majority (75-78%) belonged to the 1,000 most frequent words and much less (only 6-9%) belonged to the 1,000-2,000 most frequent words. Two known predictors of text difficulty are type-token ratio and sentence length (Klein-Braley, 1985: 91). These predictors did not seem to differ too much among the short stories either; the sentence length was relatively short (4-16 words) and type-token ratio (0.60-0.72) does not differ too much. The four stories contained 65 to 89 words (with 24 to 31 gaps). The C-test contained a

total of 109 gaps. The correct answers were counted and subjects could get a score between 0-100. Contractions such as ‘can’t’ and ‘don’t’ were counted as one word, compounds such as ‘lion tamer’ as two words. There is only one gap where it is predicted that multiple answers are possible, which is the name ‘Billy’ that occurs only once in the entire text *The Cold Planet*. Logically possible answers here or elsewhere were counted as correct. It is expected that infrequent words that do not occur elsewhere in the text such as ‘batteries’, ‘bounced’ and ‘lion tamer’ will be harder for the subjects and more often filled out incorrectly.

In addition to the C-test, a questionnaire was handed out. The questionnaire asked about the children’s exposure to English outside the classroom (e.g. via the media, such as TV, internet and games), whether the children ever attended a different primary school and what language(s) they spoke at home. The children with dyslexia were also asked to indicate this on their paper. If the children ever attended a different school, spoke different languages at home or had dyslexia, the subjects were excluded from the analysis in this study. In addition to this, the teacher also filled out a questionnaire, that concerned the following: Starter Age (SA) of instruction in English at that school, the teacher’s language background, i.e. whether the teacher is a Native Speaker (NS) or a Non-Native Speaker of English, the average instruction time per week and (if used) the teaching method.

3.3 Procedure

The Social Economic Status (SES) for all six schools was checked by requesting status scores from the ‘Sociaal en Cultureel Planbureau’. This status score is provided per postal code and the score is achieved by taking into account three factors: income, employment and level of education (Sociaal en Cultureel Planbureau, 2012). The average score is 0 and deviating scores can be seen as the standard deviation. A negative score indicates a high status, a positive score a low status. In the Netherlands, the most current status scores (those from 2010) differ from -3.4 to +5.2 (Sociaal en Cultureel Planbureau, 2012). The participating schools in this study differ from -2.8 to +1.3. The regular EIBO schools score 0.2, 0.2 and 0.9. The one subject tested outside the school environment attends an EIBO school that scores -1.2. The VVTO schools score -2.8, 0.7 and 1.3 and therefore have a wide ranging SES. Because of these differences in SES, within-group Independent Samples T-tests will be conducted to see whether schools themselves significantly differ from each other.

Before the subjects made the tests, they received some instruction. The subjects were told to fill out only one word per gap to complete the text and that they had only five minutes

per text (but in reality they received more time; the EIBO school that performed poorest also received most time, around 30 minutes). This time limit was set to prevent the subjects from lingering on certain gaps for too long, with the benefit that this way the test would not take up too much time. The subjects got all four texts and questionnaire at the same time. They were asked to first do the test and then complete the questionnaire.

A pilot was conducted to make sure the C-test was not too easy or too hard. The pilot (originally N= 25 but excluding the majority -because of other languages N= 8, different school N= 5 and native speakers of English N= 2- left N= 10) shows that the scores on the C-test is normally distributed, but generally quite difficult. Even though the C-test contains relatively easy and common words, the average score of the pilot group is low: 29.7%. SPSS 18.0.3 was used to compute the normal distribution of the C-test as a whole (i.e. the four texts taken together). The Shapiro-Wilk test shows no significant value ($p= 0.63$), thus the data is normally distributed. A histogram also shows this (see Figure 1 below). The C-test therefore sufficiently represents the children’s proficiency. In addition, an adult educated native speaker filled out the test and scored a 100% (see Appendix 3).

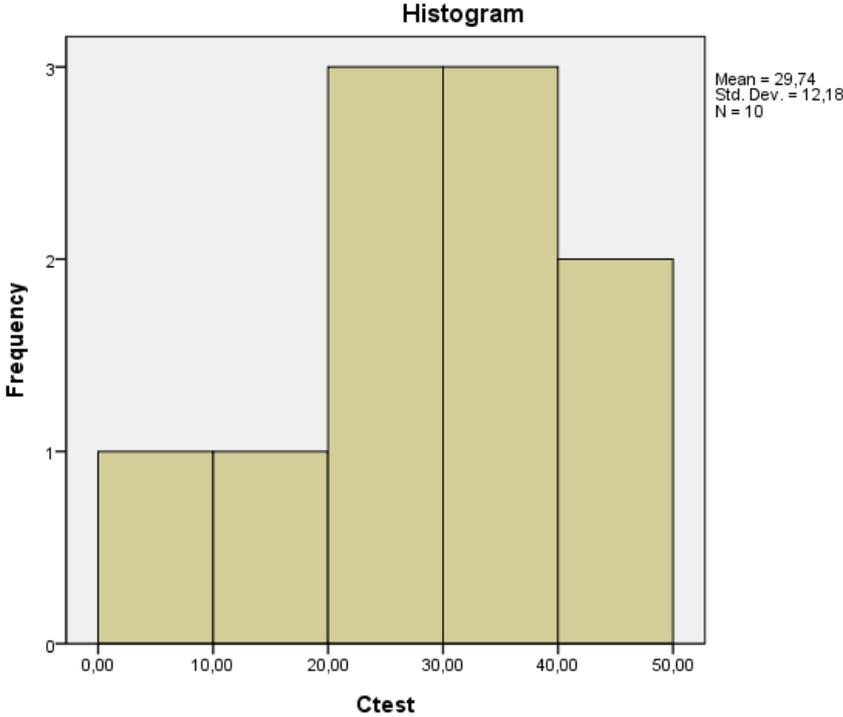


Figure 1: The histogram of normal distribution

4 Results

First, a between-group Independent Samples T-test was conducted to see whether differences in scores between the EIBO group and the VVTO group are significant. Then within-group Independent Samples T-tests were conducted to see whether EIBO schools significantly differ in scores or whether VVTO schools significantly differ in scores. Next, findings are presented from the pupils' questionnaire, followed by findings from the teachers' questionnaire. This chapter ends with two side notes that are worth mentioning, one point regarding the frequency of words and how easy or difficult they proved for the Dutch subjects, the other point regarding possible or incorrect deviating answers.

Of the total of a 109 gaps, correct answers were counted and converted into a score of a 100% (e.g. 83 correct gaps is a score of 76.1%). An Independent Samples T-test was conducted, comparing the scores of the EIBO group to the scores of the VVTO group. The T-test shows that the VVTO group scored significantly higher ($M= 62.2$, $SD= 13.2$) than the EIBO group ($M= 43.3$, $SD= 17.9$); ($t(69)= -3.909$, $p= 0.000$). The Levine's Test for Equality of Variances is not significant ($p= 0.11$), therefore the variances are assumed to be equal for both groups. The mean score of the EIBO group is 43.3%, the mean score of the VVTO group is 62.2%. The standard deviation within the EIBO group is bigger, which is surprising since the VVTO group consists of subjects that started to learn English at different ages.

To find out whether there are significant differences in results within the compared groups, three T-tests were conducted for the EIBO schools and one T-test was conducted for the VVTO schools. The Independent Samples T-test shows no significant difference in scores between EIBO school 'Tweede Openluchtschool' and 'De Nieuwe Kring' ($t(19.440)= -0.634$, $p= 0.534$). The Levine's Test for Equality of Variances is significant ($p= 0.01$), therefore equal variances are not assumed. The Independent Samples T-test also shows no significant difference in scores between EIBO school 'Tweede Openluchtschool' and 'De Ark' ($t(22.039)= 1.886$, $p= 0.072$). The Levine's Test for Equality of Variances is significant ($p= 0.03$), therefore equal variances are not assumed. The Independent Samples T-test does show a significant difference in scores between EIBO school 'De Nieuwe Kring' and 'De Ark', even with the Bonferroni correction where $p < 0.025$ ($t(37)= 3.524$, $p= 0.001$). The Levine's Test for Equality of Variances is not significant ($p= 0.38$), therefore equal variances are assumed. However, even if the less performing EIBO school is left out, the VVTO schools still perform better ($t(59)= -3.273$, $p= 0.002$) (with equal variances assumed $p= 0.14$). Finally, the Independent Samples T-test shows no significant difference in scores between VVTO

school ‘De Vlinderboom’ and ‘Willemsparkschool’ ($t(13) = -0.844$, $p = 0.414$). The Levine’s Test for Equality of Variances is not significant ($p = 0.79$), therefore equal variances are assumed. The T-Test could not be conducted with the ‘Louise de Colignyschool’ because only 1 subject of this school could be included in the analysis.

Regarding the pupils’ questionnaire, Pearson correlation tests were conducted to see whether there is a relationship between time spent on different activities and test scores, but it should be kept in mind that the times the subjects indicated are -very- rough estimates. Comparing the EIBO and VVTO groups, there are no apparent differences that emerge (see Figure 2). Both groups report to barely: read English stories, hold conversations in English or repeat the English they learned at school. Both groups report to frequently: play English (video) games, watch English TV (with or without Dutch subtitles) and listen to English music. Similar to Naber’s study (2011), VVTO pupils watch more English TV, whereas the EIBO pupils report to play more English (video) games. Opposite to Naber’s findings, however, this study shows that EIBO pupils report to listen to more English music than VVTO pupils. Naber (2011) found a positive correlation between the test score and time spent on watching English TV programmes and playing English video games. This study found no correlations between any ways of exposure to English outside the classroom and score except for test scores and reading English stories ($r = 0.28$, $p = 0.049$). This positive relationship might indicate that those who read more have a higher test score or the other way around. The relationship found is only medium, however, and since these times are rough estimates, no serious conclusions can be drawn from this. Finally, Gidding reported that VVTO pupils are exposed to English outside school for about 7.4 hours compared to 7.7 hours in this study, and the EIBO group for only 4.8 hours, compared to 9.2 hours in this study. Because of this substantial difference in the EIBO group (4.8 hours vs. 9.2 hours), it is difficult to indicate whether EIBO pupils do or do not get exposed to more English outside school than VVTO pupils. The different findings might also indicate that the times indicated in the questionnaire are too rough an estimation to be able to draw solid conclusions from this.

Time spent on... (min./week)	Eng. stories	Eng. conversations	Eng. games	Dutch TV with Eng.	Eng. TV	Eng. music	Eng. repetition	Total	
								min.	hours
EIBO	3	31	127	78	121	181	8	550	9.2
VVTO	15	11	99	31	154	136	11	459	7.7

Figure 2: Time spent on English outside school

Regarding the teachers' questionnaire, all three EIBO schools and one VVTO school had a regular teacher, but two VVTO schools had a specialist teacher. From this study, it cannot be said with certainty whether the type of teacher influences the pupils' level of proficiency, but there is a tendency for 'no'. Since none of the teachers in this study were native speakers, it unclear whether the specialist teachers have a higher proficiency of English than the regular teachers. Moreover, this study shows that the VVTO school with a regular teacher scores as high as another VVTO school with a specialist teacher (55.0% vs. 55.1% respectively). About quantity, both EIBO and VVTO schools teach English around 45 minutes a week, which is similar to the findings of the previously mentioned survey of Thijs et al (2011). Finally, little can be said about the teaching method. Because all participating schools use different teaching methods, no generalisations can be made concerning the type of input. The VVTO schools reported to use:

- kindergarten to grade 2: *Playway*, grade 3 and 4: *Just Do It!*, grade 5 and 6: *Making a Start*;
- no teaching method;
- *Bubbles*.

The EIBO schools use:

- *Alles in 1* and *Alles Apart*;
- *Take it Easy*;
- *Junior*.

Teachers reported to be unhappy with *Take it Easy* (because the method does not offer enough grammar) and *Junior* (because it is not a comprehensive method) and the schools wanted to use new teaching methods.

The expectation that mostly infrequent words such as 'batteries', 'bounced' and 'lion tamer' would be filled out incorrectly, was not fulfilled. Maybe because of contextual information such as 'torch' and 'circus', subjects did often correctly fill in 'battery/batteries' and 'tamer'. These words also resemble Dutch words 'batterij' and 'temmer' respectively. These similar words with similar meanings, also called cognates, will always remain an issue to some extent due to the language relatedness of Dutch and English. The infrequent word 'bounced', which does not resemble the Dutch equivalent ('stuitten'), did prove to be difficult. Other words that seemed difficult were 'shouted', 'suddenly', 'beam', 'moved',

‘famous’ and ‘tower’. It seems as though the subjects have not yet mastered all of the most frequent (2,000) words in the English language (see Appendix 2). The test had more possible answers than initially expected. Deviating answers that were counted as correct were for example:

- ‘her mum was *waiving* at the gates’ (instead of *waiting*);
- ‘*can*’t/could we go to the park’ (instead of *can*);
- ‘Neila put the wrong *battery* in *his* torch’ (instead of *batteries* and *her*);
- ‘Suddenly, there was a dazzling *being of life*’ (instead of *beam* and *light*);
- ‘The light hit a *bee/bat* called Billy’ (instead of *boy*);
- ‘*before* the pair were sucked up’ (instead of *but*);
- ‘It was the *director* of the circus’ (instead of *day*);
- ‘A door opened and *two/ten* lions bounced into the ring’ (instead of *the*).

Answers that were counted as incorrect were for example:

- ‘you can’t come in *her*, Jessie! *showed* Katie *at* Jaia’ (instead of *here*, *shouted* and *and*);
- ‘a boy called Billy and *hot* dog’ (instead of *his*);
- ‘the pair were *success* up’ (instead of *sucked*);
- ‘the world’s *family* landmarks’ (instead of *famous*);
- ‘*She* don’t worry’ (instead of *so*);
- ‘So don’t *work*’ (instead of *worry*);
- ‘Billy loved *to* circus’ (instead of *the*);
- ‘the lions *bought* into the ring’ (instead of *bounced*).

5 Conclusion and Discussion

Unlike similar previous studies on this topic, this study does show a significant difference in scores between VVTO schools on the one hand and EIBO schools on the other hand. In line with previous studies conducted in the area of informal language contexts, the results in this study suggest that starting sooner in a formal language context does have a positive effect on one's proficiency in English. The positive effect of starting sooner might not necessarily be due to the factor age, but due to the factor quantity. It is interesting that the little quantity of receiving English classes for 45 minutes a week in earlier grades does seem to significantly affect the child's proficiency in English. This study is not unequivocal proof that starting language instruction earlier will guarantee a higher proficiency in English (due to the small sample size, especially regarding the VVTO pupils), but it is a start. For a more extensive and reliable outcome, the results from the FLiPP project will be very relevant for schools, even though the test instruments used in this project have some methodological issues.

For practical implications, this study implies two things. Since VVTO pupils have a higher English proficiency, secondary schools should take this into account by building on the knowledge that these pupils already have. These pupils could follow a more challenging English programme than late starters, because Oller and Nagato's study (1974) already showed that the initial advantages that early starters have will disappear if they are put in the same class as late starters and receive similar English instruction.

Another practical implication of this study is that school policies are headed in the 'right' direction, which is to start offering a foreign language as soon as possible. The European Platform offers subsidies to schools that offer VVTO at their schools. Therefore, a suggestion for all EIBO schools is that it would indeed be a good idea to implement English as soon as possible in their school curriculums. Fortunately, primary schools in Amsterdam seem to have quite some interaction amongst one another. For example, one VVTO school that is really enthusiastic about teaching English as soon as possible gave a presentation for other primary schools about VVTO English. In fact, one EIBO school that went to this presentation -and which also participated in this study- is already going to change its policy by offering English in earlier grades starting from next school year.

There was also a significant difference in scores between two EIBO schools. One EIBO school, the school of the pilot, performed significantly more poorly. Reasons for why this school might perform poorer could be many, for example the teaching method *Take it Easy*. During the pilot, the teacher already reported that the pupils had poor English language

skills, and she indicated she was displeased with the teaching method because it offers little support. Another reason could be because the pilot school is situated in a town that has a less international environment than Amsterdam. There are less tourists or international companies where expats work (and less children of expats who go to school there). It would be interesting to see whether the VVTO schools in less international environments also perform better than EIBO schools. If they do not, there could be an interaction effect between exposure to English outside school and type of education. It would be interesting to compare the exposure to English outside school indicated by VVTO schools in these towns with the exposure indicated by VVTO schools in international cities. However, this study found little evidence for interaction effects between test score and exposure to English outside school, so it is expected that VVTO schools in small towns will also perform better than EIBO schools. Maybe an international environment positively influences motivation for learning English.

Regarding the relationship between test score and time spent on activities concerning English, only a relationship between test score and time spent on reading English stories was found. As mentioned before, the times indicated by the subjects are rough estimates, and it is expected that the time indicated here is underestimated (i.e. in reality the subjects will spend more time on for example playing video games and watching English TV than they indicated). However, to really see whether there is a relationship between spending time on, for example, watching English TV, playing English video games or reading English stories, a more solid experiment needs to be conducted. Instead of making use of a simple questionnaire, the subjects could be asked to keep track of a day-to-day journal about their daily activities outside school revolving English activities. This journal would probably give the researcher more reliable results.

As a side note, it is interesting to see a pattern in the acquisition of English vocabulary appear, a sort of similar interlanguage amongst the subjects. Instead of filling out the correct words, some errors that were quite consistent were ‘show(ed)’ instead of ‘shouted’, ‘being’ instead of ‘beam’, ‘mouse’ instead of ‘moved’ and ‘family’ instead of ‘famous’. The ‘mouse’ instead of ‘moved’ was related to ‘giant’ (*a naughty young giant mouse*); it seems as though Dutch children first learn the meaning of the adjective ‘giant’ and only later acquire the meaning of the noun. It also appears as though ‘show’ is acquired before ‘shout’, ‘being’ before ‘beam’ and ‘family’ before ‘famous’. It also makes sense that the subjects have not yet mastered all of the most frequent (2,000) words in the English language, because the majority has received English language instruction for only two years.

Suggestions for future research could be the influence of native(-like) input on the English proficiency or the possible influence of high-quality input in interaction with the influence of quantity. The FLiPP project is looking at both of these factors, but the testing materials they use (the PPVT-4 test amongst others) have problems with validity and reliability, so there is still room for improvement. The C-test would be a better tool to measure the subject's language proficiency. Other possibilities are investigating the relationship between time spent on English activities outside school and the proficiency in English. Intuitively it would make sense that the more time one spends on English, the higher one's proficiency. This study shows that EIBO pupils report to spend more time on English outside school than VVTO pupils, but a more reliable testing method such as a journal should give more clarity about this (since Gidding's study shows that EIBO pupils spend less time on English outside school compared to VVTO pupils). Finally, it would also be worthwhile to analyse the different teaching methods that are used in primary schools. The VVTO schools did better than the EIBO schools, but they also reported to be contented with the teaching methods they use (*Playway*, *Just Do It!* and *Making a Start*). Two out of three EIBO schools reported to be discontented with the teaching methods (*Take it Easy* and *Junior*). It would be interesting to look at the differences between these teaching methods, and what can be improved in *Take it Easy* and *Junior*.

6 Appendices

Appendix 1: The questionnaires

The questionnaire for the pupils. This questionnaire was (partly) taken from the Foreign Languages in Primary Schools Project (FLiPP) (Naber, 2011: 45).

Naam:

Ik ben een: meisje / jongen

Leeftijd:

Nationaliteit:

School:

- Zit je op deze school sinds groep 1? Ja / nee
- Spreek je thuis alleen Nederlands met je ouders? Ja / nee
- Zo **nee**, welke taal/talen spreek je dan nog meer? _____
- Met wie spreek je Engels?
 - Met niemand
 - Alleen op school met de juf/meester
 - Met _____ personen, waarvan _____ personen moedertaalsprekers van het Engels zijn.
- Ga je wel eens op vakantie naar een ander land waar je dan Engels spreekt?
 - Ik ga niet op vakantie naar een land waar ik dan Engels spreek
 - Ik ga op vakantie naar _____, ongeveer _____ weken per jaar.
- Hoeveel tijd ben je per week met Engels bezig buiten school:

		Aantal min. per week	Met wie? (meerdere opties mogelijk)	Dit is verspreid over...
Engelstalige verhaaltjes		----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag
Engelstalige gesprekken met familieleden / vrienden		----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag
Engelstalige (computer)spelletjes (Bijvoorbeeld de Wii of SimCity)		----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag
Nederlandse TV-programma's met Engelse woorden		----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per

<i>(Bijvoorbeeld Dora)</i>		<input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	week <input type="checkbox"/> Iedere dag
Engelstalige TV <i>(Bijvoorbeeld series of films)</i>	----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag
Engelstalige liedjes <i>(Bijvoorbeeld popmuziek)</i>	----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag
Het herhalen van Engels wat je op school hebt geleerd	----- min/week	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Alleen <input type="checkbox"/> Broer/Zus <input type="checkbox"/> Ouders <input type="checkbox"/> Vrienden	<input type="checkbox"/> N.v.t. <input type="checkbox"/> Ongeveer 1 dag <input type="checkbox"/> Meerdere dagen per week <input type="checkbox"/> Iedere dag

The questionnaire for the teacher.

Naam:

Geslacht:

Ik ben: groepsdocent/vakdocent

Ik geef les op school:

Nationaliteit(en):

1. Ik ben een / geen moedertaalspreker van het Engels.
2. De leerlingen beginnen met het leren van Engels in groep: _____
3. Ik geef gemiddeld _____ minuten per week Engels.
4. We maken gebruik van een / geen lesmethode Engels.
5. (Zo **ja**) we maken gebruik van _____

Appendix 2: The C-Tests.

Text 1: No Dogs! (A1)

It was half past three and Katie had just finished school. Her mum was waiting at the gates with Jessie, the dog.

“Can we go to the park with Jaia, mum?” said Katie.

“Alright, we can go for half an hour,” said mum.

When they got to the park, Katie and Jaia ran towards the swings and slides.

“You can’t come in here, Jessie!” shouted Katie and Jaia. Mum took Jessie over to the bench and tied him to it.

It was half past three and Katie had just finished school. Her m___ was wai___ at t___ gates wi___ Jessie, t___ dog.

“C___ we g___ to t___ park wi___ Jaia, m___?” said Ka___.

“Alright, w___ can g___ for ha___ an ho___,” said m___.

When th___ got t___ the pa___, Katie a___ Jaia r___ towards t___ swings a___ slides.

“Y___ can’t co___ in he___, Jessie!” sho___ Katie a___ Jaia. Mum took Jessie over to the bench and tied him to it.

	Families	Types	Tokens	Percent
K1 Words (1-1000):	39	39	63	77.78%
Function:	(39)	(48.15%)
Content:	(24)	(29.63%)
> Anglo-Sax	(16)	(19.75%)
=Not Greco-Lat/Fr Cog:
K2 Words (1001-2000):	5	5	6	7.41%
> Anglo-Sax:	(4)	(4.94%)
1k+2k			...	(85.19%)

Words in text (tokens):	81
Different words (types):	49
Type-token ratio:	0.60
Tokens per type:	1.65
Lex density (content words/total)	0.52

it was half past three and katie had just finished school her mum was waiting at the gates with jessie the dog can we go to the park with jaia mum said katie alright we can go for half an hour said mum when they got to the park katie and jaia ran towards the swings and slides you can not come in here jessie shouted katie and jaia mum took jessie over to the bench and tied him to it

(Cobb, 2012)

Text 2: The Cold Planet (A1)

Thousands of miles from Earth, on the far side of the universe, is a small planet. It is dark and cold. Little green aliens live there. They use torches to see.

One day, a young alien, Neila, put the wrong batteries in her torch. Suddenly there was a dazzling beam of light. It hit planet Earth.

The light hit a boy called Billy and his dog. Neila quickly turned the torch off, but the pair were sucked up by the light. They flew through space and landed near Neila.

Thousands of miles from Earth, on the far side of the universe, is a small planet. It i___ dark a___ cold. Lit___ green ali___ live th___. They u___ torches t___ see.

O___ day, a yo___ alien, Ne___, put t___ wrong batt___ in h___ torch. Sudd___ there w___ a dazzling be___ of li___. It h___ planet Ea___.

The li___ hit a b___ called Bi___ and h___ dog. Ne___ quickly tur___ the to___ off, b___ the pa___ were suc___ up b___ the li___. They flew through space and landed near Neila.

	<u>Families</u>	<u>Types</u>	<u>Tokens</u>	<u>Percent</u>
K1 Words (1-1000):	46	48	68	76.40%
Function:	(38)	(42.70%)
Content:	(30)	(33.71%)
> Anglo-Sax	(25)	(28.09%)
=Not Greco-Lat/Fr Cog:		
K2 Words (1001-2000):	7	7	8	8.99%
> Anglo-Sax:	(6)	(6.74%)
1k+2k			...	(85.39%)

Words in text (tokens):	89
Different words (types):	64
Type-token ratio:	0.72
Tokens per type:	1.39
Lex density (content words/total)	0.57

thousands of miles from earth on the far side of the universe is a small planet it is dark and cold little green aliens live there they use torches to see one day a young alien neila put the wrong batteries in her torch suddenly there was a dazzling beam of light it hit planet earth the light hit a boy called billy and his dog neila quickly turned the torch off but the pair were sucked up by the light they flew through space and landed near neila

(Cobb, 2012)

Text 3: Pyramids in Paris (A2)

Something very strange happened last month. A naughty, young giant moved all the world’s famous landmarks. He put the pyramids in Paris.

“And on your left the ... pyramids?”

He moved the Leaning Tower of Pisa to London.

“Here we can see the famous Leaning Tower of ... London?”

His mum told her son to put them back. So don’t worry. If you go to London, Cairo, Pisa or Paris today, you’ll see all the landmarks in their correct place.

Something very strange happened last month. A naughty, yo__ giant mo__ all t__ world’s fam__ landmarks. H__ put t__ pyramids i__ Paris.

“A__ on yo__ left th__ ... pyramids?”

H__ moved t__ Leaning To__ of Pi__ to Lon__.

“Here w__ can s__ the fam__ Leaning To__ of ... Lon__?”

His m__ told h__ son t__ put th__ back. S__ don’t wo__. If you go to London, Cairo, Pisa or Paris today, you’ll see all the landmarks in their correct place.

	Families	Types	Tokens	Percent
K1 Words (1-1000):	41	44	60	75.00%
Function:	(34)	(42.50%)
Content:	(26)	(32.50%)
> Anglo-Sax	(19)	(23.75%)
=Not Greco-Lat/Fr Cog:		
K2 Words (1001-2000):	4	4	6	7.50%
> Anglo-Sax:	(2)	(2.50%)
1k+2k			...	(82.50%)

Words in text (tokens):	80
Different words (types):	56
Type-token ratio:	0.70
Tokens per type:	1.43
Lex density (content words/total)	0.57

something very strange happened last month a **naughty** young **giant** moved all the world famous **landmarks** he put the **pyramids** in **paris** and on your left the **pyramids** he moved the **leaning tower** of **pisa** to **london** here we can see the famous **leaning tower** of **london** his mum told her son to put them back so do not **worry** if you go to **london** **cairo** **pisa** or **paris** today you will see all the **landmarks** in their **correct** place

(Cobb, 2012)

Text 4: Circus (A2)

Billy was getting excited. It was the day of the circus! Billy loved the circus – all the animals and the people. It was so much fun!

At eight o'clock Billy and his parents arrived. A door opened and the lions bounced into the ring. The man in the middle was the lion tamer. Billy watched as he put his head in the biggest lion's mouth.

Billy was getting excited. It w___ the d___ of t___ circus! Bi___ loved t___ circus – a___ the ani___ and t___ people. I___ was s___ much f___!

At ei___ o'clock Bi___ and h___ parents arr___. A door ope___ and t___ lions bou___ into t___ ring. T___ man i___ the mid___ was t___ lion ta___. Billy watched as he put his head in the biggest lion's mouth.

	Families	Types	Tokens	Percent
K1 Words (1-1000):	31	32	50	76.92%
Function:	(32)	(49.23%)
Content:	(18)	(27.69%)
> Anglo-Sax	(14)	(21.54%)
=Not Greco-Lat/Fr Cog:
K2 Words (1001-2000):	4	4	4	6.15%
> Anglo-Sax:	(2)	(3.08%)
1k+2k			...	(83.07%)

Words in text (tokens):	65
Different words (types):	42
Type-token ratio:	0.65
Tokens per type:	1.55
Lex density (content words/total)	0.51

billy was getting **excited** it was the day of the **circus** **billy** loved the **circus** all the animals and the people it was so much **fun** at eight **clock** **billy** and his **parents** arrived a door opened and the **lions** **bounced** into the ring the man in the middle was the **lion tamer** **billy** watched as he put his head in the biggest **lion** **mouth**

(Cobb, 2012)

Appendix 3: C-test by educated native speaker

No Dogs!

It was half past three and Katie had just finished school. Her mum was waiting at the gates with Jessie, the dog.

“Could we go to the park with Jaia, mum?” said Katie.

“Alright, we can go for half an hour,” said mum.

When they got to the park, Katie and Jaia ran towards the swings and slides.

“You can’t come in here, Jessie!” shouted Katie and Jaia. Mum took Jessie over to the bench and tied him to it.

The Cold Planet

Thousands of miles from Earth, on the far side of the universe, is a small planet. It is dark and cold. Little, green aliens live there. They use torches to see.

One day, a young alien, *Nebub*, put the wrong battery in his torch. Suddenly there was a dazzling beam of light. It hit planet Earth.

The light hit a boy called Billy and his dog. *Nebub* quickly turned the torch off, but the pair were sucked up by the light. They flew through space and landed near Neila.

(only found out at the end that the alien was called *Neila*; so thought of my own alien name (as you can see). I left it in, because I have no idea what you are testing and it might be an interesting outcome. BTW, I thought this was the most difficult story to fill in; especially the part where they were sucked up by the light (not even sure if I got it right))

Pyramids in Paris

Something very strange happened last month. A naughty, young giant moved all the world’s famous landmarks. He put the pyramids in Paris.

“And on your left the ... pyramids?”

He moved the Leaning Tower of Pisa to London.

“Here we can see the famous Leaning Tower of ... London?”

His mum told her son to put them back. So don’t worry. If you go to London, Cairo, Pisa or Paris today, you’ll see all the landmarks in their correct place.

Circus

Billy was getting excited. It was the day of the circus! Billy loved the circus – and the animals and the people. It was so much fun!

At eight o’clock Billy and his parents arrived. A door opened and the lions bounced into the ring. The man in the middle was the lion tamer. Billy watched as he put his head in the biggest lion’s mouth.

Appendix 4: The Results

Tweede Openluchtschool		Frederik Roeskestraat 76, 1076 ED Amsterdam				
Average 45 min. Eng. per week		Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 5	REGULAR	1	33.0	0		12
N=25		2	35.8	0		11
Dyslexia: 2		3	45.9			1 11
Native speakers: 2		4	33.9	0		11
Other language: 2		5	83.5			1 11
Different school: 4		6	67.0			1 12
N=15		7	54.1			1 11
		8	67.9	0		11
SES:		9	14.7			1 12
0.24449		10	1.8			1 12
		11	20.2	0		12
		12	57.8			1 11
		13	37.6	0		12
		14	71.6			1 ?
		15	23.9	0		11
			43.25	7		8 11.4

De Vlinderboom		Nicolaas Beetstraat 40, 1053 RM Amsterdam				
Average 45 min. Eng. per week		Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 2	SPECIALIST	1	64.2	0		12
N=14		2	45.9			1 11
Other language: 12			55.05	1		1 11.5
N=2						
SES:						
0.72011						

Willemsparkschool		Willem Witsenstraat 12, 1077 AZ Amsterdam				
Average 60 min. Eng. per week		Subject	Score	Girl: 0	Boy: 1	Age
Begin: preschool	SPECIALIST	1	79.8			1 12
N=21		2	45.9			1 12
Dyslexia: 1		3	78.9			1 12
Native speakers: 4		4	67.0	0		12
Different school: 3		5	76.1	0		12
N=13		6	56.0	0		11
		7	33.0			1 12
		8	56.9			1 12
SES:		9	58.7			1 12
-2.84047		10	70.6	0		11
		11	75.2	0		12
		12	62.4	0		12
		13	68.8	0		11
			63.8	7		6 11.8

Louise de Colignyschool
Amsterdam

Average 30 min. Eng. per week	Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 1 REGULAR	1	55.0			1 12
N=11					
Other language: 10					
N=1					

SES:
1.31470

De Nieuwe Kring	Prinses Beatrixlaan 156 c, 1111 GH Diemen				
Average. 30-45 min. Eng. per week	Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 5 REGULAR	1	52.3		0	11
N=41	2	47.7			1 12
Dyslexia: 3	3	19.3			1 12
Native speakers: 1	4	37.6		0	11
Other language: 5	5	61.5		0	12
Different school: 3	6	55.0			1 12
N=29	7	51.4		0	12
	8	64.2		0	12
SES:	9	48.6		0	12
0.19487	10	61.5		0	13
	11	76.1		0	11
	12	58.7			1 11
	13	57.8		0	11
	14	50.5		0	11
	15	20.2			1 12
	16	57.8		0	12
	17	49.5		0	12
	18	37.6		0	12
	19	58.7		0	12
	20	55.0		0	11
	21	43.1		0	12
	22	32.1		0	13
	23	26.6		0	12
	24	33.0			1 13
	25	24.8			1 11
	26	41.3			1 11
	27	57.8		0	11
	28	38.5			1 11
	29	56.9		0	12
		47.42		20	9 11.7

Pilot: De Ark	Elbestraat 4, 1966 XJ Heemskerk				
Average 45 min. Eng. per week	Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 5 REGULAR	1	29.4		0	12
N=25	2	45.0			1 11
Native speakers: 2	3	23.9			1 12
Other language: 8	4	28.4			1 12
Different school: 5	5	7.3		0	12

N=10	6	30.3	0		11
	7	38.5	0		12
SES:	8	34.9		1	12
0.947733	9	45.0		1	12
	10	14.7		1	12
		29.74	4	6	11.8

Watergraafsmeerse Scholenvereniging	Copernicusstraat 38-40, 1098 JH Amsterdam				
N=1	Subject	Score	Girl: 0	Boy: 1	Age
Begin: Grade 5	1	61.5			12
SES:					
-1.17027					

Bibliography

- Árva, V. & P. Medgyes (2000). Native and Non-native Teachers in the Classroom. *System* 28, 355-372.
- Barik, H. & M. Swain (1975). Three-year Evaluation of a Large Scale Early Grade French Immersion Program: The Ottawa Study. *Language Learning* 25, 1–30.
- Birdsong, D. & M. Molis (2001). On the Evidence for Maturation Constraints in Second-Language Acquisition. *Journal of Memory and Language* 44, 235-249.
- Bot, K. de & A. Herder (2005). Vroeg Vreemdetalenonderwijs in Internationaal Perspectief: Literatuurstudie. Rijksuniversiteit Groningen: Expertisecentrum taal, onderwijs en communicatie. 31 March 2012 <www.earlybirdie.nl/file.php?id=7>
- British Council (2012). *Learn English Kids*. 1 April 2012
<[http://learnenglishkids.britishcouncil.org/en/category/language-level/a1-medium?type\[\]=short_stories](http://learnenglishkids.britishcouncil.org/en/category/language-level/a1-medium?type[]=short_stories)>,
<[http://learnenglishkids.britishcouncil.org/en/category/language-level/a2-hard?type\[\]=short_stories](http://learnenglishkids.britishcouncil.org/en/category/language-level/a2-hard?type[]=short_stories)> and <<http://learnenglishkids.britishcouncil.org/en/parents>>
- Burstall, C., Jamieson, M., Cohen, S. & M. Hargreaves (1974). *Primary French in the Balance*. Windsor: NFER Publishing Company Ltd.
- Cenoz, J. (2003). The Influence of Age on the Acquisition of English: General Proficiency, Attitudes and Code-mixing. In M. Mayo and M. Lecumberri (eds), *Age and the Acquisition of English as a Foreign Language*. Clevedon: Multilingual Matters Ltd.
- Cenoz, J. (2007). Review on C. Munoz (ed.): AGE AND THE RATE OF FOREIGN LANGUAGE LEARNING. *Applied Linguistics* 28-4, 627-630.
- Chapelle, C. (1994). Are C-tests Valid Measures for L2 Vocabulary Research? *Second Language Research* 10-2, 157-187.
- Chevalking, J.W. (2010). *The Effect of L2 Input Quality & Quantity on Early English L2 Acquisition*. Thesis, Utrecht University. 1 April 2012,
<<http://igitur-archive.library.uu.nl/student-theses/2010-0902-200254/UUindex.html>>
- Cito (2012). *Me2! Engels voor groep 7 en 8*. 1 April 2012,
<http://www.cito.nl/nl/onderwijs/primair%20onderwijs/alle_producten/3227a460241f4f2da733ecbcab69dbff.aspx>
- Cobb, T. (2012). *Web Vocabprofile*. 1 April <<http://www.lex tutor.ca/vp/eng>>, an adaptation of Heatley & Nation's (1994) *Range*. Victoria University of Wellington, NZ. Computer programme <<http://www.vuw.ac.nz/lals/>>
- DeKeyser, R. (2000). The Robustness of Critical Period Effects in Second Language Acquisition. *Studies in Second Language Acquisition* 22, 499-533.

- Dornyei, Z (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah: Lawrence Erlbaum.
- Eckes, T. & R. Grotjahn (2006). A Closer Look at the Construct Validity of C-tests. *Language Testing* 23-3, 290-325.
- Europees Platform (2012). *Engelse Les aan Kleuters Tast Niveau Nederlands niet Aan*. 17 June 2012, <<http://europeesplatform.m13.mailplus.nl/archief/mailling-366522.html>>
- Europees Platform (2012). *Tto Tweetalig Onderwijs - Content and Language Integrated Learning*. 26 July 2012, <<http://www.europeesplatform.nl/sf.mcgi?2682>>
- Europees Platform (2012). *VVTO Engels*. 1 April 2012, <<http://www.europeesplatform.nl/sf.mcgi?3618&cat=765>>
- Flege, J., Munro, M. & I. MacKay (1995). Effects of Age of Second-Language Learning on the Production of English Consonants. *Speech Communication* 16, 1-26.
- Genesee, F. (1998). French Immersion in Canada. In J. Edwards (ed.) *Language in Canada*. Cambridge: Cambridge University Press.
- Gidding, M. (2011). *Vroeg Engels in het Basisonderwijs: Een Onderzoek naar de Opbrengst van Engels vanaf Groep 1*. Thesis, Universiteit Utrecht. 1 April 2012 <<http://igitur-archive.library.uu.nl/student-theses/2011-0802-200945/MA%20Scriptie%20Marielle%20Gidding.pdf>>
- Goorhuis-Brouwer, S. & K. de Bot (2005). Heeft Vroeg Vreemde-talenonderwijs een Negatief Effect op de Nederlandse Taalontwikkeling van Kinderen? *Levende Talen* 6-3, 3-7.
- Heesters, K., Feddema, M. Schoot, F. van der & B. Hemker (2008). *Balans van het Engels aan het Einde van de Basisschool 3: Uitkomsten van de Derde Peiling in 2006*. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.
- Holten, M. van (2011). *Kwaliteitsverschillen in VVTO: Een Case Study*. Thesis, Universiteit Utrecht. 2 May 2012 <http://igitur-archive.library.uu.nl/student-theses/2011-0826-202313/Bachelor_Scriptie_Marinka_vanHolten.pdf>
- Johnson, J. & E. Newport (1989). Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology* 21, 60-99.
- Katona, L. & Z. Dörnyei (1993). The C-Test: A Teacher Friendly Way to Test Language Proficiency. *English Teaching Forum*, 31-1. 2 April 2012, <<http://exchanges.state.gov/englishteaching/forum/archives/1993/docs/93-31-1-g.pdf>>
- Klein-Braley, C. 1985. A Cloze-up on the C-Test: A Study in the Construct Validation of Authentic Tests. *Language Testing* 2, 76-104.
- Krashen, S. (1981). *Second Language Acquisition and Second Language Learning*. Oxford:

Pergamon Press.

- Larson-Hall, J. (2008). Weighing the Benefits of Studying a Foreign Language at a Younger Starting Age in a Minimal Input Situation. *Second Language Research* 24-1, 35–63.
- May, S. & R. Hill (2005). Māori-medium Education: Current Issues and Challenges. *International Journal of Bilingual Education and Bilingualism* 8(5), 377-403. 26 July 2012. <<http://www.tandfonline.com/doi/abs/10.1080/13670050508668621>>
- Miralpeix, I. (2006). Age and Vocabulary Acquisition in English as a Foreign Language (EFL). In C. Munoz (ed), *Age and Rate of Foreign Language Learning*. Clevedon: Cromwell Press.
- Muñoz, C. (2008). Symmetries and Asymmetries of Age Effects in Naturalistic and Instructed L2 Learning. *Applied Linguistics* 29-4, 578-596.
- Naber, R. (2011). *Hoe Vroeger, Hoe Beter?: Een Onderzoek naar de Effecten van Vroeg Vreemdetalenonderwijs*. Thesis, Rijksuniversiteit Groningen. Europees Platform. 1 April 2012, <<http://www.europeesplatform.nl/sf.mcgi?3871&cat=769>>
- Oller, J. & N. Nagato (1974). The Long Term Effects of FLES: An Experiment. *The Modern Language Journal* 58, 15-19.
- Paradis, J. (2007). Second Language Acquisition in Childhood. In E. Hoff and M. Shatz (eds), *Handbook of Language Development*. Oxford: Blackwell.
- Saville-Troike, M. (2006). *Introducing Second Language Acquisition*. New York: Cambridge University Press.
- Seidlhofer, B. (1996). “It is an Undulating Feeling...” The Importance of Being a Non-native Teacher of English. *Views* 5-1&2, 63-80.
- Singleton, D. & L. Ryan (2004). *Language Acquisition: The Age Factor*. Clevedon: Multilingual Matters Ltd.
- Sociaal en Cultureel Planbureau (2012). *Statusscores*. 8 June 2012, <http://www.scp.nl/Organisatie/Onderzoeksgroepen/Wonen_Leefbaarheid_Veiligheid/Lopend_onderzoek_van_WLV/Statusscores>
- Swain, M. (1974). French Immersion Programs Across Canada: Research Findings. *The Canadian Modern Language Review* 31, 117-129.
- Thijs, A., Tuin, D. & B. Trimbos (2011). *Engels in het Basisonderwijs: Verkenning van de Stand van Zaken*. SLO - Nationaal Expertisecentrum Leerplanontwikkeling. 2 May 2012, <www.slo.nl>
- Verspoor, M., Schuitemaker-King, J., van Rein, E., de Bot, K. & P. Edelenbos (2010). *Tweetalig Onderwijs: Vormgeving en Prestaties*. Rijksuniversiteit Groningen. 26 July 2012 <<http://www.europeesplatform.nl/sf.mcgi?2682>>

Waninge, F. (2011). *The Curious Case of Early English: Teacher Beliefs about English in Primary Education in the Netherlands*. Thesis, University of Groningen. 1 April 2012
<http://scripties.let.eldoc.ub.rug.nl/FILES/root/Master/ResearchMaster/Linguistics/2011/WaningeF./MA_1554328_F_Waninge.pdf>