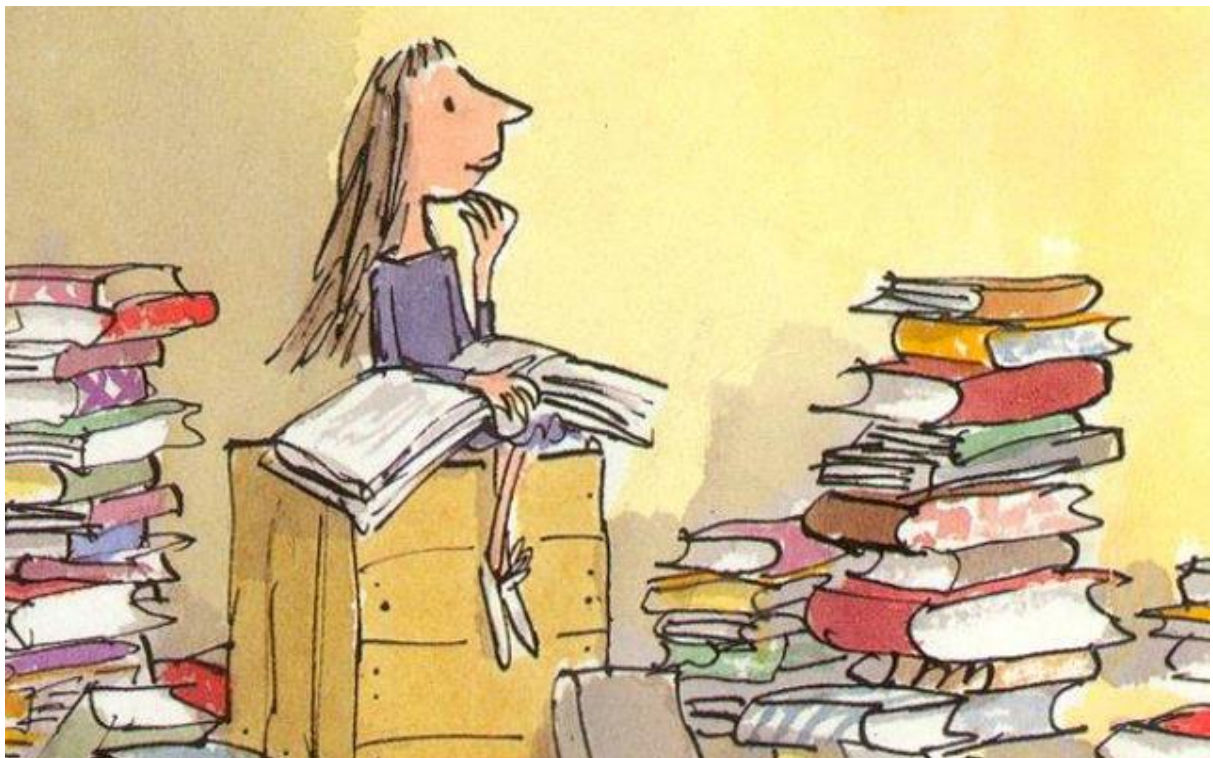




Utrecht University

**Increased Linguistic Competence as a Consequence of Extensive Reading (ER) – An Empirical Study into the Effect of ER on L2 Writing Proficiency**



Patrick Thiecke – Master's Thesis  
English Language and Culture: Education and Communication  
December 2018

Master's Thesis

English Language and Culture: Education and Communication

Utrecht University

December, 2018

Author: Patrick Thiecke (3906892)



Thesis Supervisor: Prof dr. Rick de Graaff

Second Reader: Dr. Elena Tribushinina

## **Preface**

I would be honored to take this opportunity to express my sincerest gratitude to all the individuals who, in one way or another, assisted me in the process of conceiving of, writing and finalizing my Master's Thesis. First and foremost, I would like to thank both Dr. Elena Tribushinina and Dr. Rick de Graaff for their continued support and guidance—without their help, I would not have been able to start this project or finish it appropriately. I would also like to thank Sianne Wong, Ashley Bogaski and Jake Honig for agreeing to act as correctors in order to create valid data for me to use in drawing conclusions. I would be remiss not to mention Martijn Govers, to whom I am eternally grateful for his undeniable role in helping me carry out the statistical analyses that underpin this research. Without his guidance, I would have struggled to understand how it all works.

I would also like to thank my parents for their help in keeping me grounded and focused on my project when I was going through a difficult time.

Patrick Thiecke

December, 2018

## Abstract

Extensive reading (ER) has been identified as an effective tool in teaching vocabulary and improving students' spelling habits. Despite this well-documented finding, the effect that ER has on writing proficiency—which is here seen as a complex of three interconnected variables, namely: accuracy, complexity and coherence—has not yet been investigated quite as rigorously. In particular, comparative research between, on the one hand, extensive reading and, on the other hand, traditional reading—which is here defined as standardized text comprehension activities as presented in textbooks for learners in educational settings—has so far remained elusive. In other words, the question of which method yields the greatest results with regard to writing proficiency as well as motivation to learn an L2 presents a promising field of endeavor. Although extrapolation of results from previous research can tentatively identify explicit practice—i.e. traditional reading—as the greater facilitator of learning, results from other research point in the exact opposite direction. For this reason, an extensive reading program was set up to contrast with a traditional reading program in the context of two VWO 5 classes in the Netherlands ( $N = 35$ ;  $ER = 17$ ;  $TR = 18$ ) with a focus on measuring growth in the participants' writing proficiency. Participants were asked to complete two writing assignments (cf. pre-test & post-test), one grammar test to account for inter-group comparability and two questionnaires to control for factors ranging from (prior) L2 exposure in various contexts, to various aspects of motivation.

*Key words: extensive reading, exercise-based reading, writing proficiency, linguistic competence, accuracy, complexity, coherence in writing*

## Table of Contents

1. Introduction	6
2. Theoretical Framework	8
2.1. Defining “extensive reading” as a concept and as a didactic tool	8
2.2. Known effects of extensive reading on learning in various contexts	8
2.3. Defining “writing proficiency”	10
2.4. Second Language Acquisition (SLA)	11
2.5. Motivation in the process of learning	11
3. Research Questions and Hypotheses	13
4. Method	14
4.1. Participants & intervention	14
4.2. Instruments	15
4.3. Data analysis	16
5. Results	19
Introduction	19
Writing Proficiency	20
1. Accuracy	20
2. Complexity	22
3. Coherence	23
Motivation	25
6. Discussion	26
7. Conclusion	30
Limitations of the current research	30
Recommendations for educators	31
References	32
Appendix A – Language Profile Questionnaire	39
Appendix B – Language Learning Motivation Questionnaire	40
Appendix C – Grammar Test	43
Appendix D – Writing Test A	47
Appendix E – Writing Test B	48
Verklaring kennisneming regels m.b.t. plagiaat	49

## 1. Introduction

In the Dutch educational system, English is considered an obligatory subject (DU: *kernvak*) which all students are required to take regardless of their educational level or choice of program. Students' performances are assessed according to four skills: speaking, writing, listening and reading (cf. Council of Europe, 2001); however, recent comparative studies have demonstrated that students' reading comprehension is underdeveloped (Abitzsch et al., 2017, p. 4; Reiss et al, 2016, p. 266). Therefore, it stands to reason that educators are in need of tools and techniques to effectively teach these skills, in particular reading. Moreover, there has recently been a push from the Dutch government to implement more differentiation-based teaching in the classroom so as to benefit individual students at their own level of learning (cf. Van de Pol et al., 2010).

However, with only 2 or 3 lessons a week per group taught and an enormous workload (Huyghebaert et al., 2018; Easthope & Easthope, 2000), educators often find that they are short on time to teach the curriculum effectively to all students. This reality makes research into interventions that simultaneously promote both student-aimed differentiation and two of the skills mentioned above pivotal. Additionally, it has been observed that young people in the Netherlands are less motivated to read due to less autonomy in the selection of reading materials at school (Huysmans, De Haan & Van den Broek, 2004; Verboord, 2006; cf. Abitzsch et al., 2017), which exposes them to less English and therefore fewer opportunities to learn and acquire the L2 in a stimulating environment.

In this regard, extensive reading might offer a solution as it allows students to select their own reading material, which challenges them at an appropriate level and gives them autonomy. However, little research to date exists to conclusively argue that such a reading program indeed leads to learning gains that are transferable to other domains, such as writing, nor that it may promote motivation to study the L2. In other words, the didactic value of extensive reading is doubtful if writing proficiency is tied to it, even though its effectiveness in improving learners' vocabulary and spelling habits has been thoroughly documented (Coady, 1997; Nagy, Herman & Anderson, 1985; Nation & Coady, 1988; Thornbury, 2002; Pellicher-Sánchez, 2016; Horst, 2005; Ellis, 2002), especially if a 98% word coverage is present (Hu & Nation, 2000). Such a high word coverage—i.e. the percentage of lexical items already familiar to the learner—is required according to Hu & Nation (2000) because percentages lower than 98% do not free up the learner to take in new information, and instead may cause cognitive overload from trying to interpret a large number of alien lexemes. This percentage may further be linked to Abitzsch et al.'s (2017) claim “dass Spracherwerb am besten gelingt, wenn das Lesen relativ wenig Mühe macht<sup>1</sup>” (p. 5). In other words, learning is best achieved if the learner already possesses an adequate framework upon which to expand.

The current research therefore focused specifically on differentiated reading, i.e. *extensive reading* (Abitzsch et al., 2017, p. 5), in relation to traditional reading, i.e. text comprehension activities with preselected, edited texts from a textbook, in order to measure the relative effect of both programs on writing proficiency. There is reason to assume that a positive correlation may be found because the copious texts, which students read in the extensive reading program, may subconsciously function as models and therefore facilitate the implicit development of the students' writing proficiency. The intervention took place in

---

<sup>1</sup> “*Language acquisition is best achieved when the act of reading costs relatively little effort*” (Translation courtesy of the author of this paper).

the context of two VWO 5 classes in the Netherlands ( $N = 35$ ;  $ER = 17$ ;  $TR = 18$ ) for a duration of 10 weeks ( $\approx 500$  minutes).

This research will hopefully shed light on the usability of extensive reading as a didactic tool for two purposes: (1) to promote differentiation in the sense that extensive reading allows for individualized reading at the level most suited for the individual learner; (2) to simultaneously engage students in various skills, i.e. reading and writing, and could therefore inform didactics and become a part of educators' repertoire. The relevance of this research therefore lies in the widening of scientific knowledge into the didactic value of ER in a previously unexplored area, and the possible widening of educators' repertoire by adding to it a scientifically underpinned method of differentiating that also promotes improved writing skills.

## 2. Theoretical Framework

### 2.1. Defining “extensive reading” as a concept and as a didactic tool

Various mechanisms are at play when a person is involved in the act of reading, which may manifest itself in various ways (cf. Smith, 2004). Of all these ways, extensive reading specifically is often conceived of as a method of reading that aims to promote “reading in quantity (...) to gain a general understanding of what is read. It is intended to develop good reading habits, to build up knowledge of vocabulary and structure, and to encourage a liking for reading” (Richards & Schmidt, 2002, p. 193-194). Grabe & Stoller (2002) add to this definition that the material read by the individual should be “[well] within their linguistic competence” (p. 259), in reference to Krashen’s (1982) *input hypothesis* which postulates that linguistic input must be comprehensible and decodable in order for learning to take place.

Extensive reading has, moreover, been typified as “pedagogically efficient” (Huckin & Coady, 1999, p. 182) because it simultaneously combines reading and vocabulary acquisition. In addition to this, it may be suggested that extensive reading could function as an effective tool to engage the learner and increase motivation (Dörnyei & Csizér, 1998) due to its insistence on differentiated reading relative to the learner’s level of competence in the learning process (Van de Pol et al., 2010), which removes the possibility of the task being perceived as insurmountable or cognitively taxing (Wang, 2013, p. 129; Wood & Middleton, 1975).

Day & Bamford (2002) devised ten basic tenets specifically to lay out ground rules for extensive reading to be employed in an educational, didactic setting such as a classroom. The tenets pay reference to the definitions given above while operationalizing the role played by the teacher.

1. The reading material is easy.
2. A variety of reading material on a wide range of topics must be available.
3. Learners choose what they want to read.
4. Learners read as much as possible.
5. The purpose of reading is related to pleasure, information and general understanding.
6. Reading is its own reward.
7. Reading speed is usually faster than slower.
8. Reading is individual and silent.
9. Teachers orient and guide their students.
10. The teacher is a role model of a reader.

(Day & Bamford, 2002, p. 137-140)

When implementing an extensive reading program, the teacher should be available to answer students’ questions and positively impact their sense of competence (Deci & Ryan, 2000; Zimmerman, 2008). This is because a significant amount of effort is involved in gaining modest improvement concerning L2 reading (Ngeow, 1998; Huang, 2006).

### 2.2. Known effects of extensive reading on learning in various contexts

In line with earlier observations by Huckin & Coady (1999), extensive reading is an effective promotor of vocabulary acquisition. For example, research has demonstrated that extensive reading promotes and increases sight vocabulary (Coady, 1997; Nagy, Herman & Anderson,



1985; Nation & Coady, 1988), while also promoting opportunities to encounter new words in their own relative context of use (Thornbury, 2002). In part, the increase in vocabulary knowledge may be accredited to the fact that learners will likely encounter particular words multiple times, which is congruent with research suggesting that both incidental and intentional learning of vocabulary is facilitated and strengthened by the amount of exposure resulting from a positive correlation of exposure to word retention (Pellicher-Sánchez, 2016; Horst, 2005; Ellis, 2002).

A case study performed by Maria & Norbert (2006) suggests that extensive reading, under the right circumstances, can lead to a “pickup rate of about 1 of every 1.5 words tested” (p. 1), which is a degree “not demonstrated before” (p. 8), while also promoting better spelling. However, the authors also signal potential issues with extensive reading in relation to incidental vocabulary acquisition, namely: (1) reading for meaning—which is a principle of extensive reading—has not been conclusively linked to actual vocabulary acquisition (Huckin & Coady, 1999); (2) the activity of inferring a word’s meaning through contextual clues may not in all cases lead to meaning retention (Huckin & Coady, 1999; Nation & Coady, 1988; Parry, 1993); (3) students may choose to ignore new words if uncovering their meaning is not essential to decoding the message in the text (Zahar, Cobb & Spada, 2001; Mondria & Wit-de-Boer, 1991). Furthermore, as research by Hu & Nation (2000) suggests, a 98% word coverage is also needed so as to not discourage students who might otherwise face too many alien concepts they fail to deconstruct. This means that learners must be able to recognize and understand at least 98% of the lexical items present in a text.

The rate at which linguistic abilities develop as a result of extensive reading varies according to Yamashita (2008). For example, “progress in micro-level linguistic ability such as (...) morphosyntax” may only become observable after a certain amount of time whereas other skills, such as general reading skills, are improved more rapidly (Yamashita, 2008, p. 669). Yamashita (2008) also points to research by Hafiz & Tudor (1989) where a “marked improvement (...) in terms of writing skills” (p. 4) is found for students of ESL in the UK subjected to an extensive reading program using graded readers.

Similar results were found in Hernández’ (2011) study, with specific reference to the verbal use of discourse markers by Spanish learners of English who had been subjected to a program in which the effect of input flood was contrasted with the combined effect of input flood and explicit instruction. In fact, Hernández (2011) concludes that “the combined effect of [explicit instruction] and [input flood]” was commensurate with the effect of input flood alone, even in the delayed post-test. This conclusion is in direct opposition to a multitude of other studies which attribute a significantly greater effect to explicit instruction (Ellis, 1993; Alanen, 1995; DeKeyser, 1995, 1997; Robinson, 1996, 1997; de Graaff, 1997; Yoshimi, 2001; Rosa & Leow, 2004; Hernández, 2008; de la Fuente, 2009). Hernández’ (2011) study may have focused on speaking proficiency, but the use of input flood makes it similar to extensive reading as both programs emphasize a large consumption of L2 input, albeit through different media (audiovisual versus written); therefore, the research is similar, but not identical, making it unsound to extrapolate the findings directly.

Although the two studies mentioned above purport positive correlations, their results can only be tentatively extrapolated to this research (see also *Method*). Hafiz & Tudor (1989), for example, used graded readers whereas the participants in this research were given free choice. Hernández (2011) focused on speaking and used input flood, which involves the educator supplying the learners with myriad examples of a particular (grammatical) item so that the learners may familiarize themselves with its uses and the context in which it appears (cf. Warford & White, 2012). Input flood is therefore to some extent comparable to extensive reading, but not identical. It stands to reason that changes in the genetic makeup of either the population—e.g. years of instruction, location and motivation to learn English—or the design

of the intervention (e.g. free choice versus graded readers) can have an impact on the effectiveness of extensive reading in relation to (increased) writing proficiency. Therefore, it would seem worthwhile to investigate whether the positive effects found in Hafiz & Tudor (1989) and Hernández (2011) can be replicated under divergent circumstances and with other material.

### 2.3. Defining “writing proficiency”

A working definition of *writing proficiency* can often be hard to give. However, there are certain characteristics of well-written products—which are often intuitively conceived of by educators—that may point toward a possible framework of reference (see Figure 1).

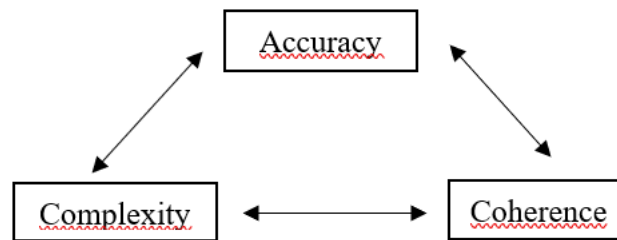


Figure 1: Visualized representation of writing proficiency, as defined in this paper (figure created by this author). Note that each node is interconnected, as there is presumed to be no hierarchy. This is because a well-written product, which evidences good writing proficiency, needs to perform adequately in all three areas.

Accuracy—i.e. the ability to correctly employ target language grammatical structures and spelling without L1 negative transfer—is often seen as an indicator of good language proficiency, and by extension good writing proficiency (cf. Williams, 1989; Raimes, 1983). Mastery of target language grammar and spelling in the written medium is a logical prerequisite in conveying relevant information and, therefore, it functions as one of three aspects that together demarcate writing proficiency.

Furthermore, Raimes (1983) suggests that complexity operates as a second aspect of writing proficiency in that the more complex a text is, the more reflective this is of the learner’s adeptness at using the L2 in writing. Complexity may be measured in a variety of ways; for example, (1) by expressing in percentages the number of correct simple, complex and compound sentences (Dülger, 2007); (2) by processing the text on the Flesch-Kincaid Readability Scale, which “assesses readability on the basis of the average number of syllables per word and the average number of words per sentence” expressed in a grade level (Paasche-Orlow, Taylor & Brancati, 2003, p. 722); (3) by calculating the token/type ratio; (4) by calculating the lexical density, which is “defined as the percentage of lexical words in a text” (Laufer & Nation, 1995, p. 309). Neither one method alone suffices to conclusively express the complexity of any given text, as each only provides a fraction of the insight required to make well-informed judgments.

Finally, the learner’s ability to promote coherence through so-called discourse markers is seen as the third aspect of writing proficiency (Corbett, 1987). Discourse markers have been found to positively “influence (...) text comprehension” (Sanders, Land & Mulder, 2007, p. 229), and their importance for writing proficiency has also been underlined by Hernández (2011) who states that “a speaker uses [them] to sequence and structure ideas and

information in a paragraph-length discourse in order to produce a cohesive and coherent narration, which is a critical feature of advanced language competence” (p. 164).

Despite the consensus regarding their utility, discourse markers have seen a number of potential, sometimes slightly conflicting definitions (Fraser, 1998, p. 302; Schiffrin, 1987, p. 41; Redeker, 1991, p. 1168). Furthermore, a distinction must be made between those discourse markers that operate at the micro-level, i.e. to link two adjacent sentences or two ideas within the same sentence, and those that operate at the macro-level, i.e. to link two paragraphs or to reveal the underlying structure of the text (cf. Ben-Anath, 2006). Often, various categories are employed to compartmentalize the discourse markers available in a given language, ranging from *additive* to *adversative* and *summarizing* (Feng, 2010, p. 300; Dülger, 2007, p. 261-262). The choice of how to categorize certain lexical expressions and the number of categories to be distinguished is subject to a certain degree of arbitrariness relative to contextual needs.

#### 2.4. *Second Language Acquisition (SLA)*

Learning, i.e. taking up new information and mapping it onto a net of pre-existing and emergent neural connections, is explicated and discussed in different theoretical frameworks—for instance, in the form of linguistic competence (White, 2003), an emergentist theory (Ellis, 1998) or a sociocultural theory (Lantolf & Thorne, 2006). Therefore, a large number of divergent teaching interventions exist to induce learning. However, in this abundance of interventions, no single one has yet been identified as producing the greatest results (De Graaff & Housen, 2009).

Broadly speaking, three different schools exist concerning the acquisition of L2 structure and grammar: (1) one school emphasizes the necessity of having the learning process resemble L1 acquisition, where exposure to and familiarity with the L2 in a communicative setting is seen as the optimal learning opportunity (Krashen, 1981; Lyster, 2007; Prabhu, 1987); (2) the other school instead emphasizes explicit instruction and form-based instruction (*FFI*; cf. Loewen, 2011; Norris & Ortega, 2000); (3) the last school emphasizes a combination of both (Ellis, 1995; Spada, 1997), which can be retraced in hypotheses about optimal retention of information such as the *counterbalance hypothesis* (Lyster & Mori, 2006). The principles advocated by these schools are rooted in divergent beliefs about the optimal way of acquiring a foreign language, with each influencing and shaping the didactic choices an educator makes differently.

How the knowledge attained is stored and called upon is a matter of much scholarly debate. Although consensus exists as to the presence of implicit and explicit knowledge, “whether [these] are to be viewed as distinct and dichotomous or intertwined and continuous” (Ellis et al., 2009, p. 335) remains a controversial question (cf. Dienes & Perner, 1999). It is also generally acknowledged that different speech acts draw on different types of knowledge (Ellis et al., 2009, p. 335; Anderson et al., 1997).

#### 2.5. *Motivation in the process of learning*

It has often been theorized and demonstrated that motivation, in particular motivation obtained through autonomy (cf. Deci & Ryan, 2010), is a huge predictor in terms of learning gains (Dörnyei & Csizér, 1998; Benson, 2007), which may be linked to Krashen’s (1981) affective filter which regulates how much learning takes place. This affective filter is tied to a multitude of factors—including motivation (Du, 2009). Consequently, the more motivated a person is to learn about a particular subject, the likelier it will be that learning takes place.

However, motivation can come from various sources (*locus of causality*), such as internal (*intrinsic*) or external (*extrinsic*), to name but a handful (Woolfolk et al., 2013). Other perspectives have been offered as well, such as the humanist view which postulates that motivation comes from opportunities to self-actualize, i.e. assuming autonomy in the learning process (Deci & Ryan, 2000).

Moreover, an individual's motivation to complete a task—thereby inducing learning—is related to their sense of self-efficacy, i.e. the set of beliefs about the individual's own competence and ability to perform the task at hand in a satisfactory manner (cf. Dweck & Leggett, 1998; Woolfolk et al., 2013, p. 454; Hsieh & Schallert, 2008, p. 523). Greater learning gains are often observed in individuals who are intrinsically motivated to perform well rather than in individuals who are not motivated (Vansteenkiste et al., 2004).

An individual's level of motivation—and even the source thereof—may be influenced by factors that are both internal and external; in fact, a person's motivation may even shift from being external to being internal due to circumstances. In this regard, rewards, i.e. “[an] attractive object or event supplied as a consequence of a particular behavior” (Woolfolk et al., 2013, p. 432), and incentives, i.e. “an object or event that encourages or discourages behavior” (p. 432), may be provided to achieve positive reinforcement or punishment, which impacts a person's level of motivation..

### 3. Research Questions and Hypotheses

In the theoretical discussion above, it was established that extensive reading as a didactic tool is positively correlated with vocabulary learning, spelling and improved reading skills. Yet, the question remains whether it is also positively correlated with improved writing proficiency in the context of the Dutch regular educational system. More specifically, it is not currently known how students in VWO 5 respond to extensive reading in terms of writing proficiency. Writing proficiency, as stated above, is defined as the sum of three separate aspects (i.e. accuracy, complexity and coherence). Therefore, the research question becomes the following:

- *Research question: How does participation in an extensive reading program relate to writing proficiency?*
  - Sub-question 1: Does extensive reading improve the accuracy of the students' written L2 production to a greater extent than TR?
  - Sub-question 2: Does extensive reading improve the complexity of the students' written L2 production to a greater extent than TR?
  - Sub-question 3: Does extensive reading improve the coherence, i.e. structure, of the students' written L2 production at the micro- and/or macro-level to a greater extent than TR?
  - Sub-question 4: Does extensive reading impact the students' motivation to learn English as an L2 to a greater extent TR?

The extensive reading group (see *Method*) is subjected to large quantities of English texts, which may consequently function as models from which to implicitly or explicitly draw key information about writing competency. Participants in the ER program may then internalize this information so as to later retrieve it during the actual writing process, improving their accuracy, complexity and coherence by mimicking what they have encountered before (cf. Lyster, 2007; Prabhu, 1987; Ellis, 1995; Spada, 1997).

Furthermore, findings from Leung & Williams' (2014) experiment demonstrate that typologically similar languages, such as English and Dutch, are likely to facilitate the L2 learning process. Consequently, students in both group should theoretically benefit from already speaking a Germanic language by either implicitly or explicitly acquiring knowledge of grammar, discourse markers and other key features to improve their accuracy, complexity and coherence. Therefore, it may be hypothesized that, under the right circumstances, both groups may improve their writing proficiency; however, learning gains should be greater in the ER group as compared to the TR group (cf. Hafiz & Tudor, 1989; Hernández, 2008).

In terms of motivation, since extensive reading promotes the students' own interests and autonomy in choosing which material to read, it is reasonable to assume that their motivation will increase since they become, in a sense, owners of their learning (Deci & Ryan, 2000; Zimmerman, 2008; Dörnyei & Csizér, 1998). Furthermore, since the reading material is picked relative to the students' own level of proficiency, their sense of competence is positively influenced, which should logically increase the students' motivation. Since this level of autonomy is not present in the TR group, who are instead exposed to preselected texts, the motivation in the ER group should theoretically be greater than in the TR group.

## 4. Method

### 4.1. Participants & intervention

For this research, two groups of VWO 5 students ( $N = 35$ ) at a regular school in Utrecht participated. The groups were split up like so: the intervention group ( $N = 17$ ;  $N_{female} = 9$ ,  $N_{male} = 8$ ), which did the extensive reading program, and the comparison group ( $N = 18$ ;  $N_{female} = 6$ ,  $N_{male} = 12$ ), which did the traditional reading program. The intervention group had a total of 3 participants who claimed to be fluent in English and Dutch, while the other 14 participants reported themselves to be fluent in Dutch only. In the comparison group, 4 participants claimed to be fluent in English and Dutch, while the other 14 claimed fluency in Dutch only. One participant in the comparison group reported their ethnicity to be Taiwanese, although this participant did not claim to speak the associated language. The number of participants is quite low due to practical reasons as the groups this research had to opportunity to work with were small.

The students in the comparison group engaged in traditional reading, i.e. they read standardized texts from a textbook and occasionally from past VWO exams. It is therefore important to note here that these students have had considerably more experience with and exposure to comprehension questions. Examples of such questions include, but are not necessarily limited to: (1) deducing the author's intentions; (2) expressing connections between paragraphs; (3) gap exercises where a word, such as a connective or fragment, has been omitted; (4) true/false questions where close reading is required; (5) multiple choice questions. This type of reading promotes close reading and a structured analysis of the content. Therefore, it may also be referred to as *structured reading*.

The students in the intervention group were engaged in extensive reading activities for one hour a week for a total of ten weeks ( $\approx 500$  minutes). These students were allowed to read whatever they wanted (e.g. books, novels, newspaper articles, online blogs, etc.) as long as five conditions were met: (1) the language of the material must be modern English; (2) the level of English must be well within their grasp so they are not overwhelmed by it, but not underwhelmed either; (3) the reading must be done silently and individually; (4) the students will not be given comprehension questions such as 'what do ll. 5-8 mean'; (5) the students must keep track of what they have been reading and share this with the teacher (cf. Day & Bamford, 2002). This type of reading may be referred to as *unstructured reading*, as students were not necessarily asked to process the content in any specific way.

To ensure inter-group comparability, near equal distributions of time spent on various activities in class were sought. For this reason, lessons were pre-planned to make sure that both groups were engaged in similar activities (e.g. listening), except for extensive reading versus traditional reading. Figure 2 below illustrates for both groups what their lessons were structured around during the period of the intervention. The percentages were calculated using the pre-planned activities divided by the total number of lessons each group had, and were adjusted if one group required a different lesson than what had been originally planned. This, however, happened minimally. It may be observed, for instance, that the intervention group spent slightly more time learning about literature because some material, which had been covered in the comparison group, had not yet been covered in the intervention group (see Figure 2). In principle, both groups received the exact same instructions and did the exact same activities in class, except for extensive reading versus traditional reading, so as to exclude the possibility of differences therein affecting the research objective.

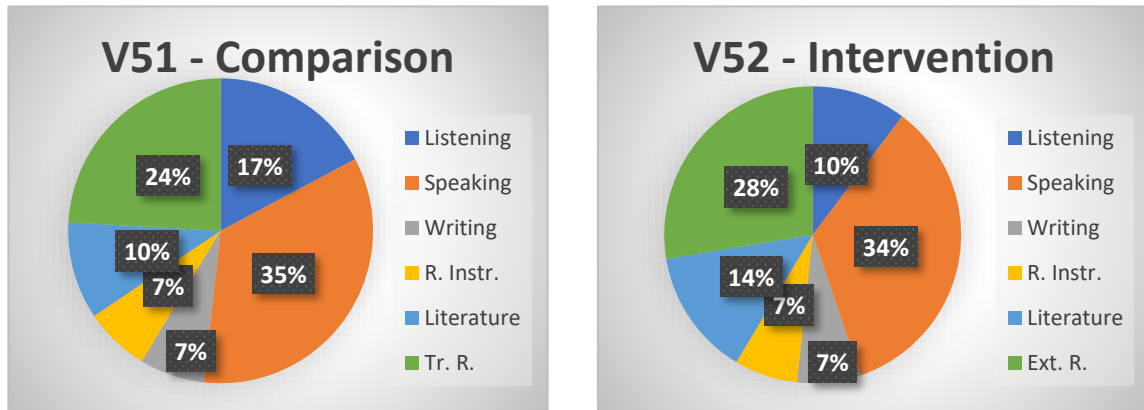


Figure 2: Activities done by V51 (comparison group) and V52 (intervention group), expressed in terms of percentages. In total, V51 has spent 31% on traditional reading and reading instructions, while V52 has spent 35% on extensive reading and reading instructions.

The activities include: (1) listening to audio fragments; (2) practicing oral skills; (3) practicing writing; (4) reading instructions: (5) traditional reading or extensive reading; (6) literary instruction. Under (4), students were given explicit training in familiarizing themselves with reading strategies and applying them—both groups received the same instructions here and did the same exercises. Under (6), students were taught about specific literary works, literary devices and how English literature in general came to develop itself over the course of the centuries.

Comparability between both groups is further promoted by the fact that the average grade per class is very similar—the comparison group, for example, has an average grade of 6.5 while the intervention group has an average grade of 6.4. This covers vocabulary, writing, listening and speaking. Both groups also scored comparatively well on the Grammar Test (see *Instruments* for a more in-depth explanation):  $M = 0.91$ ;  $Sd = 0.4$  for the comparison group and  $M = 0.87$ ;  $Sd = 0.8$  for the intervention group. An ANOVA test revealed no significant difference between groups:  $F\text{-critical} = 2.23$ ,  $f = .85$ ,  $p = .65$ . Moreover, both groups have had an average of 7 years of instruction of English.

#### 4.2. Instruments

Several instruments were used to measure various aspects. To determine that the groups were comparable in terms of prior knowledge of grammar, a Cambridge grammar placement test (see appendix C) was used consisting of cloze exercises where students are expected to choose the correct alternative from a list of four, with questions varying in degree of complexity from beginner to advanced. This was done prior to the start of the intervention. Neither group received any feedback on their performance so as to not inadvertently influence the learning process. No more than 30 minutes were allowed in class to take the test.

Furthermore, to control for variation causing factors such as previous exposure to English and motivation, two questionnaires were used (see Appendices A & B) to promote the generalizability and validity of the results (Stokking, 2016). Controlling for these factors is crucial because it may be assumed that an individual with low motivation will perhaps benefit less than a person with high motivation. To control for previous exposure, students in both groups were asked to fill in a questionnaire concerning their language profile (e.g. native language(s), exposure to L2 in various contexts outside of school, etc.). This questionnaire is

handmade by the author of this paper and requires the students to give estimates of time spent engaging with English per day and per week for objective measurement. To control for motivation, a more exhaustive questionnaire measuring motivation to learn in various contexts on a scale of 1-5 was given to students both before and after the intervention to spot potential changes. This questionnaire designed by Papi & Teimouri (2014) is standardized and validated, with Cronbach's alpha values ranging from .67 to .83 for the various items.

Papi & Teimouri's (2014) questionnaire makes it possible to measure the students' motivation in various contexts such as *instrumentality* (cf. "Can I use English to achieve certain goals?" or "Can I use English to prevent embarrassment?"), *ought-to* (cf. "I want to learn English because I am expected to speak it well"), *future plans* (cf. "I want to learn English because I can imagine myself working a job where English is needed"), *anxiety* (cf. "I want to learn English because I get anxious when I make mistakes"), *learning experience* (cf. "I want to learn English because I like the atmosphere in my English classes") and *attitudes to L2 community* (cf. "I want to learn English because I like the people that speak it"). These different types of motivation have varying degrees of influence on the individual student's learning process.

To measure the students' writing proficiency, a pre-test writing test and a post-test writing test were given (see Appendices D & E). All students received the same test at the same time so there would be no cross-contamination. Students were asked to write a short essay about a topic familiar to them—this was done to prevent the students from performing poorly because the topic is either too complicated or foreign. The students were asked, in 200 to 250 words, to give their opinion on a statement such as "homework should be banned" or "students should be required to learn two languages in school in addition to English and Dutch". The students were therefore required to first interpret the statement for themselves, to critique it in accordance with their opinion and then organize their thoughts to produce a written product, which are cognitive processes covering the lower order and higher order thinking skills on Bloom's revised taxonomy (cf. Krathwohl, 2002). Conceptual similarity between both tests is guaranteed because students are required to perform the same steps both times while dealing with a topic that is familiar to them due to being related to school, thus producing valid results. Furthermore, short essays often promote the use of discourse markers, although students from both groups were not explicitly instructed to use them.

The results obtained from the writing tests give insight into the quality of the product, not the quality of the writing process itself (Seow, 2002; Brown, 2001). Although this is not necessarily a flaw, the decision to focus on the product and not the process limits the scope of the research. However, in order to answer the research question, only the product is required.

#### 4.3. Data analysis

The written texts (both pre-test and post-test) were first graded holistically and given a mark of 1 to 5 (1 = beginner, 5 = advanced user), which is a process similar to Verspoor et al.'s (2012) scale system. Marks were given by a total of four correctors, three of whom are MA students while the other is a native speaker with a background in linguistics located in Canada, in order to combat researcher bias. The correctors were not told to which group the participants belonged, nor were they told whether they were dealing with the pre- or post-test. Therefore, the assessment took place blindly. Cronbach's alpha values were determined at  $\alpha = 0.78$  for the pre-test and at  $\alpha = 0.72$  for the post-test, which are acceptable values (cf. Tavakol & Dennick, 2011). The judgments were also taken together, after which averages



and standard deviations were calculated. Next, the texts were analyzed more thoroughly in the three areas discussed above (cf. [Figure 1](#); *accuracy, complexity, coherence*).

Complexity was expressed in percentages with regard to the number of simple, complex, and compound sentences<sup>2</sup> divided by the total number of sentences per text (cf. Dülger, 2007). The values were normalized by dividing the appropriate numbers by the word total in a given text, e.g. the total number of simple sentences divided by the total number of words. Moreover, online tools<sup>3</sup> were consulted to quickly process the morpheme complexity and lexical complexity to allow for the lexical density to be calculated. Lexical density is defined as the number of individually different, information-carrying lexical units divided by the total number of words in a given text to provide a measure of the proportion of lexical items, as opposed to function words such as determiners, prepositions, conjunctions, etc. Therefore, higher values indicate a more varied use of speech as well as fewer repetitions of a particular lexical item, which according to Didau (2013) can make a text more “concise and meaningful”, thereby increasing the quality of the text.

Additionally, these tools allowed for a text to be processed on the Flesch-Kincaid Readability Scale, with values ranging from 0.00 (=very difficult) to 100.00 (=very easy). The Flesch-Kincaid Readability Scale uses a formula that takes the average sentence length and the average number of syllables per word within the sentence or text into consideration in order to calculate the approximate reading level required to understand the content of the text (see [Figure 3](#)). These measurements were used to operationalize complexity in a variety of aspects.

$$206.835 - 1.015 \left( \frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left( \frac{\text{total syllables}}{\text{total words}} \right)$$

*Figure 3: The Flesch-Kincaid Readability formula (cf. Rudolf, 2016).*

Coherence was measured using Dülger’s (2007) technique of counting all instantiations of discourse markers (correct and incorrect) and calculating averages to measure potential growth. A decision was made to focus on discourse markers, as opposed to other elements which contribute to coherence such as parallelism and the non-excessive repetition of keywords, because the aspect of discourse markers was the easiest and least subjective to quantify. Therefore, even though discourse markers are an essential part of achieving coherence in writing, it must be pointed out that they alone do not constitute the full spectrum of coherence in writing.

Which discourse markers to focus on is a matter of selection. For this, a synthesis of Feng’s (2010) and Dülger’s (2007) categories was used for the analysis of the results. The categories include: (i) incorrect; (ii) micro-level; (iii) macro-level. A further subcategorization for micro-level discourse markers was designed in accordance with categories conceived by Feng (2010, p. 300) and Dülger (2007, p. 261-262). These categories include: (a) additive; (b) contrastive; (c) causal; (d) continuatives; (e) focusing; (f) exemplification; (g) referral to expectations; (h) generalizing; (i) summarizing. A distinction was made between connectives at the micro-level, i.e. connectives that link two adjacent sentences or two ideas within a single sentence, and connectives at the macro-level, i.e. connectives that link two paragraphs or reveal the underlying structure of the text at large (cf.

<sup>2</sup> In this paper, simple sentences are defined as (isolated) sentences containing only one main verb. Complex sentences contain one subordinate clause, whereas compound sentences may contain upwards of two subordinate clauses.

<sup>3</sup> <https://www.usingenglish.com/resources/text-statistics.php>

Ben-Anath, 2006). The emergent numbers were compared and cross-analyzed so as to measure potential growth. **Normalization was achieved by dividing the number of discourse markers, both at the micro and macro levels respectively, by the total number of words in a participant's text.**

An error analysis, inspired on Verspoor et al. (2012, p. 253) as well as Ellis' (2005, p. 155) examples of typical learner errors, was used to deal with accuracy. However, whereas Verspoor et al. identify certain mistakes as lexical mistakes, such mistakes were relegated to the area of grammar instead as they impacted the grammaticality of the sentence. To normalize the data obtained from the error analysis, the total number of mistakes were divided by the total number of words in the student's writing to generate a value between 0.00 and 1.00. The normalized error values can more accurately reflect how the participants from both groups performed with regards to accuracy.

Error type	Instantiations	Example
Grammar	<ol style="list-style-type: none"> <li>1. Wrong preposition</li> <li>2. Incorrect use of pronouns</li> <li>3. Incorrect use of singular/plural</li> <li>4. Incorrect word order</li> <li>5. Incorrect word form (cf. adverbs)</li> <li>6. Dutch constructions</li> <li>7. Wrong verb form</li> <li>8. Wrong verb use</li> <li>9. Wrong collocation / verb</li> <li>10. Wrong use of articles</li> <li>11. Use of L1</li> </ol>	<ol style="list-style-type: none"> <li>1. To be <i>on</i> school</li> <li>2. <i>It</i> are my new shoes</li> <li>3. A very cool <i>teachers</i></li> <li>4. I like it not</li> <li>5. I can see it <i>good</i></li> <li>6. <i>There</i> can be said that ...</li> <li>7. He <i>go</i> to school</li> <li>8. He <i>has gone</i> to school yesterday</li> <li>9. This will <i>follow</i> to worse results</li> <li>10. A assignment / as (<math>\emptyset</math>) student</li> <li>11. This is my <i>mening</i></li> </ol>

*Table 1: Error analysis model adapted from Verspoor et al. (2012) and Ellis (2005). Numbers 9, 10 and 11 were added during the actual analysis of the writing assignments when it was noticed that these categories constituted a recurrent mistake in the writings of various students, which could not be ignored.*

## 5. Results

### Introduction

Both groups were subjected to a baseline grammar test (see [Instruments](#)), which assesses students' ability to choose the correct alternative with regards to verb forms, pluralization, pronouns, subordinate clauses and tag questions. For each correct answer, 1 point was assigned while an incorrect answer would yield 0 instead. This purveys a set of data to analyze inter-group comparability (see Table 2). An ANOVA test revealed no significant difference between groups:  $F\text{-critical} = 2.23$ ,  $f = .85$ ,  $p = .65$ , which suggests that both groups are comparable to each other.

Group	Grammar Test			
	M	S <sub>d</sub>	S <sub>e</sub>	σ <sup>2</sup>
Comparison	0.906	0.043	0.010	0.002
Intervention	0.871	0.076	0.016	0.006

Table 2: Results from the grammar test (see Appendix C).

Moreover, as stated above, Verspoor et al.'s (2012) scale was used to determine the overall performance on both the pre-test and post-test for each individual student. Each text was holistically graded by a total of four correctors, three of whom are MA students currently studying to become teachers of English, while the other corrector is a native speaker from Canada with a background in linguistics.

After each corrector had provided a grade between 1 (=total beginner) and 5 (=advanced user), the individual average score was calculated for each student, resulting in the average scores presented in Table 3. Additionally, the differences in post-test and pre-test scores were calculated for each student, after which an ANOVA test was run to determine significance. For the comparison group, the estimator of the mean growth  $M_{\text{comp}}$  equals  $0.15 \pm 0.14$ . For the intervention group, the estimator of the mean  $M_{\text{int}}$  equals  $0.19 \pm 0.10$ . A mixed ANOVA test revealed no significant differences for the condition 'group' [ $F(1, 66) = 3.83$ ,  $p = 0.10$ ], 'time' [ $F(1, 66) = 1.88$ ,  $p = 0.18$ ] or 'interaction' [ $F(1, 66) = 0.04$ ,  $p = 0.85$ ], which suggests that neither group grew more than the other as a result of the program.

Group		Grade						
		M	S <sub>d</sub>	S <sub>e</sub>	σ <sup>2</sup>	$P_{\text{group}}$	$P_{\text{time}}$	$P_{\text{interaction}}$
Comparison	Pre-test	3.50	0.58	0.15	0.35	0.10	0.18	0.85
	Post-test	3.65	0.31	0.08	0.10			
Intervention	Pre-test	3.25	0.42	0.11	0.19			
	Post-test	3.44	0.47	0.13	0.24			

Table 3: Average performance on the pre-test and post-test expressed on a scale of 1 to 5 (cf. Verspoor et al., 2012) where 1 equals 'total beginner' and 5 equals 'advanced user'.

The similar results might be a consequence of the fact that both groups enjoy similar amounts of exposure to English outside of the classroom environment except for listening (see Table 4), expressed in the four skills often recognized by the CEFR. This finding became apparent from the questionnaire the participants were required to fill in before the start of the program (see Table 4); however, the results obtained here are the result of self-reporting, which carries with it a number of inherent characteristics which could problematize the

findings (Stokking, 2016). The only major difference, it seems, is related to listening (e.g. listening to music, podcasts, movies, etc.) as the intervention group self-reports far greater exposure than the comparison group.

Because the differences in the amount of exposure are so striking, it was decided to use this set of data as a covariate in analyzing the various aspects of writing proficiency as well as motivation. The covariance analyses are reported on in the sections to follow, expressed in correlation coefficients  $\rho$  where one consistent factor is the normalized amount of exposure in terms of listening. Values higher than -0.5 or 0.5 are subjected to a linear regression analysis, as such values indicate a strong relationship and may therefore provide interesting insights while other values are considered weak. To normalize the participants' individual exposure, the amount of exposure reported by the participant was divided by the total number of hours in one week. In this way, the correlation between amount of exposure and writing proficiency may be better explored.

	Listening	Speaking	Writing	Reading
Intervention Group	M = 17.0; S <sub>d</sub> = 12.5	M = 1.0; S <sub>d</sub> = 1.2	M = 1.0; S <sub>d</sub> = 1.1	M = 3.0; S <sub>d</sub> = 5.3
Comparison Group	M = 8.6; S <sub>d</sub> = 7.2	M = 1.2; S <sub>d</sub> = 1.4	M = 0.7; S <sub>d</sub> = 0.9	M = 2.2; S <sub>d</sub> = 3.1

Table 4: Weekly distributions of exposure to English outside of school in various contexts. Results are obtained from a questionnaire students filled in before the start of the research.

### Writing Proficiency

As stated above, writing proficiency is seen as a complex interconnected mechanism consisting of three elements: (1) accuracy; (2) complexity; (3) coherence. This part of the results section will report on the findings specific to these three elements.

#### 1. Accuracy

Accuracy was measured by identifying the number of mistakes made by the participants according to the error analysis model adapted from Verspoor et al. (2012) and Ellis (2005) (see Table 1). Each mistake was categorized and counted, after which normalized errors were calculated by dividing the total number of mistakes per participant by the total number of words in the participant's written text (see Table 5). A mixed ANOVA test revealed no significant difference for the condition 'group' [ $F(1, 66) = 1.45, p = 0.23$ ], 'time' [ $F(1, 66) = 1.11, p = 0.30$ ] or 'interaction' [ $F(1, 66) = 0.02, p = 0.88$ ], which means that no improvement in terms of accuracy is observed that is greater in one group than in the other. A covariance analysis between exposure to listening and accuracy for both groups combined put the  $\rho$ -value for the pre-test at 0.14, and for the post-test at -0.18. These correlations are weak, suggesting that there is a low likelihood of exposure being correlated with accuracy.

Group		Accuracy							
		M	S <sub>d</sub>	S <sub>e</sub>	$\sigma^2$	$\rho$	$P_{group}$	$P_{time}$	$P_{interaction}$
Comparison	Pre-test	0.0269	0.0192	0.0048	0.0004	0.10	0.23	0.30	0.88
	Post-test	0.0214	0.0125	0.0031	0.0002	0.34			
Intervention	Pre-test	0.0321	0.0212	0.0055	0.0005	-0.10			
	Post-test	0.0278	0.0187	0.0048	0.0004	-0.49			

Table 5: Normalized error values obtained after applying the error analysis adapted from Verspoor et al. (2012) and Ellis (2005). The closer to 0 the value is, the fewer mistakes were identified.

Figures 4 & 5 give a visual overview of the number of mistakes identified per category and per group. In total, for the pre-test, the comparison group made 100 mistakes distributed over 11 categories, with an average of 6.25 mistakes per participant. The intervention group had a total of 112 mistakes distributed over 11 categories, with an average of 7.47 mistakes per participant. It should be noted that the categories the comparison group performed worse in when compared to the intervention group are *collocation*, *plural/singular* and *word form (adverb)*, whereas this does not seem to be the case for the post-test where the category *preposition* seems to be causing the most trouble along with *pronouns*.

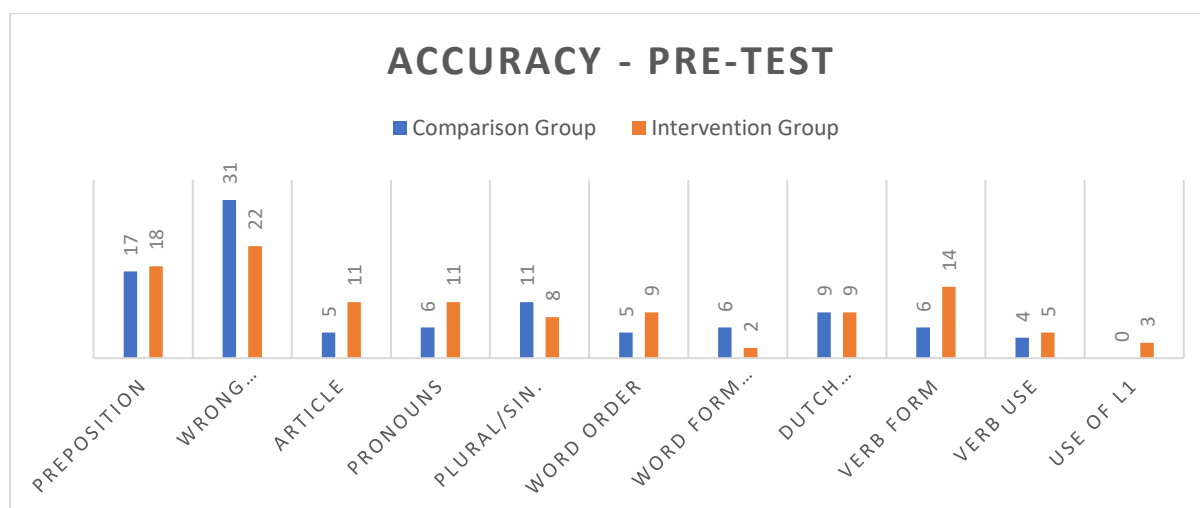


Figure 4: Visualized overview of errors (cf. Verspoor et al., 2012; Ellis, 2005) for the pre-test.

For the post-test, the total number of errors identified in the comparison group was 81, which constitutes a 19% decrease compared to the pre-test (normalized decrease: -0.005). In the intervention group, the total number of errors was coincidentally also determined to be 81, which constitutes a 28% decrease compared to the pre-test (normalized decrease: -0.006). However, as Figure 5 demonstrates, despite the total number of errors being equal amongst both groups, their distributions are not. This set of data was not tested for statistical significance, so no further observations may be made.

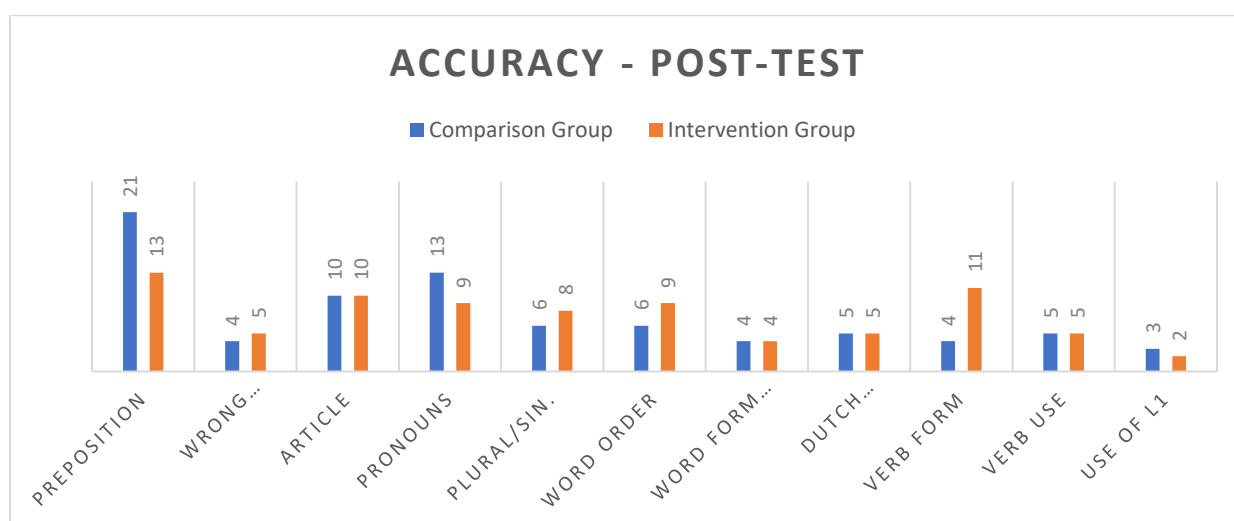


Figure 5: Visualized overview of errors (cf. Verspoor et al., 2012; Ellis, 2005) for the post-test.

## 2. Complexity

A total of five aspects—readability (as measured on the Flesch-Kincaid scale), lexical density, simple sentence use, complex sentence use and compound sentence use—were analyzed with regards to complexity.

Table 6 demonstrates the participants' scores on the Flesch-Kincaid readability scale, where values closer to zero indicate a higher level of education required to understand the lexicon. A mixed ANOVA test, however, revealed no significant differences between the intervention group and the comparison group for 'group' [ $F(1, 66) = 0.37, p = 0.55$ ], 'time' [ $F(1, 66) = 0.91, p = 0.34$ ] or 'interaction' [ $F(1, 66) = 0.02, p = 0.89$ ]. Therefore, no observations may be made concerning greater improvements in one group compared to the other. The  $\rho$ -value for the groups combined is -0.02 for the pre-test and 0 for the post-test, which shows a very weak correlation between exposure to English outside of school and readability.

After doing a linear regression analysis for the comparison group concerning the pre-test, it was found that the normalized grade  $G$  for readability can be expressed as a linear function of the normalized time spent exposed to listening to English as follows:  $G = (-0.040 \pm 0.027) * T + (0.010 \pm 0.002)$ , with a correlation coefficient  $\rho = 0.52$ . This is not significant, so further observations cannot be made.

Group		Flesch-Kincaid					
		M	S <sub>d</sub>	$\rho$	$P_{group}$	$P_{time}$	$P_{interaction}$
Comparison	Pre-test	67.78	8.01	0.52	0.55	0.34	0.89
	Post-test	65.57	7.07	-0.05			
Intervention	Pre-test	68.35	6.49	0	0.55	0.34	0.89
	Post-test	66.32	6.87	0.38			

Table 6: Student results on the Flesch-Kincaid scale. The closer the values are to zero, the more difficult a text is to read, i.e. more complex.

In terms of lexical density, a mixed ANOVA test revealed no significant differences either for the condition 'group' [ $F(1, 66) = 1.43, p = 0.24$ ], 'time' [ $F(1, 66) = 0.24, p = 0.94$ ] or 'interaction' [ $F(1, 66) = 3.97, p = 0.05$ ]. Therefore, no further observations may be made concerning what effect may have been achieved. A covariance analysis revealed the  $\rho$ -value for the groups combined to be at 0.04 for the pre-test and at -0.01 for the post-test, which—as with readability—shows little correlation. Therefore, no further observations may be made concerning the way in which lexical density and exposure to English are interwoven.

Group		Lexical Density					
		M	S <sub>d</sub>	$\rho$	$P_{group}$	$P_{time}$	$P_{interaction}$
Comparison	Pre-test	50.28	7.62	-0.30	0.24	0.94	0.05
	Post-test	48.12	4.97	0.17			
Intervention	Pre-test	49.46	5.54	0.01	0.24	0.94	0.05
	Post-test	52.77	5.42	0.35			

Table 7: Student results with regard to lexical density.

Table 8 reports on the distribution of simple, complex and compound sentence use in the participants' writing assignments, where all the values have been normalized. A mixed ANOVA test revealed no significant differences between both groups for any of the

conditions (see Table 8), which means that no further observations may be made in terms of whether or not, or in which way, sentence structure use has improved in either group.

		<i>Simple Sentence</i>					
<b>Group</b>		<b>M</b>	<b>S<sub>d</sub></b>	<b>ρ</b>	<b>P<sub>group</sub></b>	<b>P<sub>time</sub></b>	<b>P<sub>interaction</sub></b>
Comparison	Pre-test	0.049	0.014	0.23	0.314	0.082	0.711
	Post-test	0.042	0.020	0.23			
Intervention	Pre-test	0.044	0.015	-0.19			
	Post-test	0.035	0.016	0.10			
		<i>Complex Sentence</i>					
<b>Group</b>		<b>M</b>	<b>S<sub>d</sub></b>	<b>ρ</b>	<b>P<sub>group</sub></b>	<b>P<sub>time</sub></b>	<b>P<sub>interaction</sub></b>
Comparison	Pre-test	0.026	0.008	-0.19	0.132	0.593	0.682
	Post-test	0.023	0.007	0.32			
Intervention	Pre-test	0.028	0.008	-0.31			
	Post-test	0.028	0.007	0.16			
		<i>Compound Sentence</i>					
<b>Group</b>		<b>M</b>	<b>S<sub>d</sub></b>	<b>ρ</b>	<b>P<sub>group</sub></b>	<b>P<sub>time</sub></b>	<b>P<sub>interaction</sub></b>
Comparison	Pre-test	0.008	0.006	-0.09	0.993	0.558	0.320
	Post-test	0.008	0.005	-0.37			
Intervention	Pre-test	0.007	0.005	0.11			
	Post-test	0.009	0.005	0			

Table 8: Student results with regard to the distribution of sentence types, expressed as normalized values.

### 3. Coherence

Coherence, i.e. the linking of adjacent sentences and overall structure in an explicit manner through the use of discourse markers, was analyzed in imitation of Dülger's (2007) method of counting all instantiations of discourse markers. Each instantiation was grouped under one of three categories: (i) incorrect<sup>4</sup>; (ii) micro-level; (iii) macro-level. A further subcategorization for micro-level discourse markers was designed in accordance with categories conceived by Feng (2010, p. 300) and Dülger (2007, p. 261-262). These categories include: (a) additive; (b) contrastive; (c) causal; (d) continuatives; (e) focusing; (f) exemplification; (g) referral to expectations; (h) generalizing; (i) summarizing.

After counting the instantiations and classifying them in accordance with the groups mentioned above, the difference of all instantiations—i.e. total use—between the post-test normalized results and the pre-test normalized results were calculated for both groups. The values were normalized by dividing the number of instantiations by the total number of words in a given participant's text. After this, a mixed ANOVA test was run to analyze this set of data, yielding the results outlined in Table 9. Unfortunately, no statistically significant differences were found for any of the conditions, either for micro-level discourse markers or macro-level discourse markers, making further observations impossible. A covariance analysis revealed the  $\rho$ -value for both groups combined to be 0.04 for the pre-test and 0.13 for the post-test concerning micro-level discourse markers, while the values are 0.06 for the pre-test and -0.11 for the post-test concerning macro-level discourse markers. These values indicate little correlation.

<sup>4</sup> Even though the category "incorrect" was conceived for the purpose of inclusiveness, there were no instances of participants misusing a particular discourse marker.

		<i>Discourse Markers at Micro-Level</i>						
<b>Group</b>		<b>M<sub>d</sub></b>	<b>S<sub>d</sub></b>	<b>S<sub>e</sub></b>	<b>ρ</b>	<b>P<sub>group</sub></b>	<b>P<sub>time</sub></b>	<b>P<sub>interaction</sub></b>
Comparison	Pre-test	0.033	0.012	0.003	-0.27	F = 0.70	F = 2.02	F = 0.21
	Post-test	0.030	0.006	0.002	-0.02			
Intervention	Pre-test	0.031	0.008	0.002	-0.21			
	Post-test	0.027	0.009	0.002	-0.31			
		<i>Discourse Markers at Macro-Level</i>						
<b>Group</b>		<b>M<sub>d</sub></b>	<b>S<sub>d</sub></b>	<b>S<sub>e</sub></b>	<b>ρ</b>	<b>P<sub>group</sub></b>	<b>P<sub>time</sub></b>	<b>P<sub>interaction</sub></b>
Comparison	Pre-test	0.004	0.004	0.001	0.06	F = 0.02	F = 1.08	F = 0.01
	Post-test	0.003	0.005	0.001	0.42			
Intervention	Pre-test	0.004	0.004	0.001	-0.39			
	Post-test	0.003	0.006	0.001	-0.43			

Table 9: Normalized values for coherence in terms of micro-level and macro-level markers.

Furthermore, some other observations were made during the analysis of the data even though these were not originally questions to focus on. First, a number of participants misused particular discourse markers that typically appear in pairs, such as *on the one hand* with the accompanying *on the other hand* (or conversely, *firstly* without *secondly/thirdly*). They would use only one half and leave the other unmentioned during both the pre-test and post-test, suggesting a possible fossilization of interlanguage error (cf. Han & Tarone, 2014; Jordens, 1997), which cannot be assumed to correct itself without exterior intervention from an educator.

Secondly, the participants in this research showed an overall preference for simple discourse markers (e.g. *but*, *because*, *so*), with only few opting to use more advanced synonyms (e.g. *however*, *therefore*, *thus*) (cf. Table 10). Table 10 shows the categories “contrastive” and “causal” because, percentage-wise, the other categories were used very minimally and therefore could not produce relevant data. However, a mixed ANOVA test revealed no significant difference between groups for the contrastive markers for the condition ‘group’ [F(1, 66) = 1.52, p = 0.22], ‘time’ [F(1, 66) = 3.76, p = 0.15] or ‘interaction’ [F(1, 66) = 3.32, p = 0.88]. The same goes for the causal markers for the condition ‘group’ [F(1, 66) = 1.56, p = 0.12], ‘time’ [F(1, 66) = 2.85, p = 0.10] and ‘interaction’ [F(1, 66) = 1.82, p = 0.18].

Thirdly, despite the essay-like format of the pre-test and post-test, hardly any summarizing discourse markers were utilized. In fact, the percentage of such markers in relation to all markers at the micro-level never exceeds 0.05 for either group and for either test.

<b>Discourse Marker</b>		<b>Pre-test total</b>	<b>Percentage</b>	<b>Post-test total</b>	<b>Percentage</b>
Intervention Group + Comparison Group	BUT	58	0.91	50	0.88
	HOWEVER	6	0.09	7	0.12
	BECAUSE	53	0.87	58	0.97
	THEREFORE	8	0.13	2	0.03

Table 10: Use of simple versus complex discourse markers with regards to the categories “contrastive” (*but/however*) and “causal” (*because/therefore*).



### Motivation

For this research, participants were subjected to Papi & Teimouri's (2014) questionnaire (see Appendix B) and asked to evaluate on a scale of 1 (= 'completely disagree') to 5 (= 'completely agree') to what degree they identified with the statements. Table 11 below reports on the average scores as well as standard deviations for each group and for each category. For each category, a mixed ANOVA test was run with Group as a between-subjects variable and Time as a within-subjects variable in order to calculate statistical significance. The results of this analysis are reported on in Table 12, along with the  $\rho$ -values per group, per test. The results reported on in Table 12 reveal no significant differences between both groups, so no further observations may be made.

	Comparison Group				Intervention Group			
	Pre-test		Post-test		Pre-test		Post-test	
<i>Type of motivation</i>	<b>M</b>	<b>S<sub>d</sub></b>	<b>M</b>	<b>S<sub>d</sub></b>	<b>M</b>	<b>S<sub>d</sub></b>	<b>M</b>	<b>S<sub>d</sub></b>
Future Image	3.5	1.1	3.8	1.1	3.7	1.0	3.5	1.0
Ought-to	2.2	1.2	2.5	1.3	2.2	1.3	1.9	1.1
Intrinsic Motivation	3.5	1.1	3.3	1.1	3.3	1.2	3.0	1.1
Instrumentality – Promotion	3.5	1.2	3.8	1.2	3.8	1.1	3.4	1.2
Instrumentality – Prevention	3.2	1.4	3.2	1.3	3.2	1.3	2.9	1.4
Family Influence	2.6	1.2	2.3	1.2	2.4	1.3	2.2	1.2
Anxiety	3.1	1.3	2.4	1.2	2.4	1.3	2.7	1.3
Learning Experience	0.6	0.5	0.6	0.5	0.7	0.4	0.5	0.5
Attitudes to L2 community	0.9	0.3	0.9	0.3	0.9	0.3	0.9	0.3

Table 11: Overview of motivation in various contexts for both classes (cf. Papi & Teimouri, 2014). Note that "Learning Experience" and "Attitudes of L2 community" are graded on a scale of 0-1 as students could only answer no (=0) or yes (=1).

A linear regression analysis revealed that the grade G for the various values in Table 12 may be expressed as a linear function of the normalized time spent exposed to listening to English as follows:

- *Intervention, future image, pre-test*:  $G = (-4.764 \pm 1.860) * T + (3.996 \pm 0.234)$ , with a correlation coefficient  $\rho = -0.57$ .
- *Intervention, anxiety, pre-test*:  $G = (9.710 \pm 3.735) * T + (2.155 \pm 0.469)$ , with a correlation coefficient  $\rho = 0.58$ .
- *Comparison, anxiety, pre-test*:  $G = (-11.889 \pm 4.852) * T + (3.285 \pm 0.373)$ , with a correlation coefficient  $\rho = -0.56$ .
- *Comparison, learning experience, pre-test*:  $G = (0.067 \pm 0.539) * T + (0.616 \pm 0.068)$ , with a correlation coefficient  $\rho = 0.67$ .

Correspondingly, there is a slight indication that exposure to English outside of school may be correlated to the future image a person has concerning English (see Table 12 on the next page). Interestingly, exposure may also be correlated with the level of anxiety a person feels and the level of enjoyment experienced when learning English. Perhaps, the more exposure a person has to English in a natural context, the more at ease a person feels when using English. The correlation observed with regards to learning experience is also present for the comparison group, who had less exposure to English outside of school than

the intervention group. Interestingly, the correlation goes in the exact opposite way, being negatively correlated.

Category	$P_{group}$	$P_{time}$	$P_{interaction}$		$\rho_{pre}$	$\rho_{post}$
Future image	0.07	0.80	0.86	Intervention	<b>-0.57</b>	0.11
				Comparison	0.05	-0.14
Ought-To	0.11	0.81	0.46	Intervention	-0.09	0.19
				Comparison	-0.35	-0.03
Intrinsic Motivation	0.10	0.45	0.27	Intervention	0.15	0.42
				Comparison	0.27	-0.05
Instrumentality – Promotion	0.06	0.64	0.74	Intervention	0.02	-0.37
				Comparison	0.07	-0.16
Instrumentality – Prevention	0.61	0.51	0.31	Intervention	0.46	0.03
				Comparison	-0.36	-0.04
Family Influence	0.81	0.32	0.32	Intervention	0.19	-0.17
				Comparison	-0.11	-0.36
Anxiety	0.08	0.52	0.97	Intervention	<b>0.58</b>	0.03
				Comparison	<b>-0.56</b>	0.35
Learning Experience	0.12	0.22	0.92	Intervention	0.03	0.03
				Comparison	<b>0.67</b>	-0.31
Attitudes to L2 Community	0.61	0.87	0.87	Intervention	0.25	0.22
				Comparison	0.31	-0.28

Table 12: Results from mixed ANOVA test and covariance analysis on the various categories measured by the questionnaire. Bolded correlation coefficients indicate interesting values.

## 6. Discussion

This research focused on the question of how an extensive reading program might impact the writing proficiency and motivation of the participants compared to a traditional reading program. In order to answer this question, four sub-questions were formulated, three of which emphasized elements identified as constituting writing proficiency while the last one emphasized motivation to learn English from a variety of angles. Each sub-question will be discussed separately below.

### *1. Does extensive reading improve the accuracy of the students' written L2 production to a greater extent than traditional reading?*

In this research, a synthesized error identification model was used (cf. Verspoor et al., 2012; Ellis, 2005) to analyze the participants' performances concerning proper use of grammar in English. It was hypothesized that participants in the extensive reading program would improve more in terms of accuracy, as findings from a number of experiments can be extrapolated to suggest that such is the case (Hafiz & Tudor, 1989; Hernández, 2008), especially since typologically similar languages—such as English and Dutch—facilitate noticing, and thus the learning process (cf. Leung & Williams, 2014).

No statistically significant differences were found between the intervention group and the comparison group in this research, which makes it impossible to put into evidence a supposed link between extensive reading and improved accuracy with the data collected here. Therefore, this question cannot be conclusively answered, even if at first glance accuracy seemed to improve in both groups. Perhaps, the small population ( $N = 35$ ) that this research worked with may account for the lack of statistically significant findings, as small populations typically lead to issues such as non-generalizability. Additionally, this research did not use graded readers, unlike other research (cf. Hernández, 2008), which might account for the different results found here. Furthermore, the intervention period was perhaps not long enough to effect noticeable differences between both groups, as Yamashita (2008) suggests that extensive reading programs typically require a long time before changes and/or improvements may be observed.

The aspect of accuracy was also correlated with exposure to listening to English outside of school. A covariance analysis revealed weak  $\rho$ -values, which suggests that no conclusions may be drawn concerning a participant's exposure outside of school and that participant's accuracy in this particular research.

### *2. Does extensive reading improve the complexity of the students' written L2 production to a greater extent than traditional reading?*

Complexity, i.e. the ability to use a varied speech to express thoughts on paper, was theorized to improve to a greater extent in the extensive reading program because this group was subjected to large quantities of English texts at the participants' respective, individual levels of proficiency. Since such texts could consequently function as models from which to distill essential information (cf. Lyster, 2007; Prabhu, 1987; Ellis, 1995; Spada, 1997) about writing proficiency and vocabulary (cf. Coady, 1997; Nagy, Herman & Anderson, 1985; Nation & Coady, 1988), it was theorized that the participants in the extensive reading program would experience greater leaps forward in terms of complexity, since one element—lexical density—is closely tied to a person's lexical repertoire.

However, no statistically significant differences were found between both groups, which unfortunately means that this sub-question cannot be conclusively answered. As stated before, the small population as well as the length of the intervention period may have problematized this. Concerning the correlation between exposure to listening outside of school and complexity, weak  $\rho$ -values were found, which again makes it hard to draw any conclusions pertaining to this.

*3. Does extensive reading improve the coherence, i.e. structure, of the students' written L2 production at the micro- and/or macro-level to a greater extent than traditional reading?*

The final element of writing proficiency was identified in this research as coherence, i.e. the linking of adjacent sentences and overall structure in an explicit manner through the use of discourse markers. A synthesis of categories, which built on work by Dülger (2007) and Feng (2010), was used to compartmentalize and identify discourse markers in the written productions of participants.

No statistically significant differences were detected after a mixed ANOVA test, which makes it impossible to draw any conclusions. However, some observations were made concerning the use of discourse markers by the participants. For example, it was observed that participants often misused particular discourse markers that come in pairs (e.g. *on the one hand + on the other hand*) by using only one half and omitting the other. This is congruent with Feng's (2010) suggestion that not explicitly teaching discourse markers will either lead to students seldomly using them or using them incorrectly (p. 301-302), possibly because this might be one type of information that participants are unlikely to pick up on their own and then apply in their writing. It often is the case that knowledge is only retained if it is practiced sufficiently, so no explicit training can have a detrimental effect on the use and acquisition of discourse markers in an L2.

Concerning the correlation between exposure to English outside of school and coherence, the weak  $\rho$ -values make it hard to describe any effects. Therefore, no observations can be made at the present moment. Regardless, it is often believed that the more exposure one has to an L2 in a natural context, the smoother and the better the acquisition process will be (cf. Krashen, 1981; Lyster, 2007; Prabhu, 1987). Unfortunately, this research is not able to contribute to the ongoing discussion of how a language may be acquired implicitly.

*4. Does extensive reading impact the students' motivation to learn English as an L2 to a greater extent than traditional reading?*

Motivation plays a role in regulating a person's learning and information intake (Dörnyei & Csizér, 1998; Benson, 2007). Motivation obtained through autonomy is a particularly powerful factor in this regard (cf. Deci & Ryan, 2010), so it stands to reason that increasing a person's autonomy in completing a task should have a positive impact on that person's motivation. Since extensive reading allows for participants to select their own reading material, which matches their own interests and preferences, it was argued that extensive reading should theoretically have a greater impact on the participants' motivation to study English than a traditional reading program, which forces participants to read preselected, edited texts that may or may not match their personal interests.

However, the questionnaires participants were asked to fill in revealed that extensive reading did not actually have the theorized impact. The data were unfortunately shown to be statistically insignificant. Concerning the correlation between exposure to English outside of

an education setting and motivation, there is a slight indication that it is correlated to the participants' future image, level of anxiety experienced, and the level of enjoyment experienced when learning English. At face value, it would seem logical to assume that a person who is exposed to large quantities of English will feel more at ease navigating in an English-speaking context, whereas someone with little exposure—and therefore less practice—will feel uncomfortable if positioned in the same context.

As an aside, it may be suggested that perhaps the participants did not experience any actual autonomy, as they were still forced to read at particular times. If this is true, then the sheer ability to choose what to read cannot be seen as a kind of motivator, nor can it be claimed to be conducive to increased (intrinsic) motivation. Conversely, perhaps participants in an extensive reading program may not fully understand the benefit of an extensive reading program, instead preferring a traditional reading program because such a program might better conform to their expectations of what efficient learning looks like in the classroom. Unfortunately, since no interviews were held, it is impossible to draw any conclusions concerning the outcome of the questionnaires. More research would have to be carried out in order to more accurately reflect what factors are contributing here.

## 7. Conclusion

The question of how participation in an extensive reading program, as compared to participation in a traditional reading program, relates to writing proficiency in the context of two 5VWO classes was the main issue investigated in this research. Three interrelated variables—accuracy, complexity and coherence—were used to operationalize this question, while the impact of extensive reading on motivation was also investigated. It was hypothesized that learning gains in the extensive reading program would be greater than those in the traditional reading program because of increased autonomy, greater ease of reading and greater exposure to authentic English language sources.

However, no statistically significant differences were found between both groups after running mixed ANOVA tests. Therefore, this research cannot give any definitive or conclusive answers concerning extensive reading and writing proficiency, even though prior research has suggested a link between these factors (Hafiz & Tudor, 1989; Hernández, 2008). It must be pointed out that the small samples in both reading programs complicates the generalizability of results while also minimalizing the chance of determining statistical significance. For this reason, the research would have to be replicated—not just with greater sample sizes, but also in divergent contexts to investigate whether extensive reading impacts other age groups and/or other educational levels to the same degree, or perhaps to a greater or lesser extent.

Furthermore, the correlation between exposure to English outside of school—with particular reference to listening—and the various aspects which this research focused on was analyzed as well. Tentatively, it may be suggested that increased exposure to English outside of school has an impact on the level of anxiety and the level of enjoyment respectively that an individual experiences when dealing with English. The same goes for future image, meaning that a person who has increased exposure to English is likelier to expect to use English (professionally) in their future compared to someone who is not exposed to English as much. These claims come from analyses performed on the data obtained from the motivation questionnaires (cf. Papi & Teimouri, 2014).

### *Limitations of the current research*

Although every measure was taken to ensure validity, there are a number of methodological flaws that need to be discussed so future research can eliminate them and produce more detailed results.

The questionnaire about motivation was presented to the students with items in a non-randomized manner due to a technical fault, so that questions belonging to the same cluster appeared one after the other. This may potentially have had an influence on the results and should be eliminated in future research.

To analyze coherence, this research focused on discourse markers. However, it may be suggested that discourse markers alone do not represent the full spectrum of what constitutes coherence in writing. For example, the appropriate use of parallelism as a rhetorical device as well as tactical repetition of keywords constitute two other aspects of coherence which future research would also have to look at in order to paint a more representative picture.

In the absence of a third group whose members were comparable to the intervention group and comparison group, it was difficult to firmly establish the comparability of the pre-

test and post-test. Therefore, it could not be demonstrated whether the tests were conceptually similar enough before they were administered.

Due to time constraints, no possibility was present for the administration of a delayed post-test, which means it was impossible to establish whether the changes observed in this research actually stabilized in either group, or whether they were retained to some degree. Therefore, future research into the same or a similar topic should seek to plan a delayed post-test to account for this.

The participants in both groups were not subjected to a reading comprehension or vocabulary test before or after the intervention took place. This problematizes certain findings, in the sense that the changes observed here cannot be automatically attributed to participation in either reading program. In other words, the lack of data from such tests makes it harder to determine how much participation in either program actually influenced the participants' writing proficiency concerning complexity.

### *Recommendations for educators*

Despite the limitations discussed above, certain recommendations may be given for educators who are seeking to implement extensive reading into their teaching. These recommendations are summed up below and are based on findings from this research and other research (see #3).

1. Students need to be made aware of the potential benefits of extensive reading, which include increased vocabulary and increased exposure to the target language as used in an unedited, native context. Otherwise, the students' motivation to participate is likely to be low, which is detrimental for their learning gains.
2. Extensive reading programs are likely to produce greater learning gains if they are supplemented with explicit instruction, especially concerning coherence markers. Therefore, it might be a good idea to collect certain material the students have been reading, and to use that material to closely examine, for example, the use of coherence markers.
3. In order to improve the students' sense of success, it might be a good idea to have them keep track of what they have been reading in a portfolio and to ask them to self-report on the vocabulary they have learned. This way, students will be made aware of their own learning gains, which is likely to promote a positive sense of self-efficacy and a better understanding of the usefulness of the program.

### References

- Abitzsch, D., Van der Knaap, E., Abbate, R., Dawidowicz, M., Feld-Knapp, I., Hoffmann, S., Perge, G., & Schramm, K. *Freies Lesen im LEELU-LehrerInnenbildungsprojekt. Vom Forschungsstand zu einer Handreichung für den Unterricht*. Retrieved from [https://leelu.eu/wp-content/uploads/sites/164/2018/03/Konzeptpapier-dO1\\_12030218.pdf](https://leelu.eu/wp-content/uploads/sites/164/2018/03/Konzeptpapier-dO1_12030218.pdf)
- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning*, pp. 259-302. Honolulu, HI: University of Hawai'i Press.
- Anderson, J. C., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction*, 12, 439-462.
- Ayoun, D. (2004). The effectiveness of written recasts in the second language acquisition of aspectual distinctions in French: A follow-up study. *Modern Language Journal*, 88, 31-55.
- Ben-Anath, D. (2006). The Role of Connectives in Text Comprehension. *TESOL & Applied Linguistics*, 5(2), 1-27.
- Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching*, 40(1), 21-40.
- Brown, H. D. (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. Addison Wesley Longman, Inc. (Second edition).
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 225-237). Cambridge: Cambridge University Press.
- Corbett, E. P.J. (1987). *The Little English Handbook: Choices and Conventions*. U.S.A.: Scott, Foresman and Company.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Day, R., & Bamford, J. (2002). Top Ten Principles for Teaching Extensive Reading. *Reading in a Foreign Language*, 14(2), 136-141.
- de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249-76.



- de Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. *The handbook of language teaching*, 726-755.
- de la Fuente, M. (2009). The role of pedagogical tasks and focus on form in acquisition of discourse markers by advanced learners. In R.P. Leow, H. Campos & D. Lardiere (Eds.), *Little words: Their history, phonology, syntax, semantics, pragmatics, and acquisition*, pp. 211-221. Washington, DC: Georgetown University Press
- Deci, E. L., & Ryan, R. M. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379-410.
- DeKeyser, R. (1997). Beyond explicit rule learning: Automating second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195-221.
- Didau, D. (2013). Black space: improving writing by increasing lexical density. *The Learning Spy: Brain Food for the Thinking Teacher*. Accessed on November 27, 2018. <https://learningspy.co.uk/writing/black-space-increasing-lexical-density/>
- Dörnyei, Z., & Csizér, K. (1998). Ten commandments for motivating language learners: results of an empirical study. *Language Teaching Research*, 2(3), 203-229.
- Du, X. (2009). The affective filter in second language teaching. *Asian Social Science*, 5(8), 162-165.
- Dülger, O. (2007). *Discourse markers in writing*. Üniversitesi Eğitim Fakültesi
- Dweck, C. S., & Leggett, E. L. (1998). A Social-Cognitive Approach to Motivation and Personality. *Psychological Review*, 95(2), p. 256-273.
- Easthope, C., & Easthope, G. (2000). Intensification, Extension and Complexity of Teachers' Workload. *British Journal of Sociology of Education*, 21(1), 43-58.
- Ellis, N. (1995). Consciousness in second language acquisition: a review of field studies and laboratory experiments. *Language Awareness*, 4(3), 123-146.
- Ellis, N. (1998). Emergentism, connectionism, and language learning. *Language Learning*, 48, 631-664.
- Ellis, N. C. (2002). Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188.
- Ellis, N. (1993). Rule and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5, 289-318.

- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition*, 27, 141-172.
- Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Bristol, UK: Multilingual Matters.
- Feng, L. (2010). Discourse markers in English writers. *The Journal of International Social Research*, 3(11), 209-304.
- Fraser, B. (1998). Contrastive discourse markers in English. In A. H. Jucker & Y. Ziv (ed.), *Discourse Markers: Descriptions and Theory*, pp. 301-320. Amsterdam: Benjamin.
- Grabe, W. & Stoller, F. (2002). *Teaching and researching reading*. Harlow, UK: Longman.
- Hafiz, F. M., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal*, 43(1), 4-13.
- Han, Z., & Tarone, E (eds.). (2014). *Interlanguage: Forty years later*. John Benjamins Publishing Co.: Amsterdam.
- Hernández, T.A. (2008). The effect of explicit instruction and input flood on students' use of discourse markers on a simulated oral proficiency interview. *Hispania*, 91, 665-75.
- Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, 15(2), 159-d 182.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61, 355-382.
- Hsieh, P-H. P., & Schallert, D. L. (2008). Implications from self-efficacy and attribution theories for an understanding of undergraduates' motivation in a foreign language course. *Contemporary Educational Psychology*, 33(4), 513-532.
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Huang, S. (2006). Reading English for academic purposes: What situational factors may motivate learners to read? *System*, 34, 371-383.
- Huckin, T. & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21, 181-193.
- Huyghebaert, T., Gillet, N., Beltou, N., Tellier, F., & Fouquereau, E. (2018). Effects of workload on teachers' functioning: A moderated mediation model including sleeping problems and overcommitment. *Stress and Health*, 0(0), 1-11.

- Huysmans, F., De Haan, J., Van den Broek, A. (2004). *Achter de schermen: een kwart eeuw lezen, luisteren, kijken en internetten*.
- Jordens, P. (1997). Introducing the Basic Variety. *Second Language Research*, 13(4), 289-300.
- Krashen, S (1982). *Principles and Practices in Second Language Acquisition*. Pergamon Press, Oxford.
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212-218.
- Lantolf, J., & Thorne, S. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.
- Laufer, B., & Nation, P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*, 16(3), 307-322.
- Leung, J., & Williams, J. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition*, 36, 733-755.
- Loewen, S. (2011). Focus on form. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning*, Vol. 2, 576-592. London: Routledge.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269-300.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.
- Maria, P., & Norbert, S. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Mondria, J. A. & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and retention of words in a foreign language. *Applied Linguistics*, 12, 249-267.
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233-253.
- Nation, I. S. P. & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.
- Ngeow, K. (1998). *Motivation and transfer in language learning*. ERIC Digest. Retrieved from <http://www.ericdigests.org/1999-4/motivation.htm>
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528.

- Paasche-Orlow, M. K., Taylor, H. A., & Brancati, F. L. (2003). Readability Standards for Informed-Consent Forms as Compared with Actual Readability. *The New England Journal of Medicine*, 348(8), 721-726.
- Papi, M., & Teimouri, Y. (2014). Language Learner Motivational Types: A Cluster Analysis Study. *Language Learning. A Journal of Research in Language Studies*, 64(3), p. 493-525.
- Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109-129). Norwood, NJ: Ablex.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eyetracking study. *Studies in Second Language Acquisition*, 38, 97-130.
- Pica, T., Kang, H. S., & Sauro, S. (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition*, 28, 301-308.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Raimes, A. (1983). *Techniques in Teaching Writing*. Hong Kong: O.U.P.
- Redeker, G. (1991). Review Article: Linguistic markers of discourse structure. *Linguistics*, 29(6), p. 1139-1172.
- Reinders, H., & Ellis, R. (2009). The effect of two types of enhanced input on intake and the acquisition of implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elders, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 281-302). Bristol: Multilingual matters.
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E., & Köller, O. (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Münster, New York: Waxmann.
- Richards, J.C., & Schmidt, R. (2002). *Longman Dictionary of Language Teaching and Applied Linguistics*. Pearson Education, London.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27-68.
- Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 19, 223-47.
- Rosa, E.E. & Leow, R.P. (2004). Computerized task-based exposure, explicitness and type of feedback on Spanish L2 development. *Modern Language Journal*, 88, 192-217.

- Rudolf, F. (2016). How to Write Plain English. *University of Canterbury*. Retrieved on November 27, 2018.
- Sanders, T., Land, J., & Mulder, G. (2007). Linguistics markers of coherence improve text comprehension in functional contexts. *Information Design Journal*, 15(3), 219-235.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University.
- Seow, A. (2002). The Writing Process and Process Writing. In Jack C. Richards, & Willy A. Renandya, (Eds.), *Methodology in Language Teaching: An Anthology of Current practice* (p. 315-320). New York: Cambridge University Press
- Smith, F. (2004). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Routledge.
- Spada, N. (1997). Form-Focused Instruction and Second Language Acquisition: A Review of Classroom and Laboratory Research. *Language Teaching*, 30(2), 73-87.
- Stokking, K. (2016). *Bouwstenen voor onderzoek in onderwijs en opleiding*. Antwerpen-Apeldoorn: Garant.
- Tavakol, M., & Dennick, R. (2011). Making Sense of Cronbach's Alpha. *International Journal of Medical Education*, 2, 53-55.
- Thornbury, S. (2002). *How to teach vocabulary*. London: Longman
- Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271- 297. Doi: 10.1007/s10648-010-9127-6
- Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K.M., & Deci, E.L. (2004). Motivating Learning, Performance, and Persistence: The Synergistic Effects of Intrinsic Goal Contents and Autonomy-Supportive Contexts. *Journal of Personality and Social Psychology*, 87(2), 246-260.
- Verboord, M. (2006). Leesplezier als sleutel tot succesvol literatuuronderwijs. In K. Hilderdink & S. Wagenaar (Eds.), *Leescultuur onder vuur*, 35-52.
- Verspoor, M., Schmid, M., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239-263.
- Wang, S. (2013). Reading and Grammar Learning Through Mobile Phones. *Language, Learning and Technology*, 17(3), p. 117-134.
- Warford, M. K., & White, W. L. (2012). Reconnecting proficiency, literacy, and culture: From theory to practice. *Foreign Language Annals*, 45(3), 400-414.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

- William, J. D. (1989). *Preparing to Teach Writing*. California: Wadsworth Publishing Co.
- Wood, D., & Middleton, D. (1975). A study of assisted problem-solving. *British Journal of Psychology*, 66, 181-191. doi: 10.1111/j.2044-8295.1975.tb01454.x
- Woolfolk, A., Hughes, M., & Walkup, V. (2013). *Psychology in Education* (2<sup>nd</sup> Edition). Pearson Education Limited.
- Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, 36(4), 661-672.
- Yoshimi, D.R. (2001). Explicit instruction and JFL learners' use of interactional discourse markers. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching*, p. 223-44. Cambridge: Cambridge University Press.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, 57, p. 541-572.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166-183.

### Appendix A – Language Profile Questionnaire

#	Vraag	Antwoord
1	Hoe oud ben je?	
2	Hoeveel jaar krijg je al Engelse les? (Neem ook je tijd op de basis-school, je tijd op deze school en overige cursussen mee in je antwoord).	
3	Gemiddeld genomen, hoeveel minuten per week krijg je Engelse les? Hoe was dit op de basisschool?	
4	Welke taal / talen spreek je vloeiend?	
5	Is Nederlands je moedertaal? Zo niet, welke dan?	
6	Hoeveel minuten per dag luister je gemiddeld naar Engelstalige bronnen? (Series, programma's, reclames, Netflix, podcasts, familieleden, etc.)	
7	Hoeveel minuten per dag lees je gemiddeld Engelstalige teksten? (Boeken, magazines, blogs op het internet, etc.)	
8	Hoeveel minuten per dag schrijf je gemiddeld in het Engels om wat voor reden dan ook?	
9	Hoeveel minuten per dag spreek je gemiddeld Engels?	
10	Hoeveel uur per week luister je gemiddeld naar Engelstalige bronnen?	
11	Hoeveel uur per week lees je gemiddeld Engelstalige teksten?	
12	Hoeveel uur per week schrijf je gemiddeld in het Engels?	
13	hoeveel uur per week spreek je gemiddeld in het Engels?	
14	In welke klas zit je?	

**Appendix B – Language Learning Motivation Questionnaire**

Due to formatting issues, the questionnaire can be found on the next page.



#	Vraag	Antwoord
1	Ik kan mezelf voorstellen als iemand die Engels spreekt als een moedertaalspreker.	
2	Ik zie wel voor me dat ik Engels spreek met internationale vrienden of collega's.	
3	Als ik nadenk over mijn toekomstige carrière, dan stel ik mezelf voor dat ik Engels spreek.	
4	Ik zie het voor me dat ik studeer op een universiteit of HBO waar al mijn cursussen in het Engels worden gegeven.	
5	Ik kan mezelf voorstellen als iemand die Engelse e-mails vloeiend schrijft.	
6	Ik stel mezelf voor als iemand die in het buitenland kan wonen en daar Engels gebruikt om effectief te communiceren met de mensen.	
7	Ik leer Engels omdat goede vrienden van mij dit belangrijk vinden.	
8	Als het mij niet lukt Engels te leren, heb ik het gevoel dat ik andere mensen teleurstel.	
9	Ik vind het belangrijk om Engels te leren omdat de mensen, voor wie ik respect heb, vinden dat ik deze taal moet leren.	
10	Ik studeer Engels omdat ik het belangrijk vind om de goedkeuring van vrienden, leraren en/of familie te krijgen.	
11	Ik vind het belangrijk Engels te leren omdat de mensen in mijn omgeving dit ook van me verwachten.	
12	Engels leren is voor mij belangrijk omdat andere mensen mij meer zullen respecteren als ik kennis heb van het Engels.	
13	Ik zou graag meer tijd willen hebben om Engels te leren.	
14	Ik ben bereid om me flink in te zetten om Engels te leren.	
15	Ik zou me graag meer focussen op Engels dan op andere vakken.	
16	Als er een vak wordt aangeboden in het Engels, dan zou ik die volgen.	
17	Als de leraar een niet verplichte opdracht geeft aan de klas, dan zou ik deze zeker vrijwillig maken.	
18	Ik zou Engels willen leren op school, ook al was ik daar niet toe verplicht.	
19	Engels is belangrijk voor me omdat ik denk dat ik met kennis van deze taal later een goede baan kan krijgen.	
20	Engels is belangrijk voor me omdat goede beheersing van het Engels belangrijk is voor een promotie later.	
21	Engels is belangrijk voor me omdat ik het nodig zal hebben om later onderwijs te volgen.	
22	Engels is belangrijk voor me om een speciaal doel te bereiken (bijvoorbeeld slagen voor school of een beurs ontvangen).	
23	Ik leer Engels om me op de hoogte te kunnen houden van het meest recente nieuws in de wereld.	
24	Engels is belangrijk voor me omdat ik van plan ben om in het buitenland te gaan studeren.	
25	Ik moet Engels leren omdat ik geen slechte cijfers wil.	
26	Ik moet Engels leren omdat ik mijn diploma niet kan halen als ik hierin faal.	
27	Engels leren is belangrijk omdat ik gezien word als een zwakke leerder als ik weinig kennis heb van het Engels.	
28	Engels leren is belangrijk voor me omdat ik geen slecht cijfer wil in een 'proficiency test' zoals de TOEFL, IELTS, Cambridge exam, etc.	
29	Ik moet Engels leren omdat ik dit vak niet wil falen.	
30	Engels leren is belangrijk omdat ik me zou schamen als ik slechte cijfers haalde voor Engels.	

31	Engels leren is belangrijk omdat ik niet gezien wil worden als laag opgeleid.	
32	Mijn familie vindt dat ik Engels moet leren om een goed opgeleide persoon te zijn.	
33	Engels leren is belangrijk voor me om de familie in ere te houden.	
34	Ik vind het belangrijk goed te presteren qua Engels om mijn ouders of familieleden tevreden te houden.	
35	Mijn familie zet veel druk op me om goed te presteren qua Engels.	
36	Mijn ouders moedigen me aan om zo veel mogelijk Engels te oefenen.	
37	Ik moet Engels leren omdat mijn ouders anders teleurgesteld zijn.	
38	Ik voel me ongemakkelijk als een buitenlander me in het Engels vraagt om een routebeschrijving.	
39	Ik voel me ongemakkelijk als ik Engels spreek met een moedertaalspreker.	
40	Ik word nerveus en verward als ik tijdens de les in het Engels moet spreken.	
41	Ik ben bang om dom te klinken als ik Engels spreek vanwege fouten.	
42	Ik ben bezorgd dat andere sprekers van het Engels mijn Engels raar vinden.	
43	Ik ben bang dat mensen me zullen uitlachen als ik in het Engels praat.	

#	Vraag	Ja / Nee
1	Vind je de sfeer tijdens de lessen aangenaam?	
2	Vind je het interessant om Engels te leren?	
3	Denk je dat de tijd sneller voorbij vliegt wanneer je Engels leert?	
4	Heb je altijd zin om naar de Engelse les te gaan?	
5	Zou je meer Engelse lessen willen op school?	
6	Vind je het aangenaam om Engels te leren?	
7	Vind je de mensen die in Engelstalige landen wonen aardig?	
8	Vind je het leuk om mensen uit Engelstalige landen te ontmoeten?	
9	Vind je het leuk om te reizen naar Engelstalige landen?	
10	Wil je graag meer weten over mensen die in Engelstalige landen wonen?	
11	Vind je Engelstalige muziek leuk?	
12	Vind je Engelstalige films leuk?	
13	Vind je Engelstalige programma's leuk?	
14	Vind je Engelstalige magazines, kranten en/of boeken leuk?	

## Appendix C – Grammar Test

## Student Test A

There are 60 multiple-choice questions in the test. Read the example and choose the correct answer to each question: a, b, c or d.

Time limit: 30 minutes

- 0 Tomasz \_\_\_\_\_ from Poland.  
a) are            b) is            c) aren't            d) am
- 1 Maria and Fernando \_\_\_\_\_ Spanish.  
a) is            b) isn't            c) are            d) am
- 2 They've got three \_\_\_\_\_  
a) child's.            b) childrens.            c) children.            d) child.
- 3 There's \_\_\_\_\_ pencil on the table.  
a) a            b) two            c) some            d) an
- 4 My brother's sixteen. \_\_\_\_\_ called Tom.  
a) She's            b) He's            c) It's            d) You're
- 5 I've got two sisters. \_\_\_\_\_ bedroom is very big.  
a) His            b) Your            c) Their            d) Her
- 6 \_\_\_\_\_ are you from?  
a) Where            b) What            c) When            d) Who
- 7 This is my book. \_\_\_\_\_ are your books on the table.  
a) This            b) That            c) It            d) Those
- 8 \_\_\_\_\_ are twenty students in my class.  
a) They            b) There            c) We            d) It
- 9 There's a blackboard in the classroom but there aren't \_\_\_\_\_ shelves.  
a) any            b) some            c) a            d) the
- 10 My parents have got blue eyes but my \_\_\_\_\_ hair is black.  
a) father            b) fathers            c) fathers'            d) father's
- 11 \_\_\_\_\_ you got any apples?  
a) Has            b) Have            c) Is            d) Do
- 12 They speak English but they \_\_\_\_\_ speak French.  
a) don't            b) do            c) does            d) doesn't
- 13 \_\_\_\_\_ he play the guitar?  
a) Do            b) Does            c) Is            d) Don't

- 14 I \_\_\_\_\_ up at 7 o'clock.  
a) usually get    b) get sometimes    c) get often    d) get usually
- 15 We like him but he doesn't like \_\_\_\_\_.  
a) we.    b) he.    c) they.    d) us.
- 16 She \_\_\_\_\_ a black T-shirt today.  
a) wears    b) doesn't wear    c) is wearing    d) are wearing
- 17 I don't like \_\_\_\_\_ football.  
a) play    b) playing    c) to playing    d) doing
- 18 My friend, Jack, \_\_\_\_\_ at school yesterday because he was ill.  
a) isn't    b) was    c) were    d) wasn't
- 19 Where \_\_\_\_\_ last night?  
a) did you go    b) do you go    c) you go    d) does she go
- 20 What \_\_\_\_\_ to do next weekend?  
a) do you go    b) are you going    c) are you doing    d) did they go
- 21 She \_\_\_\_\_ the piano very well.  
a) does    b) can play    c) play    d) can
- 22 We usually go to the disco on Saturdays but we \_\_\_\_\_ today.  
a) don't go    b) doesn't go    c) isn't going    d) aren't going
- 23 \_\_\_\_\_ tennis with us tomorrow?  
a) Are they playing    b) Do we play    c) You are doing    d) Does he do
- 24 She's more \_\_\_\_\_ than her sisters.  
a) big    b) taller    c) oldest    d) intelligent
- 25 London is the \_\_\_\_\_ city in Britain.  
a) most expensive    b) more expensive    c) bigger    d) beautiful
- 26 I \_\_\_\_\_ to Warsaw last week.  
a) go    b) was    c) went    d) am not going
- 27 Her Spanish is very good. She speaks it very \_\_\_\_\_.  
a) badly.    b) good.    c) quickly.    d) slowly.
- 28 We \_\_\_\_\_ a coffee in the café when we saw Tom.  
a) had    b) was having    c) are having    d) were having
- 29 The music is very loud, Bob. \_\_\_\_\_ it down, please.  
a) Turned    b) Turning    c) Turn    d) Don't turn

- 30** You \_\_\_\_\_ take your passport when you travel to another country.  
 a) must            b) should            c) mustn't            d) don't have to
- 31** We \_\_\_\_\_ you next week.  
 a) see            b) will see            c) is going to see            d) is seeing
- 32** If she \_\_\_\_\_ the exam, she'll go to university.  
 a) is passing            b) will pass            c) passes            d) won't pass
- 33** I'll buy \_\_\_\_\_ milk if I go to the supermarket.  
 a) a            b) an            c) some            d) any
- 34** \_\_\_\_\_ you ever met a famous person?  
 a) Has            b) Do            c) Did            d) Have
- 35** They've never \_\_\_\_\_ to a rock concert.  
 a) saw            b) seen            c) gone            d) been
- 36** It's not my bag. It's \_\_\_\_\_  
 a) hers.            b) her.            c) him.            d) mine.
- 37** He hasn't phoned \_\_\_\_\_  
 a) just.            b) already.            c) ever.            d) yet.
- 38** I'm not hungry. I \_\_\_\_\_ had lunch.  
 a) have yet            b) have just            c) already have            d) just have
- 39** You don't \_\_\_\_\_ go now. You can go tomorrow.  
 a) must            b) mustn't            c) have to            d) have
- 40** This is the best chocolate in the world. It \_\_\_\_\_ in Switzerland.  
 a) were made            b) is made            c) makes            d) made
- 41** The book \_\_\_\_\_ in 1954.  
 a) is written            b) were written            c) was written            d) wrote
- 42** If you see a snake, \_\_\_\_\_  
 a) 'll run!            b) running!            c) to run!            d) run!
- 43** At school last year I \_\_\_\_\_ wear black shoes.  
 a) must            b) mustn't            c) have to            d) had to
- 44** This jacket is \_\_\_\_\_. It's too short for me.  
 a) not long enough            b) long enough            c) enough long            d) too long
- 45** It's \_\_\_\_\_ beautiful day. Let's go out.  
 a) so            b) such            c) such a            d) very

- 46 The film \_\_\_\_\_ already started when we got to the cinema.  
a) has            b) was            c) have            d) had
- 47 If you went to bed earlier, you \_\_\_\_\_ so tired.  
a) wouldn't feel b) will feel            c) would feel            d) didn't feel
- 48 They \_\_\_\_\_ live in Buenos Aires before they went to Madrid.  
a) were            b) used to            c) had used to            d) use to
- 49 She said she \_\_\_\_\_ at 9 o' clock.  
a) was coming b) coming            c) come            d) has come
- 50 I'm not sure if I'll go to the party. I \_\_\_\_\_ stay at home.  
a) must            b) has to            c) mustn't            d) might
- 51 That's Mr Thomson. He's the teacher \_\_\_\_\_ gives us a lot of tests.  
a) when            b) which            c) who            d) where
- 52 They don't live here, \_\_\_\_\_  
a) don't they? b) do they?            c) are they?            d) aren't they?
- 53 I \_\_\_\_\_ for three hours. I haven't finished it yet!  
a) read            b) have been reading c) was reading d) am reading
- 54 I went to the supermarket \_\_\_\_\_ some food.  
a) for to buy b) to buying            c) to buy            d) for buying
- 55 They come from Rome. They \_\_\_\_\_ be Italian.  
a) could            b) can            c) must            d) can't
- 56 They're late. They must \_\_\_\_\_ the train.  
a) miss            b) missing            c) had missed            d) have missed
- 57 If I \_\_\_\_\_ known he was going to the disco, I wouldn't have gone.  
a) had            b) would have            c) has            d) have
- 58 The teacher \_\_\_\_\_ do the test again.  
a) make us            b) made us            c) makes            d) made us to
- 59 **A** I don't want to go there again. **B** \_\_\_\_\_ It was horrible.  
a) Neither I do. b) So do I.            c) Neither do I.            d) So I do.
- 60 She asked me \_\_\_\_\_ to play tennis the next day.  
a) do I want            b) if I wanted            c) do you want            d) if

### Appendix D – Writing Test A

When it comes to doing homework, opinions are rather split. Should students be given homework and if so, how many minutes or hours a day? Should certain grades (like 2 VWO and 5 VWO) be given more or less homework?

You have decided to answer this question as part of a contest to win money. You have already done a little bit of research and you have found the text below. Read this text first.

The most comprehensive research on homework to date comes from a 2006 meta-analysis by Duke University psychology professor Harris Cooper, who found evidence of a positive correlation between homework and student achievement, meaning students who did homework performed better in school. The correlation was stronger for older students—in 7th through 12th grade—than for those in younger grades, for whom there was a weak relationship between homework and performance.

Cooper’s analysis focused on how homework impacts academic achievement—test scores, for example. His report noted that homework is also thought to improve study habits, attitudes toward school, self-discipline, inquisitiveness and independent problem-solving skills. On the other hand, some studies he examined showed that homework can cause physical and emotional fatigue, fuel negative attitudes about learning and limit leisure time for children. At the end of his analysis, Cooper recommended further study of such potential effects of homework.

Despite the weak correlation between homework and performance for young children, Cooper argues that a small amount of homework is useful for all students. Second-graders should not be doing two hours of homework each night, he said, but they also **shouldn’t** be doing **no** homework.

Not all education experts agree entirely with Cooper’s assessment. Cathy Vatterott, an education professor at the University of Missouri-St. Louis, supports the “10-minute rule” as a maximum, but she thinks there is not sufficient proof that homework is helpful for students in elementary school.

“Correlation is not causation,” she said. “Does homework cause achievement, or do high achievers do more homework?”

In **200 to 250 words**, write a reply to the following question: “Should homework be banned?” Make sure to give arguments that support your opinion. Remember that you are writing to an unknown person, so don’t be impolite. Start your email with the following line: “To whom it may concern.”

*Note: Students in seventh grade are typically 12 or 13 years old.*

#### Vocabulary:

Comprehensive	Including all elements or aspects	Uitgebreid
Correlation	A connection between two or more things	Verband
Inquisitiveness	Being eager for knowledge	Nieuwsgierigheid
Sufficient	Enough; adequate	Voldoende
Causation	The action of causing something	Oorzakelijk verband

### Appendix E – Writing Test B

There is a belief that, as European citizens, we have a responsibility to know and speak at least two languages other than our first language in order to be effective global citizens. Usually, one of these languages will be English. What is your personal view on this?

You have decided to answer this question as part of a contest to win money. To participate, you have to write an email to the person hosting the contest. In **200 to 250 words**, write a reply to the following question: **“Should Dutch students be required to learn at least two foreign languages in school in addition to English and Dutch?”**

If you say yes, please specify which languages you think should be taught and why.

If you say no, explain why and what you think education should focus on instead.

Make sure to give arguments that support your opinion in a clear and understandable way. Remember that you are writing to an unknown person, so don't be impolite. Start your email with the following line: “To whom it may concern.”

*You have found the following text online as you were doing research. Read this text.*

In a world that is becoming ever more globalized, people need to communicate more. Very often, it is assumed that this communication should be facilitated with the use of English... But this preconceived notion needs to be flushed down the toilet and forgotten!

First of all, it is pure imperialism to expect everyone else to conform to one language—in our case, English. On top of that, it is silly to hold the belief that people with all kinds of different language backgrounds will be able to learn English just like that. In fact, it is unfair to expect a Japanese person to learn English with the same kind of ease as a Swedish person because their languages are fundamentally different.

Did you know that learning a foreign language actually broadens your horizon? It opens up your mind to think in new and unforeseen ways. For example, there are languages—such as Spanish—where you don't even need to use personal pronouns such as “I” or “he”. Languages such as Turkish even order their thoughts in a completely different way because the verb always comes last, not second.

In other words, people could all benefit from learning languages other than just English for the purposes of international communication, but also for themselves because it will enrich their lives.

### Vocabulary

To facilitate	To make easier	Vergemakkelijken
Preconceived notion	An idea someone has before checking the facts	Vooraf gevormd idee
To benefit from	To be helped by something / someone	Baat hebben bij
To enrich	To make richer, to make more varied	Verrijken



## Verklaring kennisneming regels m.b.t. plagiaat

### Fraude en plagiaat

Wetenschappelijke integriteit vormt de basis van het academisch bedrijf. De Universiteit Utrecht vat iedere vorm van wetenschappelijke misleiding daarom op als een zeer ernstig vergrijp. De Universiteit Utrecht verwacht dat elke student de normen en waarden inzake wetenschappelijke integriteit kent en in acht neemt.

De belangrijkste vormen van misleiding die deze integriteit aantasten zijn fraude en plagiaat. Plagiaat is het overnemen van andermans werk zonder behoorlijke verwijzing en is een vorm van fraude. Hieronder volgt nadere uitleg wat er onder fraude en plagiaat wordt verstaan en een aantal concrete voorbeelden daarvan. Let wel: dit is geen uitputtende lijst!

Bij constatering van fraude of plagiaat kan de examencommissie van de opleiding sancties opleggen. De sterkste sanctie die de examencommissie kan opleggen is het indienen van een verzoek aan het College van Bestuur om een student van de opleiding te laten verwijderen.

### Plagiaat

Plagiaat is het overnemen van stukken, gedachten, redeneringen van anderen en deze laten doorgaan voor eigen werk. Je moet altijd nauwkeurig aangeven aan wie ideeën en inzichten zijn ontleend, en voortdurend bedacht zijn op het verschil tussen citeren, parafaseren en plagiëren. Niet alleen bij het gebruik van gedrukte bronnen, maar zeker ook bij het gebruik van informatie die van het internet wordt gehaald, dien je zorgvuldig te werk te gaan bij het vermelden van de informatiebronnen.

De volgende zaken worden in elk geval als plagiaat aangemerkt:

- het knippen en plakken van tekst van digitale bronnen zoals encyclopedieën of digitale tijdschriften zonder aanhalingstekens en verwijzing;
- het knippen en plakken van teksten van het internet zonder aanhalingstekens en verwijzing;
- het overnemen van gedrukt materiaal zoals boeken, tijdschriften of encyclopedieën zonder aanhalingstekens en verwijzing;
- het opnemen van een vertaling van bovengenoemde teksten zonder aanhalingstekens en verwijzing;
- het parafaseren van bovengenoemde teksten zonder (deugdelijke) verwijzing: parafrasen moeten als zodanig gemarkeerd zijn (door de tekst uitdrukkelijk te verbinden met de oorspronkelijke auteur in tekst of noot), zodat niet de indruk wordt gewekt dat het gaat om eigen gedachtengoed van de student;
- het overnemen van beeld-, geluids- of testmateriaal van anderen zonder verwijzing en zodoende laten doorgaan voor eigen werk;
- het zonder bronvermelding opnieuw inleveren van eerder door de student gemaakt eigen werk en dit laten doorgaan voor in het kader van de cursus vervaardigd oorspronkelijk werk, tenzij dit in de cursus of door de docent uitdrukkelijk is toegestaan;
- het overnemen van werk van andere studenten en dit laten doorgaan voor eigen werk. Indien dit gebeurt met toestemming van de andere student is de laatste medeplichtig aan plagiaat;
- ook wanneer in een gezamenlijk werkstuk door een van de auteurs plagiaat wordt gepleegd, zijn de andere auteurs medeplichtig aan plagiaat, indien zij hadden kunnen of moeten weten dat de ander plagiaat pleegde;
- het indienen van werkstukken die verworven zijn van een commerciële instelling (zoals een internetsite met uittreksels of papers) of die al dan niet tegen betaling door iemand anders zijn geschreven.

De plagiaatregels gelden ook voor concepten van papers of (hoofdstukken van) scripties die voor feedback aan een docent worden toegezonden, voorzover de mogelijkheid voor het insturen van concepten en het krijgen van feedback in de cursushandleiding of scriptieregeling is vermeld.

In de Onderwijs- en Examenregeling (artikel 5.15) is vastgelegd wat de formele gang van zaken is als er een vermoeden van fraude/plagiaat is, en welke sancties er opgelegd kunnen worden.

Onwetendheid is geen excuus. Je bent verantwoordelijk voor je eigen gedrag. De Universiteit Utrecht gaat ervan uit dat je weet wat fraude en plagiaat zijn. Van haar kant zorgt de Universiteit Utrecht ervoor dat je zo vroeg mogelijk in je opleiding de principes van wetenschapsbeoefening bijgebracht krijgt en op de hoogte wordt gebracht van wat de instelling als fraude en plagiaat beschouwt, zodat je weet aan welke normen je je moeten houden.

Hierbij verklaar ik bovenstaande tekst gelezen en begrepen te hebben.
Naam: Patrick Thiecke
Studentnummer: 3906892
Datum en handtekening: 10-09-2018

Dit formulier lever je bij je begeleider in als je start met je bacheloreindwerkstuk of je master scriptie.

Het niet indienen of ondertekenen van het formulier betekent overigens niet dat er geen sancties kunnen worden genomen als blijkt dat er sprake is van plagiaat in het werkstuk.