The effect of omission of grammar instruction on oral fluency

Comparing the effect of implicit and explicit English language teaching on Dutch secondary school pupils' oral fluency in English

Manouk Bakermans

Master thesis Education and Communication: Education and Communication

June 2017

This page is intentionally left blank.

Master thesis Education and communication: English language and culture

Graduate school of teaching

Utrecht – June 29, 2017

Author: Manouk Bakermans (3978834)

Thesis supervisor: Prof. dr. Rick de Graaff

Second reader: Dr. Elena Tribushinina

This page is intentionally left blank.

Preface

I would hereby like to express thanks to everyone who helped me during the process of writing this thesis. First of all I want to thank Leslie Piggott, without her data and support this thesis would not exist. I also want to thank Pieternel van Braak and Sharyn Moesman for working together on the same data. Among other things, we divided up the transcriptions and fortunately, they agreed to also mark disfluencies in the transcripts, which made it possible for me to include so many participants in this study. I want to express my extreme gratitude to Nivja de Jong for the immense help with all of my questions on fluency and for writing me a Praat script to cut the oral data files. Without her help I would still be editing each sound file individually at this very moment. I want to thank Rick de Graaff for his feedback and guidance and Elena Tribushinina for her tips, advice, and for being my second reader. And last but not least, I also want to thank my boyfriend, Luuk Mandemakers, for his (mental) support.

Manouk Bakermans

June 2017

Abstract

For several decades, researchers have compared the learning effects between form-focused and meaning-focused language education. Whereas one of the main arguments of the supporters of meaning-focused education is that explicit knowledge of grammar rules will impede on the communicative flow, few studies have actually looked at the effect of type of instruction on oral fluency. The present study is part of a bigger longitudinal study that looks at the effect of delaying grammar instruction for Dutch EFL learners in a regular classroom setting. More specifically, the present study investigates the differences in oral fluency between two groups of students in their first year of secondary school (N=319), one of which had received traditional English education (including grammar instruction) for a full year, whereas for the other group the grammar instruction was omitted from the lessons. The following research question was posed: what is the effect of one year of implicit English education on the English fluency of Dutch pupils in their first year of secondary school compared to Dutch pupils who have experienced traditional, English language education in which grammar instruction is included? In line with the arguments raised by supporters of meaning-focused education, it was expected that the group that had not received grammar instruction would outperform the group that had. The students were subjected to a narrative elicitation task. The stories were transcribed in CLAN and fluency measures were calculated by use of Praat and CLAN. Three different fluency variables (1. average number of words; 2. average length of silent pause; 3. breakdown and repair fluency) were gathered for each participant and t-tests were conducted to the compare the two groups on each variable. For the first variable it was found that the group that had received no grammar instruction significantly outperformed the group that had. For the second variable, no significant differences were found. However, for breakdown and repair fluency it was found that the group that had received grammar instruction significantly outperformed the group that had not. It was concluded that the omission of grammar instruction has both a positive and a negative effect on oral fluency.

Table of content

This page is intentionally left blank.

1. Introduction

In the Netherlands, English is an obligatory subject for every secondary school pupil. Unlike many other European countries, there is not a national curriculum that all schools have to follow. The primary objectives and final attainment level goals concerning English language proficiency are nationally established. However, schools are allowed to decide how exactly they teach their students English and how they prepare them for the national exams at the end the secondary school period. Many Dutch schools choose a very consistent way of teaching English: 2 or 3 lessons a week, following a book, teaching many grammar rules and assigning the students long lists of vocabulary to learn by heart.

This traditional way of teaching a language and its rules explicitly is not the only possible type of instruction. Although there is enough evidence to support the claim that language instruction has a positive effect on language learning, there is no consensus about which way of teaching has the greatest effect (De Graaff & Housen, 2009). There are several theories on the most effective ways of teaching a language (Doughty & Varela, 1998; Ellis, 2003; Gass, 1997; Gass & Mackey, 2000; Krashen, 1982; Krashen & Terell, 1983; Long, 1991; Nunan, 1989; Prabhu, 1987). Some scholars have claimed that the language learning process should resemble the acquisition of a first language, and that therefore the learners should not get explicit language instructions but rather should be provided with a communicative setting in which they can pick up on the L2 (Ellis, 2003; Krashen, 1982; Krashen & Terell, 1983; Lyster, 2007; Schwartz, 1993; Nunan, 1989; Prabhu, 1987). Other researchers claim that learning a language does not necessarily mean learning explicit grammar rules and vocabulary, but that language learners can also pick up on these language features by implicit instruction (Long, 1991; Gass, 1997; Gass & Mackey, 2000; Doughty & Varela, 1998). Various ways for realising this implicit way of teaching exist.

Currently, there are three main varieties on explicit and implicit learning that are supported by scholars: Focus on Forms (FonFS), Focus on Form (FonF) and Focus on Meaning (FoM) (Loewen, 2011). FonFS is a more traditional way of teaching a language: there is much focus on grammar rules, and the teacher teaches the students explicitly about the forms of the language. FonM is the exact opposite: it entails no focus on grammar rules or linguistic forms at all. Instead, the lessons are only focused on communicating meaning, and the only language being used is the language needed to communicate successfully. FonF is the middle road: whereas there is some focus on form, this focus is not as explicit as it is in FonFS language teaching. Instead, the main focus remains on meaningful communication of

content, not of form rules. Explicit teacher explanations of grammar rules in a non-communicative setting for instance, or drilling exercises, do not occur in this type of teaching. This variety has by far been interpreted in the most various ways, which is probably because the concept remains very broad. Ellis (2001) distinguished between teacher-initiated FonF, in which the teacher expects students will have trouble with certain linguistic forms and therefore starts teaching these forms implicitly, and reactive FonF, in which the teacher only implicitly teaches about a form if it has wrongly occurred in the students' communicative output. Not all FonF-supporters agree on which type should occur (Loewen, 2011). Another topic on which the scholars are divided concerns the question whether FonF should be intensive, that is "one or two linguistic forms are targeted continually during an activity", or extensive, that is " no single linguistic item is targeted, rather limited attention may be given to a variety of structures" (Loewen, 2011, p. 579).

Most of these language-teaching theories are supported by a thorough body of evidence gathered from research. Norris and Ortega (2000), for instance, provided an overview of research into FonF and FonFS and found that both of these varieties have a positive influence on language learning. However, not each theory is supported thoroughly with empirical research. Whereas research on FonF and FonFS generally seems to result in supportive evidence for these theories, the outcomes of research into FonM are rather inconclusive (De Graaff & Housen, 2009). Moreover, there are some downfalls to most studies that have found evidence for FonM, FonF or FonFS. Most studies use a set-up that does not resemble a language teaching classroom at all. Instead, they use a laboratory to conduct a short-term study focused on one feature. The types of instruction are often carried out during one or two sessions only. Besides that, the tests after the instruction are often flawed, being more in line with one of the two types of instruction. Therefore, outcomes of these studies, and the implications for real-life language education, remain quite questionable. The aim of the present study is to empirically investigate the effect of including or excluding grammar instruction in a real-life language learning setting for a longer period of time. The research question is as follows: what is the effect of one year of implicit English education on the English fluency of Dutch pupils in their first year of secondary school compared to Dutch pupils who have experienced traditional, explicit English language education?

2. Overview of academic discussion

2.1 Focus on form, forms or meaning

A much debated question in research is which type of instruction yields the best results in language learning. Is it the implicit FonM, the implicit FonF or the explicit FonFS? Supportive empirical evidence exists for both implicit and explicit types of instruction, and no firm conclusion has been drawn thus far. A few of the main studies and their outcomes in the field are listed below.

2.1.1 Focus on form(s)

As has been mentioned above, some scholars claim that it is best to focus on language rules and forms. In a review article about the effectiveness of type of instruction, Norris and Ortega (2001) found that the average effect sizes of both FonF and FonFs instructional types were large compared to FonM instructional types. This means that the studies which they used for their review provided outcomes in favour of FonF and FonFS instruction. There was no significant difference between the effect sizes of FonF and FonFS, indicating that at the moment of their review, none of the two types of instruction appeared to be superior to the other (Norris & Ortega, 2001). However, the researchers did have some points of criticism towards most studies that found a large effect size for the more explicit forms of education: the tests after the instruction mostly tested explicit, declarative knowledge, which was not in any way communicative (Norris & Ortega, 2001). Therefore, the outcome might be a distorted picture of the actual effect. Besides that, the "implicit" types of instructions were usually much more restricted in comparison with the more explicit forms (Norris & Ortega, 2001). Researchers often focused on creating a well-balanced and diverse type of FonF or FonFS instruction, whereas the way the implicit instruction was designed was one-sided and sometimes did not even involve the target construction (Norris & Ortega, 2001). Like Norris and Ortega (2001), Spada and Tomita (2010) conducted a meta-analysis of studies investigating the difference in learning gains between explicit and implicit instruction. They found that, in general, the effect sizes of the explicit types of instructions were larger than those of the implicit types of instruction (Spada & Tomita, 2010). Spada and Lightbown (2008) presented an overview of theories and empirical studies that have supported the claim that explicit language teaching has a positive effect on language proficiency.

These large overview studies already suggest that a lot of studies have pointed out an advantage of explicit instruction over implicit instruction. A few individual studies will be mentioned here. Laufer and Girsai found evidence that a group of students that received the

form of education that was most form-focused outperformed two groups of students who received more meaning-focused education on a translation task (2008). Likewise, Ellis, Loewen, and Erlam (2009) found that students who received explicit corrective feedback outperformed the group of students that received implicit corrective feedback on the targeted grammar feature (regular past tense –ed). A study by Alanen (1995) showed that explicit instruction of a grammar rule before giving students enhanced input (meaningful L2-input in which the target item occurs more often than in general L2-input) increased students' ability to recall and apply the rule as opposed to input enhancement without the instruction beforehand. Scott (1989) looked at the different effect sizes of implicit and explicit instruction on two French grammar structures (subjunctive and relative pronoun) and found that the group of students receiving explicit instruction outperformed the group of students receiving implicit instruction.

2.1.2 Focus on Meaning

Other researchers state that implicit language teaching is more effective than explicit teaching. One of the 'founding fathers' of this position is Stephen Krashen. In 1982 Krashen introduced his theory in his book *Principles and practice in second language acquisition*. In this book he explains that students do not need explicit grammar teaching but rather need comprehensible input, that is, meaningful input that is one step further than the current level the students are at (Krashen, 1982). That way, the students acquire implicit knowledge about the language and learn how to use it in naturalistic settings (Krashen, 1982). The explicit rules taught in traditional education can only be used in optimal situations in which students are allowed time to think about their output and its form, and these situations are not very likely to happen in reality, according to Krashen (1982). An example of such a situation would be an untimed grammar test, in which students are allowed enough time to access their explicit knowledge of grammar rules (Krashen, 1982). Likewise, Schwartz argues that the 'data' learned from explicit language instruction does not lead to actual linguistic competence, which is needed to communicate in a second or foreign language (1993).

A number of studies compared the difference in learning benefits between classroom settings and natural settings, and these natural settings mostly entailed that the 'natural setting students' stayed in the country of the language they were learning for some time (Freed, Segalowitz & Dewey, 2004; Llanes & Serano, 2014). The outcomes of these studies generally indicate that the latter setting has a more positive effect on language learning than the former (Freed, Segalowitz & Dewey, 2004; Llanes & Serano, 2014). However, most of these results

could also have been caused by a larger amount of input and more opportunities for practice (Llanes & Serano, 2014). Many other studies also found the language benefits of studying abroad for some time (e.g. DeKeyser, 2007; Dywer & Peters, 2004; Llanos & Muñoz, 2009). However, the current study is limited to classroom teaching only. Nonetheless, this argument is often used to emphasize the benefits of more implicit, communicatively meaningful instruction, which resembles the naturalistic setting.

A study by Reinders and Ellis (2009) went into the difference in learning results on the grammatical feature 'negative adverbs' between a group that received enriched input only (meaningful input in which the target feature occurs many times) and a group that first received an explicit instruction to notice the target feature and then was subjected to enriched input. They found that the group that received the noticing instruction did not have any benefits over the group that only received enriched input (Reinders & Ellis, 2009). However, they did not find any evidence that the implicit group outperformed the explicit group either (Reinders & Ellis, 2009). Ayoun (2004) found similar results in a study into the effect of different types of instruction on the acquisition of past tense grammar structures: the type of instruction did not have much influence on the students' success in acquisition.

A study by Macaro and Masterman (2006) looked at the different language outcomes between a group of students who received an "intensive burst of explicit instruction" before starting their study of French at a UK university and a group that did not (p. 297). The results indicated that the group who did receive this burst outperformed the control group on grammar tests, but not on translation or free composition tasks (Macaro & Masterman, 2006). This is in line with Krashen's theory that students do not experience any benefits of their explicit knowledge in communicative situations which was, in this case, writing.

What the researchers who support FonM education have in common is their belief that knowledge learned during explicit language instruction cannot be used in real communicative situations (Krashen, 1982; Schwartz, 1993). The main argument behind this claim is that accessing the explicit knowledge takes too much time, which disrupts the communicative flow (Krashen, 1982; Schwartz, 1993). It can be argued that if these students intend to apply their explicit knowledge in a communicative setting, their fluency would be disrupted, because of the excessive amount of time it takes to access their knowledge. Surprisingly, currently there are no studies that specifically focused on the effect of implicit versus explicit instruction on a language learner's oral fluency. This paper focuses on the effect of explicit and implicit teaching on fluency.

2.1.3 Effect of implicit versus explicit education on oral proficiency

Most of the studies that compare implicit and explicit education look at specific grammar items. It seems surprising that there are not many studies that explicitly go into the difference between how implicit and explicit instruction influence oral proficiency. This appears to be an extremely interesting gap in research, especially since the FonM supporters claim that explicit knowledge cannot be used in communicative setting, of which the ability to communicate orally is probably the first and foremost representative. The reason for this focus on grammar instead of oral proficiency is probably because it can easily be measured whether or not the acquisition of a grammar item did or did not take place, whereas overall oral proficiency is a lot broader and not measured in a standard way. Some studies did look at whether or not subjects orally produced the targeted grammar items correctly. For example, a study by Hernandez (2008) went into the difference in effect of explicit and implicit language teaching on L2 students' use of discourse markers in an oral test. The experimental group received explicit instruction on discourse markers besides input flood, opportunity for practice and corrective feedback, whereas the control group received the same types of instruction except for the explicit instruction. The experimental group outperformed the control group, indicating that, considering oral discourse markers, the inclusion of explicit instruction has a benefit over limiting teaching to implicit instruction only.

2.1.4 Attitudes towards implicit and explicit education

All of these studies into the most effective ways of teaching are very important for our understanding of what types of instruction would be ideal to use in real-life settings. However, many of the research findings into second language acquisition have little or no effect on the language teaching practice at all (Larsen-Freeman, 2015). Therefore, it is important to keep in mind the attitudes of the gatekeepers of education. If their attitudes towards a change that is suggested by research are not positive, the change is not likely to happen. Burgess and Etherington (2002) did research into English language university teachers' attitudes towards explicit and implicit grammar education. They found that the teachers preferred explicit teaching of grammar over implicit teaching (Burgess & Etherington, 2002). Besides that, the teachers involved in the study also indicated that they noticed that their students also preferred getting explicit rather than implicit education (Burgess & Etherington, 2002). In other words, although there is academic support for both explicit and implicit education, the attitudes of teachers and students appear to favour explicit education.

2.2 The interface debate

Another concept on which scholars do not seem to agree, which is tightly linked to implicit and explicit education, concerns the interface debate. The interface debate concerns the mobility of explicit and implicit knowledge. Again, there are two opposing views and many scholars that base their opinion somewhere in the middle. On the one hand some scholars advocate that implicit and explicit knowledge are fluent constructs that do not necessarily remain implicit or explicit forever. On the other hand opposing scholars propagate that explicit knowledge could never become implicit knowledge and vice versa.

Usually, the supporters of FonM instruction adhere to the non-interface position. These scholars argue that usable language knowledge can only be acquired in communicative situations. The lack of mobility between the two types of knowledge implies that explicit knowledge, most often learned in classroom settings, cannot be accessed during such communicative situations, in which only implicit knowledge acquired during meaningful communication can be used (Krashen, 1981; Schwartz, 1993; Truscott, 1996; Paradis, 2009).

Not all researchers that do suggest there is interaction between implicit and explicit knowledge agree on how this interaction takes place. Some scholars claim that though the interaction between the two types of knowledge is restricted, and most explicit knowledge is not accessible in communicative situations, L2 learners can make some explicit knowledge implicit by practicing it a lot (N. Ellis, 2005; 2006; 2007; R. Ellis, 1994; 2005; 2006). This is called the weak interface position. According to the weak interface position "explicit knowledge may help where implicit knowledge fails" and "everything is learnable" (Han & Finneran, 2004, p. 4). Other scholars are of the opinion that, normally, implicit knowledge can become explicit knowledge via practice and automatisation (DeKeyser, 2007). This is called the strong interface position. One of the assumptions this position relies on is the noticing hypothesis, which argues that people cannot learn anything without explicitly paying attention to it (Han & Finneran, 2004, based on Schmidt, 1990; 1995; 2001).

Recently, Han and Finneran (2014) took a closer look at the interface debate and the empirical evidence that seems to exist for all points of view. They argued that since evidence exists for all of the positions, all of these might co-exist (Han & Finneran, 2014). The results of their case study indicated existence of all of the three interface positions in one person's interlanguage at one specific moment in the learning process (Han & Finneran, 2014). Han and Finneran (2014) argue that the lack of consensus between scholars can be ascribed to this presence of all of the interface positions in every language learner. If this statement is correct,

it may be argued that the same goes for type of instruction: in that case some features can be learned explicitly whereas others would have to be acquired from input.

2.3 Fluency in foreign language research

2.3.1 The concept 'fluency'

Various ways of defining the concept "fluency" exist. In popular speech, fluency generally refers to overall language proficiency. It is not uncommon to hear people comment on their own language proficiency that they speak a language "fluently', by which they indicate that they are able to communicate without much difficulty and on an overall high level. The term can even be used on a curriculum vitae to indicate a level of foreign language proficiency. In such instances, 'fluently' indicates a proficiency level that is almost as high as near-nativeness and clearly much higher than merely basic language proficiency. This interpretation of the concept is called global fluency, and is different from componential fluency (Koponen & Riggenbach, 2000). Componential fluency is a narrowed down definition, in which fluency is not equal to proficiency but is part of it (Koponen & Riggenbach, 2000). Koponen and Riggenbach state about most descriptions of componential fluency that it is associated "with flows, currents, or motion" of speech (2000, p. 8).

Componential fluency is one of the language proficiency components represented in the Common European Framework of Reference (CEFR). The CEFR is used in The Netherlands to determine the national primary objectives and the final attainment level goals for English. The CEFR subdivides language proficiency into six varying levels, ranging from A1/A2 (basic user) via B1/B2 (independent user) to C1/C2 (proficient user), A1 being the lowest and C2 being the highest level. Roughly, the aspects of language that are taken into account to establish these levels are reading, listening, speaking, and writing. Within spoken language use, the CEFR distinguishes between different aspects of speech, such as range, accuracy, fluency, interaction and coherence. From the presence of this diversity of aspects to be taken into account it can already be concluded that fluency is a separate aspect of speech that is not the same as general language proficiency. Besides that, the fluency indicators used in the descriptions of the levels all concern flow of speech and ease of production, and there is even mention of disfluencies such as pauses, false starts, and repairs. Hence, it can be stated that since the Dutch school system uses the CEFR as an indicator for the nationally established levels, componential fluency is also an aspect for which objectives are set nationally.

In research, the two-fold distinction between global and componential fluency often becomes clear in the way researchers use the concept in their studies. On the one hand, some researchers interpret or use the concept of fluency as global fluency. There are studies in which fluency is regarded as a broad concept related to "a person's general language proficiency, particularly as characterized by perceptions of ease, eloquence, and 'smoothness' of speech or writing (Housen & Kuiken, 2009). Kormos and Dénes (2004), for example, compared the different measurable aspects related to fluency to subjective fluency ratings by native and non-native language teacher judges. They did not give these judges any indication of how fluency should be rated, and it turned out that the judges took various aspects of language proficiency into account that, in many other studies, are strictly separated from fluency, such as accuracy and lexical diversity. This study is similar to an earlier study by Lennon (1990), which followed the same set-up and suggested that the factors underlying general fluency might be subdivided into "'core' and 'peripheral' fluency variables" (Lennon, 1990). Core, in this definition, relates to the componential fluency, whereas peripheral entails the broader sense of language proficiency. Freed did a similar study and also found that judgments by the native speaker and non-native speaker judges were related to "a multitude of linguistic, psychological, and sociolinguistic factors, each of which contributes to the full construct of fluency" (Freed, 2000, p. 261). Researchers such as Chambers (1997), Guillot (1999), Freed (2000) and Koponen & Riggenbach (2000) discuss the different possible ways of defining the concept "fluency" and also take these broad definitions into account. Hilton (2008) suggests that lexical competence plays a very important part in how fluent L2-speakers sound.

These broad interpretations of the concept are not in line with other definitions that are strictly refined to measurable subdimensions, hence to componential fluency. De Jong (2016) explains that disfluencies occur when a speaker cannot keep up with the articulation of what he or she is trying to say while conceptualizing what to say or formulating how to say it (p. 203). As a consequence, the speaker becomes disfluent, possibly by slowing down, being silent for a while, using a filled pause, or repeating something (De Jong, 2016). Another form of disfluency is reformulating or correcting something, which happens when a speaker notices an error in his or her speech and decides to correct it (De Jong, 2016). Componential fluency could roughly be subdivided into speed fluency, which relates to how quickly someone produces speech units, breakdown fluency, which relates to the number and duration of pauses, and repair fluency, which relates to the number of reparations a speaker makes in his or her speech (Lambert & Kormos, 2014; Tavakoli & Skehan, 2005). Some studies

investigated the link between one or several of these measurable subdimensions and global fluency (Kormos & Dénes, 2004; Lennon, 1990; Freed, 2000). Many other studies simply calculate fluency as a measurement of speed fluency, breakdown fluency, repair fluency or a combination of these components (e.g. De Jong et al., 2012; Gan, 2012; Housen et al., 2011; Lambert & Kormos, 2014; Larsen-Freeman, 2006). Componential fluency can be calculated from speech data once it has been defined which components should be taken into account. For an overview of possible calculations to arrive at a measure of componential fluency, see De Jong (2016), table 1.

2.3.2 Disfluencies in speech
Whereas global fluency is more or less equal to L2 proficiency, componential fluency only plays a small part in L2 proficiency. In fact, it has been argued that disfluencies are not necessarily indicators of lower proficiency as such (De Jong, 2016). Although "fluency is at least partly dependent on L2 proficiency", it is also part of fluent L1 speech (De Jong, 2016, p. 206, based on De Jong et al., 2013; De Jong et al., 2015; Riazantseva, 2001). In fact, some studies have found that disfluencies can actually carry meaning (Clark & Fox Tree, 2002; Clark, 2000; Collard et al., 2008; De Jong, 2016). De Jong mentions that several studies have found that after disfluencies in native speech, the content of speech is often more complex (De Jong, 2016). Consequently, a disfluency might make the listener aware of an upcoming passage that requires the listener to pay closer attention. Disfluencies, in other words, do not necessarily affect the performance of the speaker, but can be part of a certain speaking style and add meaning to a message.

However, in studies that looked into which aspects predict ratings of L2 proficiency best, fluency, among other proficiency aspects, turned out to be one of the most important indicators (De Jong, 2016, based on Iwashita et al., 2008). In other words, though fluency is part of natural L1 speech, it is also a good indicator of foreign language proficiency. Moreover, fluency has been found to improve along with proficiency in L2 learning (De Jong, 2016). So whereas some part of fluency may depend on personal speaking style, at least some part of fluency measures or judgements are related to actual L2 proficiency. De Jong (2016) states that (inverse) articulation rate is most strongly related to L2-proficiency and least to personal speaking style. In order to not measure the personal fluency (hence the general fluency of a speaker, in L2 as well as L1), it may be best to use a metric that is hardly related to personal speaking style in the present study.

2.3.3 Effects on fluency

Many researchers have included or specifically focused on the effect of different variables on fluency of foreign or second language learners in their study (Foster &Skehan, 1996; Kormos & Dénes, 2004; Yuan & Ellis, 2003). Examples are Yuan and Ellis (2003) who studied the effect of pre-task and online planning on language learners' fluency. They found that pre-task planning has a positive effect on the fluency of language learners when performing a monologic oral production task (Yuan & Ellis, 2003). An article by Foster and Skehan (1996) also reported that planning has a positive effect on fluency, and reported that this effect is stronger in a narrative task and a decision making task than in a task which concerns giving personal information. Gilabert (2007) also found that pre-task planning has a positive effect on fluency. Besides that, task complexity concerning context also influences fluency: if a task demands from a subject that he or she talks about the past rather than about the present, this influences the subject's fluency negatively (Gilabert, 2007). A study by De Jong et al. (2012) found that task complexity influences different aspects of fluency in different ways, and that task complexity influences the fluency of native speakers and non-native speakers in different ways. Michel, Kuiken, and Vedder (2007) compared complex and less complex tasks and found that the complexity of tasks affects fluency negatively. The study also looked at the difference in fluency in monologic and dialogic tasks and found that subjects were more fluent in dialogic tasks than in monologic tasks (Michel, Kuiken & Vedder, 2007). Llanes and Muñoz (2009) reported that study programs in the country of the FL have a beneficial effect on language learners' fluency.

3. Hypotheses

The current study investigates the effect of implicit learning on the oral fluency of Dutch secondary school students of all levels. As has been mentioned before, the research question investigated in this paper is as follows: what is the effect of one year of implicit English education on the English fluency of Dutch pupils in their first year of secondary school compared to Dutch pupils who have experienced traditional, explicit English language education?

The present study is restricted to the calculative concept of componential fluency. Following the theories that support FonM, it is expected that the students that have received implicit language education, will be more fluent than their peers who have received traditional explicit education. According to Krashen's theory, the explicit knowledge of grammar rules creates a filter. This filter is not useful in many situations, because accessing the explicit

knowledge takes up a lot of time, so that in actual communication the "flow of speech" would be interrupted. This could implicate that students would have more disfluencies such as (filled or silent) pauses, repetitions, and repairs in their speech. Following De Jong's (2016) explanation of disfluencies, the students might slow down and become disfluent the moment they are trying to formulate how to say what they intend to say. The disfluency would then occur because the information is not accessed rapidly enough, hence impeding the flow of speech. If a student's speech rate is faster than the speed at which the outcome can be formed in the mind, disfluencies occur (De Jong, 2016). Another way this filter would affect fluency is when a speaker notices an error based on his/her explicit knowledge and starts correcting it, which would relate to the repair fluency. Therefore, the hypothesis of this study is that the group of students that received implicit education will be more fluent in the oral task than the group of students that received explicit education.

Hypothesis 1. The intervention will have a positive effect on the average number of words uttered per minute of speaking time.

Hypothesis 2. The intervention will have a negative effect on the mean length of silent pause.

Hypothesis 3. The intervention will have a negative effect on the average number of pauses, filled pauses, repetitions and revisions per minute of speaking time.

These hypotheses are in line with the main assumption that knowledge of grammar rules negatively influences the speaking fluency of second language learners. Therefore, each hypothesis is stated in a way that reflects the expectation that the intervention group will outperform the control group on fluency.

4. Method

4.1 Participants

The current study uses data gathered in a larger, longitudinal study by Piggott (in preparation). Piggott investigates the effects of a two-year grammar instruction delay on Dutch EFL learners' proficiency.

All participants (N=319) studied at the same Dutch secondary school in the Netherlands. The participants in this study were in the end of their first year during the gathering of the data. All students were between the age of 11 and 13. The intervention group

did not receive any grammar instruction and the control group did receive grammar instruction throughout their first year of education. The textbook used during the classes was *More,* a method designed by Cambridge (Puchta, Stranks, Gerngross, Holzmann, & Jones, 2008). For the intervention group, the pages and text blocks that contained explicit grammar explanations were deleted from the textbook. There were two teachers who only taught the intervention group classes, two teachers who only taught the control group classes and one teacher who taught classes to both groups. The two groups did not receive their education during the same year; there was a one-year delay for the intervention group. This should not affect the results in any way since there is no reason to believe that this made the population any different. The data for the control group was gathered at the end of the academic year 2014-2015 and the data for the intervention group was gathered at the end of the academic year 2015-2016.

The study's population consisted of a selection of students from ten intervention group classes and ten control group classes. Both groups consisted of a diversity of classes representing a large part of the regular Dutch school system. Five different levels of education can be distinguished and are represented in both the intervention and the control group. The levels are vocational education (*vmbo-tl*, N=32), vocational education/higher general secondary education (*vmbo-tl*/*havo,* N=64), higher general secondary education (*havo,* N=74), higher general secondary education/pre-university education (*havo/vwo,* N=64) and pre-university education (*vwo*, N=86). These levels differ in how students are prepared for their tertiary education and/or their profession. Some educational levels prepare students for vocational professions, whereas other levels are oriented towards university or university of applied sciences. One pre-university level student in the control group had to be excluded from the study for sound-technical reasons.

4.2 Materials & procedure

An oral proficiency test existing of two parts was conducted at the end of the first year of the students' English education. The test was conducted as part of the students' regular curriculum; the grades would count towards their final mark for the English course for that school year. The tests were recorded and the students were informed about this beforehand. The students were not aware that their test output would be used for research. However, parents were informed for ethical reasons, and were given the opportunity to withdraw their child's data from the research database.

The students were tested in pairs in a quiet room in the school during regular school time. For each pair of students, one teacher would be present, who would lead the test and explain the tasks. It was made sure that the teacher would not be the students' regular teacher, since the teacher-student bond might influence the student's performance. In the first part of the test, the students were asked to answer some simple questions about their daily life and preferences. This was done in order to make them feel at ease and to get "warmed up" for the second part of the test. The first part was not taken into account for the grade and was not used for research. The second part consisted of a narrative elicitation task. One of the participants received a storyboard containing several pictures and was told he/she should tell a coherent story based on the pictures. The storyboards were shortened versions of the stories *Frog, where are you?* (Mayer, 1969) and *A boy, a dog, a frog and a friend* (Mercer & Mayer, 1993). The students were allowed some time to look at the pictures before they got started with telling the story. After the first student was done, the second student received the storyboard with the remaining story and was told to do the same thing. During the graded tasks, the teacher sometimes made encouraging remarks such as "go on", or "good". Sometimes the teacher would even help the student out. This only occurred rarely, if the student really could not go on with the task without help.

For each participant, the story was transcribed using the CLAN programme of CHILDES (MacWhinney, 2000). Three MA-students including the author of the current study divided up the participants' stories to be transcribed. Filled pauses, revisions and repetitions were marked in the script using symbols that CLAN recognises. Both when words and sentences or fragments were repeated or revised, this was coded as one repetition or revision. This was done in order to get an actual reflection of the number of times a student would produce disfluencies as opposed to a reflection of how many actual words the student repeated or revised. Once the stories were transcribed, the transcriptions were converted into *praat.texgrid* files, which can be read by the programme Praat (Boersma & Weenink, 2007). Subsequently, both the sound files and the transcripts were cut so that only the participant's speech and the pauses he/she made were left. This was done by use of a Praat script created for this purpose. The Praat script first cuts off the beginning and ending in which the teacher and participants are having conversations that are not of importance to the results. Then the script cuts the moments during the participant's story in which the teacher speaks, as well as other sounds/noises that are irrelevant for the data. These moments had been coded under the 'teacher-tier' (TEA), so that the script could recognise which parts should be deleted from the sound file. During the next step, the sound was filtered by use of Praat's 'remove noise'

function. After the noise had been removed, De Jong and Wempe's (2013) *Praat Script Syllable Nuclei* was used to detect the total duration in seconds, the phonation time in seconds and the number of pauses for each student. The silence threshold was set at -25 decibel. The minimum dip between peaks was set at 4 decibel. The lower minimum pause duration threshold was set at 300 milliseconds, following guidelines proposed by De Jong and Bosker (2013).

Besides these steps in Praat, the total number of words, filled pauses, repetitions and revisions in each transcript was calculated by use of the MOR-function, Flucalc-function, and freq-function of CLAN. The data were put together in one data list containing for each participant the following variables: number of words uttered, the number of pauses, the complete duration of the sound file and the speaking time within this sound file, and the number of repetitions, revisions and filled pauses.

To be able to control for education, an extra value was added to each participant's values. The different school levels were each given a value, the lowest level being the lowest value, and the highest level being the highest value. The first two school levels were combined into one level for statistical reasons, due to the low number of students enrolled in vocational education. The new list was then as follows: 1 = vocational education/higher general secondary education (*vmbo-tl/havo, N=96*), 2 = higher general secondary education (*havo, N=74*), 3 = higher general secondary education/pre-university education (*havo/vwo*, N=64) and 4 = pre-university education (*vwo*, N=86).

Then, the independent variable of intervention/control group was added to each participant's list of values. Each participant that was in the control group got the value 0, whereas each participant that was in the intervention group had the value 1. This was used later on to check for the differences between the intervention and control group.

Lastly, the three dependent fluency variables were calculated and added to the list. The first variable, the mean number of words per minute of speaking time, was calculated by dividing the number of words by the speaking time and then multiplying it by 60 (nwords/phonation_sec x 60). The second variable, the mean length of a silent pause, was calculated by dividing the silent time by the number of pauses ([duration_sec– phonation_sec]/npause). The last variable that was calculated was the breakdown and repair fluency. This was calculated by first finding the average number of silent pauses, filled pauses, repetitions and revisions per minute of speaking time. This was done by adding up the number of silent pauses, filled pauses, repetitions and revisions and dividing this by four. Subsequently, the number was divided by speaking time, and then multiplied by 60

([[npause+filledpauses+repetitions+revisions]/4]/speaking time x 60). Besides that, for the third variable, a reliability analysis was carried out to check for internal consistency between the four different breakdown and repair items used. Though the Cronbach's alpha was not extremely high at α=0.63, it is considered acceptable. However, if the variable number of silent pauses is deleted from the scale, the Cronbach's alpha increases to α=0.76. Although there might be statistical reason to omit this variable from the scale, this step was not taken. In her study into the best fluency measures, De Jong (2016) proposes a subset of fluency measures that encompasses all measures of fluency that are currently used in the field. She explains that by both including the number of silent pauses and the mean length of a pause, none of the measures measure the same construct, yet all of the possible fluency aspects are taken into account (De Jong, 2016). Therefore, the number of silent pauses was kept as item within the breakdown and repair fluency variable.

Subsequently, a factor analysis was done in order to test whether it was likely that one underlying dimension was represented with the subset of variables chosen. All four items loaded on a single factor with an eigenvalue of above one. In total, the factor explained 65.28% of the variance in the four items. Considering the extensive study into fluency measures by De Jong (2016), this underlying factor can be referred to as breakdown and repair fluency.

4.3 Analytical strategy

For the relation between intervention group and all dependent variables, independent variable t-tests were carried out in order to compare the two groups. A regression analysis was carried out to check for education effects. Independent sample t-tests were carried out for each level of education in order to check for effect of education on the effect of the intervention on the fluency variables.

5. Results

The aim of this study was to compare the intervention group and the control group on fluency measures. In the results section, differences between the intervention group and the control group are examined based on 1) the number of spoken words, 2) the length of a silent pause and 3) breakdown and repair fluency. The results are interpreted in light of the theoretical background of this paper. A Table with the mean scores and standard deviations is presented below.

Table 1. Descriptive statistics of all variables (N=319).

	M	SD
Dependent Variables		
Number of words per minute	199.65	233.93
Length of silent pauses	.85	.22
Breakdown and repair fluency	22.04	8.52
Independent variable		
Intervention Group[1]	.50	.50
Control Variable		
Education	3.46	1.29

[1]Reference category = group that received grammar instruction.

Mean number of words per speaking time minute

An independent variable t-test was conducted to calculate if there was a significant difference in mean number of words between the two groups of students. The mean number of words differed significantly between the two groups. On average, the intervention group uttered 224.62 words per minute, and the control group 174.52. As the hypothesis was unilateral, the p-value of the two-tailed t-test was divided by two. The difference in means therefore turned out to be significant, with $t(161)=-1.93$, $p=0.028$ (one-tailed). The effect size (Cohen's $d = 0.22$) can be considered relatively small (Lakens, 2013). Therefore, the hypothesis can be accepted and it can be concluded that the intervention had a positive effect on the mean number of words per speaking time minute.

Mean length of silent pause

An independent variable t-test was conducted in order to calculate any significant differences between the intervention group and the control group. The intervention group scored slightly lower on this variable, but the results were not significant. The second hypothesis should therefore be rejected. The effect size (Cohen's $d = 0.02$) can be considered very small (Lakens, 2013). The mean number of pauses was very high for both groups, at respectively 0.85 pauses per minute for the intervention group and 0.85 pauses per minute of speaking time for the control group. The lower threshold of 300 milliseconds appears to account for this high mean. De Jong and Bosker (2013) investigated what lower threshold should be used in second language research, and found that if the threshold is set higher than 250 to 300 milliseconds, the number of silent pauses and their length will confound. Therefore, the lower threshold was kept at 300 milliseconds.

Breakdown and repair fluency

The breakdown and repair fluency measures were subjected to an independent variable t-test. The results showed that the control group outperformed the intervention group significantly: on average, the control group had a significantly lower breakdown and repair fluency score (M=20.91, SD =4.20) than the intervention group (M=23.16, SD=11.18), with t(203)=2.37, p=.02 (two-tailed). The effect size (Cohen's *d* = 0.27) is relatively small again (Lakens, 2013). It can therefore be concluded that the intervention did not have a decreasing effect on the breakdown and repair fluency of the students, as was expected. This entails that, on average, it was the intervention group instead of the control group that produced more disfluencies in speech. Consequently, hypothesis 3 must be rejected.

Table 2. Independent sample T-tests tests for mean differences in control group and intervention group (N=319).

	T-Test			
	Mean Difference[a] (SE Difference)	DF	T	*d*
Number of words	50.10 (26.00)	161	-1.93*	0.22
Mean length of pause	-.004 (.025)	317	.18	0.02
Breakdown fluency	-2.24 (.94)	203	2.37*	0.27

*p<.05
[a]Reference category = group that received grammar instruction.

The educational level was not constant within the two groups. Fluency is related to overall language proficiency (De Jong, 2016, based on Iwashita et al., 2008; Oh, 2006). Since it may be expected that the higher-level students have a higher proficiency of English, regression analyses were carried out to control for educational level. No significant effect was found, which indicates that the educational level did not influence the students' fluency. The effect size of education and intervention on number of words (R^2 = .01) is relatively small, the effect size on mean length of silent pause (R^2 = 0.003) is very small and the effect size on breakdown fluency (R=0.2) is relatively small as well. Furthermore, after including education in the regression analyses, there was still a positive relation between the intervention group and breakdown and repair fluency, with *B*=2.10, p=0.016 (one-tailed). Likewise, the average

number of words per minute of speaking time is also still significant, with B=50, p=.031 (one-tailed).

Table 3. Regression analyses controlling for effects of education on fluency variables.

	Model 1 *Linear Regression* **DV: Number of Words**	Model 2 *Linear Regression* **DV: Mean length of pause**	Model 3 *Linear Regression* **DV: Breakdown fluency**
	B (SE)	B (SE)	B (SE)
Intervention[a]	50.00 (26.67)*	-.01 (.03)	2.10 (.97)*
Education	-.19 (10.35)	-.01 (.01)	-.28 (.38)
N	319	319	319
F-test (df)	1.84 (2;316)	.44 (2;316)	3.07 (2;316)*
R^2	.01	.003	.02

*p<.05
[a]Reference category = control group. .

6. Discussion and conclusion

Krashen and other supporters of FonM education have stated that there are, if any, very few benefits to educating L2 grammar to foreign language learners (Ellis, 2003; Krashen, 1982; Krashen & Terell, 1983; Lyster, 2007; Nunan, 1989; Schwartz, 1993; Prabhu, 1987). The grammar rules learned would not be applicable in any real communicative situation. Moreover, if foreign language learners try to access their explicit grammar knowledge, this would have a negative effect on the communicative flow. The results of this study differed per variable and the overall outcome provides both support for and against this theory. It appears that the omission of grammar instruction influences the average number of spoken words per minute positively, since the intervention group, on average, uttered 50 words more per minute of speaking time. However, the students that did receive grammar instruction significantly outperformed the students that did not receive grammar instruction on breakdown and repair fluency. However, these results, especially those concerning the third variable, should not be seen as hard evidence. The reason for this is that it could be concluded from a pre-test concerning both general language proficiency in the mother language and some aspects of

English that the group that received grammar instruction had a language proficiency advantage over the group that did not receive grammar instruction (Piggott, in preparation).

The main aim of this study was to find out whether or not the omission of grammar instruction would benefit students' L2 oral fluency. For the first fluency variable, the average number of words, the intervention group significantly outperformed the control group (H1). On average, the control group produced 175 words per minute of speaking time whereas the intervention group produced an average number of 225 words per minute of speaking time. This difference is strikingly big. This result adds to the previously discussed theory that knowledge of grammar influences the communicative flow negatively. Following this theory, the most immediate interpretation would be to state that the control group would utter fewer words per minute because they would be busier with thinking about grammar rules, therefore speaking slower, whereas the intervention group could easily speak without a constant urge to access explicit knowledge. Another explanation might be that because the intervention group students had more time to practice their oral production during classes, they were more trained to speak quickly without much hesitation. Yet another possible interpretation of the results is that the intervention group students are more confident in their oral production, simply because they are unaware of any errors they could make. According to the teachers involved in Piggott's larger study, the intervention group students were less preoccupied with worrying about grades and making mistakes, which might have been a consequence of the omission of grammar lessons and the omission of grammar sections in the tests these students had to make. This could lead to the students worrying less during the oral exam as well, resulting into a more confident, less anxious flow of words. On the other hand, it is also quite clear that the average number of words differed a lot, not only between the groups, but also within the groups. A possible interpretation of this outcome could be that fluency as measured by the average number of words might be less compliant to grammar instruction than it was expected on the basis of previous research. In both groups, there were students who uttered many words per minute of speaking time, and students who uttered very few words per minute, which can be concluded from the high standard deviations in both groups. It is also quite possible that other factors existed that influenced this fluency measure. Examples of other possible explanations of the large variance within groups are general language aptitude, vocabulary knowledge, and L1 fluency of the students. Lastly, it must be noticed that average number of words per minute is a less representative measure of fluency than average number of syllables per minute. The syllable count outcome of the Praat Script Syllable Nuclei (De Jong & Wempe, 2013) was tested but proved to be unreliable (average deviation of 19% from

a syllable count from the transcript on a 5% data check), which is why the current study had to use average number of words instead of syllables. It would be interesting to see if the results would change if syllables were used instead of words, though there is no reason to believe this would lead to different results other than the higher representational value of the metric (De Jong, 2016). Average number of words per speaking minute is a variable that can be categorised under speech rate, which, according to Kormos and Denes' (2004) study into the way fluency measures are related to perceived fluency, is the best predictor of perceived fluency by native and non-native speaker judges. It would be interesting for future research to do similar studies but instead of focussing on of number of words, these could focus on average number of syllables. Besides that, it would be interesting to see if a study with a larger group of participants would have the same results. A longitudinal study into a similar group of participants would be interesting in order to find out if this positive effect of omission of grammar teaching only exists in the beginning stage of language learning or if it stretches to different learner levels as well.

   For the second fluency variable, the mean length of silent pause, no significant difference between the intervention and control group was found (H2). The fact that there was no significant difference between the two groups is interesting on itself. As has been stated before, the FonM-supporting researchers would expect the intervention to cause students to have shorter pauses on average. The reason for expecting this, as has been mentioned before, is that the students do not have to dig deep into their explicit knowledge of English grammar while speaking because they do not have this explicit knowledge. Therefore, the students would be quicker to pick up where they have left after a short pause, whereas students who are still pondering on how the grammar rule goes will stay quiet for a longer period of time. The lack of supportive evidence for this hypothesis in this study supports the point of view of the FonF and FonFS supporters. Whereas the FonM criticise teaching grammar for having a negative effect on fluency, this effect was definitely not found for this particular aspect of fluency in this study. A possible interpretation of this could be that maybe the control group students did not spend much time contemplating on the correct grammar rule after all. It may also be the case that the students did have to think of grammar rules but remembered them quite quickly, leading to no significant differences in the mean length of silent pause. The mean length of silent pause variable only reflects the average length of a pause, and not the number of pauses in total. Therefore, a combination between this variable and the third variable seems of importance to see if the control group students were more disfluent in other terms. Therefore, a more inclusive contemplation on this will follow in the interpretation of

the third variable. Of course, this study is limited as it only sheds light on the effect of omission of grammar instruction on L2 students' oral fluency. It would be interesting to compare the two groups on grammatical accuracy as well, since the finding of the current study could also be interpreted as the students not being able or trying to access their explicit grammatical knowledge. If the control group students indeed did not take the time to access their explicit knowledge of English grammar, the group's fluency would not be interrupted but the students' grammatical accuracy would not be higher either. It would be interesting to combine grammatical accuracy and fluency in future research.

On breakdown and repair fluency, the control group students significantly outperformed the intervention group students (H3). This means that not only is the hypothesis disproved, the opposite was found to be true. Considering the second variable, for which it was found that on average, the control group did not produce longer pauses than the intervention group, this finding is interesting. Apparently the control group had even fewer disfluencies in their speech than the intervention group, which makes it unlikely that they produced more pauses. Therefore, a likely explanation is that the control group produced slower speech, in which the students produced fewer words but also fewer disfluencies. Possibly, the control group students were less likely to repeat or repair their utterances, because they were more familiar with grammar rules and therefore could confidently tell the story without thinking too much and revising their output too much. Another possibility is that the control group simply needed less correcting and repeating because they would make fewer mistakes that needed correcting, since they explicitly knew the grammar rules. Yet another possible explanation can be derived from the finding that the control group averagely produced fewer words per speaking minute than the intervention group. Both variables were controlled by minute of speaking time, and a simple explanation could be that, since the control group uttered fewer words, the control group students therefore also uttered fewer disfluent words (filled pauses, repetitions and revisions). Moreover, the difference between the two groups was a lot higher concerning the variable average number of words than the breakdown and repair fluency, which makes it even more likely that the disfluency effect that has been found had to do with a higher number of words. Although these, and many other explanations can be sought for this finding, another important explanation that should be mentioned is the 0-level of both groups. Although there was no similar English narrative elicitation task before the year of intervention/control education took place, there were other tests that can arguably provide some information about the groups' average language aptitude and knowledge. At the start of the study, the control group had higher scores on L1

proficiency, on vocabulary, and on listening skills in English (Piggott, in preparation). Previous research has indicated that L1 fluency is an important indicator of L2 fluency (De Jong, 2016, based on Iwashita et al., 2008; Oh, 2006). Therefore, it may be best to not regard the results found for breakdown and repair fluency as solid evidence for the notion that excluding grammar instruction from L2 education has a negative effect on breakdown and repair fluency.

6.1 Limitations

This study does not provide clear evidence that omission of grammar instruction leads to a more fluent oral production of students' L2. Instead, evidence was found both for and against the FonM point of view that grammar instruction omission leads to more fluent L2 communication. The reliability of the results should be considered in the light of the research details and previous research.

First, the teachers present in the school during the gathering of data and the researchers transcribing the stories noticed that some children seemed to know many more words in comparison to other students. It appeared as if these students were told in advance about the content of the story and therefore were able to look up some words and come to the test a bit more prepared than other students. Since the number of participants was so high, the communication between students could not be prevented. Consequently, the students who had their tests on later days or at later moments in the day, might have had an advantage over students who were tested in the beginning of the day or testing period. Besides this possibility of oral transfer of information, the fact that the students were tested in pairs leads to the same adverse point in the research: students could have picked up on information given by other students. These limitations may have influenced the degree of variation between the individual students. However, since this goes for both groups, it does not affect the general results in any way.

Second, some teachers were more eager to help out students than others. Since the distribution of teachers present at the tests differed between the two groups, one of the two groups might have had an advantage over the others. No tests were done to check this.

Third, the participants in this study were all living in the same province in the Netherlands (Overijssel) and all attended the same school. This makes the results of this study narrow and causes reason to question the generalizability of the conclusions to other (English) language learners in the Netherlands and language learners in general.

Fourth, because of the omission of grammar instruction, intervention group teachers had extra time for other activities during the lessons. An explorative bachelor thesis in which Piggott's data were used found that teachers in the control group averagely spent 15% of their classroom time on speaking proficiency, whereas the teachers in the intervention group averagely spent 20% on speaking time (Abrahamse, 2016, bachelor thesis; Piggott, in preparation). The effect this might have caused would be in favour of the intervention group's fluency, but regarding the fact that no statistically significant evidence was found that suggested that the intervention group outperformed the control group, this turns out to not be crucial to the study's results.

In conclusion, this study first provided an overview of previous research from which it could be concluded that overall, FonF an FonFS approaches were found to be more successful in teaching foreign languages than FonM approaches, for which there appears to be a lesser body of academic evidence. In order to further go into the FonM supporters' argument that students' communicative flow would be disrupted if they were taught grammar rules, this study compared the fluency results of a group of students that did receive grammar instruction in regular classroom setting for a year and a group of students that did not receive grammar instruction in regular classroom setting for a year. The outcomes varied in whether or not they provided support for an advantage of FonM-based education over FonF or FonFS-based education on fluency. For one fluency variable, average number of words per speaking time minute, it was found that the omission of grammar instruction led to a higher number of words per speaking time minute. For another fluency variable, breakdown and repair fluency, it was found that students who had received grammar instruction outperformed students who had not. It remains the subject of future research to find out why the omission of grammar instruction seems to have a positive effect on one variable of fluency, and a negative effect on the other. Besides that, future research could focus on combining fluency variables with accuracy measurements in order to find out how the two aspects of language proficiency interact.

References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. *Attention and awareness in foreign language learning* (pp. 259-302).

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.

Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. Available from http://www.praat.org/

Burgess, J., & Etherington, S. (2002). Focus on grammatical form: Explicit or implicit? *System, 30*(4), 433-458.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535-544. Retrieved from: http://dx.doi.org/10.1016/S0346-251X(97)00046-8

De Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. *The handbook of language teaching* (pp. 726-755).

De Jong, N.H. (2016). Fluency in second language assessment. In D. Tsagari & J. Banerjee (Eds.), *Handbook of Second Language Assessment* (pp. 203-218). Mouton de Gruyter.

De Jong, N. H. and Bosker., H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (ed.), *Proceedings of Disfluency in Spontaneous Speech Diss 2013* (pp. 17–20). Stockholm: Royal Institute of Technology (KTH)

De Jong, N., Groenhout, R., Schoonen, R. & Hulstijn, J. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior, *Applied Psycholinguistics 36*(2), 223-243.

De Jong, N., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In A. Housen, I. Vedder & F. Kuiken (Eds.), *Dimensions of L2 Performance and Proficiency.Complexity, Accuracy and Fluency in SLA* (pp. 121-142). Amsterdam: John Benjamins Publishing Company.

De Jong, N., Steinel, M. P., Florijn, A. F., Schoonen, R. & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language, *Applied Psycholinguistics 34*, 893-916.

Jong N.H. de & Wempe T. (2008), *Praat script syllable nuclei*.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J.

Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge: UP.

DeKeyser, R. M. (2007). Skill acquisition theory. In B. Van Patten and J. Williams (Eds.), *Theories in second language acquisition* (pp. 97-113). Mahwah: Erlbaum.

DeKeyser, R. M. (2007). Study abroad as foreign language practice. *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208-226).

Dwyer, M. M., & Peters, C. K. (2004). The benefits of study abroad. *Transitions abroad, 37*(5), 56-58.

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition 27*(2), 305–352.

Ellis, N. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics 27*(2), 164–194.

Ellis, N. (2007). SLA: The associative-cognitive creed. In B. VanPatten and J. Williams (eds.), *Theories in second language acquisition: an introduction*, 77-96. Cambridge: UP.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (ed.), *Implicit and explicit learning of languages* (pp. 79-114). London: Academic Press.

Ellis, R. (2001). Introduction: Investigating form‐focused instruction. *Language learning, 51*(1), 1-46.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford: University Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: a psychometric study. *Studies in Second Language Acquisition 27*(2), 141–172.

Ellis, R. (2006). Modeling learning difficulty and second language proficiency: the differential contributions of implicit and explicit knowledge. *Applied Linguistics 27*(3), 431–463.

Ellis, R., Loewen, S. & Erlam, R. (2009). Implicit and explicit corrective feedback and the acquisition of L2 grammar. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 303-332). Bristol: Multilingual matters.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second language acquisition, 18*(3), 299-323.

Freed, B. F. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 243-265). Michigan: UP.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in second language acquisition, 26*(2), 275-301.

Gan, Z. (2012). Complexity measures, task type, and analytic evaluations of speaking proficiency in a school-based assessment context. *Language Assessment Quarterly, 9*(2), 133-151.

Gass, S. (1997). *Input, interaction, and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and [+/-Here-and-Now]: Effects on L2 oral production. In M. Del Pilar García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44-68). UK: Multilingual matters ltd.

Guillot, M. N. (1999). *Fluency and its teaching* (Vol. 11). UK: Multilingual Matters.

Han, Z., & Finneran, R. (2014). Re-engaging the interface debate: Strong, weak, none, or all? *International Journal of Applied Linguistics 24*(3), 370-389.

Hernández, T. (2008). The effect of explicit instruction and input flood on students' use of Spanish discourse markers on a simulated oral proficiency interview. *Hispania 91*(3), 665-675.

Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *The language learning journal 36*(2), 153-166.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461-473.

Housen, A., Schoonjans, E., Janssens, S., Welcomme, A., Schoonheere, E., & Pierrard, M. (2011). Conceptualizing and measuring the impact of contextual factors in instructed SLA - the role of language prominence. *IRAL - International Review of Applied Linguistics in Language Teaching, 49*(2), 83-112.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied linguistics*, *29*(1), 24-49.

Koponen, M. & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5-24). Michigan: UP.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164. Retrieved from: http://www.sciencedirect.com/science/article/pii/S0346251X0400017X

Krashen, S. (1981). *Second language acquisition and second language learning.* New York: Pergamon Press.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom.* Oxford: Pergamon.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in psychology, 4,* 1-12.

Lambert, C. & Kormos, J. (2014) Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics, 24,* 1-9.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590-619.

Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching, 48*(2), 263-280.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, *29*(4), 694-716.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language learning*, *40*(3), 387-417.

Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference?. *System, 37*(3), 353-365.

Llanes, À., & Serrano, R. (2014). The effectiveness of classroom instruction 'at home' versus study abroad for learners of English as a foreign language attending primary school, secondary school and university. *The Language Learning Journal*, 1-13.

Loewen, S. (2011). Focus on form. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning, Vol. 2* (pp. 576 – 592). London: Routledge.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach.* Amsterdam: John Benjamins.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52) Amsterdam: John Benjamins.

Macaro, E., & Masterman, L. (2006). Does intensive explicit grammar instruction make all the difference? *Language Teaching Research, 10*(3), 297-327.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Mahwah, NJ: Erlbaum.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *IRAL-International Review of Applied Linguistics in Language Teaching*, *45*(3), 241-259.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*, 417–528.

Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta‐analytic review. *Language learning, 51*(1), 157-213.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Oh, S. (2006). *Investigating the relationship between fluency measures and second language writing placement test decisions.* Doctoral dissertation, University of Hawaii.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam, John Benjamins.

Piggott, L. E. B. (in preparation). First meaning then form: A longitudinal research on the effects of delaying the explicit focus on form for young adolescent L2 learners. Doctoral Dissertation. Utrecht University.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Puchta, H., Stranks, J. Gerngross, G., Holzmann, C.; Lewis-Jones, P. (2008). *More! Level 1 Student's book*. Cambride: University press.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in second language acquisition*, *15*(02), 147-163.

Reinders, H., & Ellis, R. (2009). The effects of two types of enhanced input on intake and the acquisition of implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 281-302). Bristol: Multilingual matters.

Riazantseva, A. (2001). Second language proficiency and pausing a study of Russian speakers of English. *Studies in Second Language Acquisition*, *23*(4), 497-526.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158.

Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning* (pp. 1-64). Honolulu: UP.

Schmidt, R. (2001). Attention. In P. Robinson (ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition 15*(2), 147–163.

Spada, N., & Lightbown, P. M. (2008). Form-Focused Instruction: Isolated or Integrated? *Tesol Quarterly*, *42*(2), 181-207.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language learning*, *60*(2), 263-308.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-273). Philadelphia : John Benjamins Pub. Co.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning 46*(2), 327–369.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied linguistics*, *24*(1), 1-27.