

**Motivation and Second Language Acquisition:
A Study on the Relation between Motivation and Oral Proficiency**

by

Django de Smet

4136721

BA Scriptie Engelse Taal en Cultuur,

Universiteit Utrecht 2016

9332 words

Supervisor:

dr. Elena Tribushinina (Universiteit Utrecht)

Aril 6, 2016

Table of Contents

1. Theoretical Background	2
2. Methodology	12
3. Results	26
4. Discussion	29
5. Conclusion	37
References	38
Appendices	42

Abstract

Motivation has been shown to be an important factor in the process of second language (L2) acquisition. Previous research has indicated a relation between motivation and general L2 proficiency. However, less is known about the relation between motivation and specific aspects of L2 proficiency (e.g. oral proficiency). Furthermore, no previous studies have investigated motivation in relation to L2 proficiency in the Netherlands. Therefore, this study tried to fill this gap by administering a motivation questionnaire and an oral proficiency task to Dutch high school students learning English as a second language in an instructed setting. The results of this study suggest that a relation exists between motivation and L2 oral proficiency for Dutch learners of English. In particular, the results showed that participants with a high language anxiety performed worse on an oral proficiency task than participants with a lower language anxiety. Furthermore, learners who identify themselves more with the L2 culture performed better than learners who identify themselves less with the L2 culture. Overall, the results seem to confirm that motivation plays an important role in the L2 acquisition process.

1. Theoretical Background

1.1 Introduction

Previous research has shown that there are substantial individual differences in achievement between learners acquiring a new language on top of their native language. Some learners go on to reach a native-like level of proficiency, seemingly without much effort, while others struggle to reach a satisfactory level of proficiency. Research on individual differences in second/foreign language (L2) acquisition has shown that motivation is an important determinant of L2 achievement (e.g. Gardner, 2010). Given that there are many different reasons as to why an individual might be motivated to learn a second language, it is extremely difficult to provide a satisfactory definition of motivation: motivation is necessarily multi-faceted and complex (Crookes & Schmidt, 1991; Gardner, 2010; Gardner, 2007; Kleinginna Jr & Kleinginna, 1981). It is beyond the scope of this paper to provide an overview of all definitions and types of motivation, the focus of this paper is on those constructs of motivation that might be relevant to L2 learning and proficiency in classroom settings. I use the term L2 learning to refer to the learning of a language, other than an individual's native language, in an instructed setting, while at the same time acknowledging the fact that this might also be the third or fourth language for some learners.

In this paper, the relation between motivation and one specific type of second language proficiency, namely oral proficiency, will be investigated. In the remainder of this section, I will provide a brief overview of the various constructs of motivation, discuss previous studies that have investigated motivation in relation to second language proficiency and oral proficiency, and present the research questions that will be investigated in this paper. In section 2, I will present the methodology that was used to gather the data and explain how the data was analysed. Subsequently, in section 3, I will present the results from the analyses.

Finally, in section 4, the major findings of this study will be discussed and compared with previous findings from similar studies.

1.2 Constructs of motivation

Broadly speaking, two major approaches toward attitude and motivation can be distinguished. The most recent one is the L2 motivational self system developed by Dörnyei (Dörnyei, 2005; Dörnyei, 2009). In this approach, the focus is on the psychological concept of ‘selves’ in relation to second language learning. Most notably, Dörnyei (2009) proposes three components of the L2 motivational self system: (1) *Ideal L2 self*, which refers to the person we would ideally like to be, (2) *Ought-to L2 self*, which refers to the traits we would need to accomplish a goal, and (3) *L2 learning experience*, which refers to the learning environment. According to Dörnyei (2005; 2009) these components determine to what extent a second language learner is motivated to learn the L2.

The other approach, and perhaps the most widely studied one, is the socio-educational approach proposed by Gardner (1985). Initially, the main focus of this approach was on the concept of integrativeness. Integrativeness refers to the desire of a language learner to identify with the cultural group of the L2 language (Gardner, 2010). Over the years more constructs were considered in this approach and eventually this resulted in the development of the Attitude/Motivation Test Battery (AMTB), which is a questionnaire consisting of multiple statements that measure several constructs of motivation and attitude (see Gardner, 2010 for a detailed overview). Crucially, the AMTB was developed for use in language situations in which the L2 that is being acquired is English.

The socio-educational approach, in the form of the AMTB, distinguishes six constructs of motivation: (1) *Integrativeness*, this construct is measured by considering the integrative orientation (i.e. reasons to learn the L2 with the purpose of communicating with

native speakers), attitudes towards English speaking people, and interest in foreign languages. (2) *Attitudes towards the learning situation*, this construct considers attitudes towards the language teacher as well as attitudes towards the language course. (3) *Motivation*, this is measured by taking into account motivational intensity (i.e. how much effort does the language learner invest in learning the L2), desire to learn English (i.e. to what extent does the language learner want to learn the L2), and attitudes toward learning English (i.e. to what extent does the language learner enjoy learning the L2). (4) *Language anxiety*, this refers to the anxiety a language learner might experience when learning or using the L2 in the classroom and outside the classroom. (5) *Instrumentality*, this construct focuses on the material, or practical, gain that might motivate a language learner to learn the L2. (6) *Parental encouragement*, this is measured by assessing the role of the language learners' parents or caregivers.

In the present study, Gardner's (1985) socio-educational approach will be adopted because the data that will be analysed was gathered as part of a larger research project by Piggott (in preparation), in which the AMTB was used to measure motivation and attitude. The shortcomings of this approach will be discussed in section 4.5 below.

1.3 Motivation and general second language proficiency

Previous studies on the influence of motivation on the language proficiency of L2 learners have reported intriguing results. Gardner and Lambert (1959) found a relation between motivation to learn French and proficiency scores in French in Canadian high school children learning French as a second language. In another study, Gardner, Day, and MacIntyre (1992) found a positive correlation between motivation and the ability to learn word pairs in a second language. In yet another study, Gardner, Tremblay, and Masgoret (1997) also found positive correlations between several measures of motivation and performance of university students

on various French proficiency tests. On the other hand, Papi and Teimouri (2014) did not find a relation between motivation and level of proficiency for Iranian learners of English in an instructed setting. The difference between this study and previously mentioned studies, however, is that Papi and Teimouri (2014) used self-ratings as a measure of proficiency whereas other studies have commonly used standardised language tests. This indicates that motivation might not affect all types of language proficiency and operationalisations of language proficiency to the same extent. Nevertheless, most of the previous findings suggest that there is a relation between motivation and L2 proficiency.

Another important finding is that correlations between language proficiency and motivation differ as a function of the context where testing is done. For example, results may vary between learners from Quebec, a bilingual city where learners might be acquiring two languages simultaneously, and learners from other non-bilingual cities where a second language is acquired during high school. Gardner and Lambert (1959) found that motivation was consistently related to language proficiency but different correlations were found between subtypes of motivation and language proficiency in different settings (i.e. settings that differ from each other regarding language context) (as cited in Gardner, 2010). This might indicate that findings regarding motivation and language proficiency are not easily generalisable to other countries, cultures, and language situations. Therefore, it is important to investigate whether a relation between motivation and L2 proficiency exists in other countries and language contexts as well.

In a classroom setting, motivation might play an especially important role since learning a second language is inherently different from learning other school subjects (e.g. mathematics) (Gardner, 2010). Acquiring another language provides learners with much more than simply the ability to speak another language, it enables them to adopt something from another culture and offers them new ways to express themselves. For other school subjects,

however, these cultural and expressive components are not as significant (Williams, 1994; Gardner, 2010). This shows that motivation is a powerful concept and one which deserves to be investigated with regard to L2 learning in instructed settings.

1.4 Motivation and oral second language proficiency

Research on the relation between motivation and specific types of proficiency (e.g. oral proficiency) is limited. Achieving oral proficiency in a second language is often seen as the most challenging part of the L2 acquisition process and not many L2 learners ultimately acquire a native-like oral proficiency. Some studies have investigated whether a relation between motivation and oral proficiency exists. Hernández (2010), in a study on the effects of motivation on oral proficiency in a study-abroad context, found that students with a high integrative motivation (i.e. wanting to learn the L2 to be able to interact with the L2 culture) had more contact with the L2 than students with a lower integrative motivation. Furthermore, the highly motivated learners outperformed the learners with a lower motivation on an oral proficiency task. In another study, Hernández (2006) also found integrative motivation to be a significant predictor of L2 oral proficiency. These findings seem to suggest that there is a relationship between motivation and oral proficiency.

However, Hernández (2006; 2010) did not take into consideration other factors that might have influenced the learners' proficiency and motivation. Language anxiety, for example, has been shown to negatively influence motivation as well as proficiency (Gardner, 2010). Language anxiety, as identified by Gardner (2010) in his socio-educational model of language acquisition, is the anxiety that a learner might experience when using the L2 either in an instructed setting (i.e. language class anxiety) or outside of the classroom (i.e. language use anxiety). Papi and Teimouri (2014), in a study on motivational types, found that the group with the highest overall motivation did not have higher L2 proficiency scores than the group

with the second highest overall motivation. They suggest this could be due to the fact that the second highest motivation group showed significantly less language anxiety than the highest motivation group. Thus, these findings indicate that anxiety is an important aspect of motivation that could have strongly affect L2 proficiency. Language anxiety might play an especially important role with regard to L2 oral proficiency since achieving oral proficiency in another language necessarily requires the learner to actively use the language. Therefore, it is necessary to investigate the relation between multiple constructs of motivation and oral proficiency.

Research on motivation in relation to proficiency in English as a second language in the Netherlands seems to be lacking. An interesting finding from a recent study by Elzenga and de Graaff (2015) is that Dutch high school students are, in general, more motivated to learn English as a second language than French. However, since the goal of the study by Elzenga and de Graaff (2015) was to find differences in motivation between single-language education and dual-language education students, the relation between motivation and language proficiency was not investigated.

1.5 The present study

To date there are only some studies on the relation between motivation and specific types of proficiency. Moreover, in the case of oral proficiency, these studies either only investigated proficiency and motivation in study-abroad contexts (Hernández, 2010), or considered only one construct of motivation (i.e. integrativeness) (Hernández, 2006). Furthermore, there are currently no studies on the relation between motivation and oral L2 proficiency, or any other type of L2 proficiency, in the Netherlands. This study aims to fill these gaps by investigating the relation between multiple constructs of motivation and oral proficiency in learners of English as a second language in an instructed setting in the Netherlands.

While earlier studies mostly used more formal types of assessment (e.g. objective tests) or broad measures of proficiency (e.g. course grades), in this study, oral proficiency will be measured by administering a story-telling task to the participants. Story-telling tasks are natural speaking tasks and have proven to be a reliable measure of linguistic skills (Iwashita, 2010; Duinmeijer, de Jong & Scheper, 2012). Furthermore, a translated version of the international form of the Attitude/Motivation Test Battery will be used to assess motivation (Gardner, 2010).

1.6 Research questions and hypotheses

To investigate the relation between motivation and second language proficiency in the Netherlands, the following research question and sub-questions will be addressed in this paper:

Research question: Is there a relation between motivation and level of oral proficiency in Dutch learners of English as a foreign language in an instructed setting?

Given that previous research has consistently indicated a relation between motivation and overall L2 proficiency it would be reasonable to assume that a relation also exists between motivation and oral proficiency. However, as indicated above, previous measures of proficiency were obtained either through objective tests or by using final course grades.

Therefore, investigating only one specific type of proficiency (i.e. oral proficiency) might yield different results. Since there is only limited research on oral proficiency in relation to motivation, I can only make assumptions regarding possible outcomes. In personal observations I have often noticed that people with a high oral proficiency in their L2 express considerable interest in the country (and its inhabitants) where their L2 is the native language. Oftentimes they are also involved with the language outside the instructed setting, for

example, in the form of being a member of a drama club that performs in the L2 or having relatives and friends living in the L2 country. These kinds of orientations (i.e. reasons to learn a language) is what Gardner (2010) refers to in his socio-educational model as integrativeness (see above) and he argues that individual differences in integrativeness in turn lead to motivational differences.

Most importantly, however, almost all research on L2 proficiency and motivation has in common that highly motivated learners score higher on proficiency measures than less motivated learners. Based on these findings, I expect to find a relation between motivation and oral proficiency in the sense that highly motivated learners score higher on a task measuring oral proficiency than less motivated learners.

Sub-question 1: Are some constructs of motivation more strongly related to oral proficiency than others?

Previous research indicates that integratively motivated language learners have higher L2 proficiency scores than learners that are instrumentally motivated (i.e. wanting to learn a language as a means of achieving individual goals) (Gardner & Lambert, 1959; Hernández, 2010; Hernández, 2006; Spolsky, 1969). However, other research suggests that instrumentally motivated learners perform better on L2 proficiency tasks than integratively motivated learners (Lukmani, 1972).

In addition to the distinction between integrativeness and instrumentality, there are other dimensions of motivation that have to be considered. Given that a story-telling task requires the learner to tell a story based on pictures on the spot, it might be that some learners feel anxious at the thought of speaking up in their L2 in front of others. Gardner (2010) refers to this aspect of motivation as language anxiety. This type of motivation could be considered negative since a high language anxiety generally leads to decreased performance on language

tasks. This motivational aspect might not be as relevant when proficiency is measured by objective language tests because in those cases language use is non-oral and the language learner could therefore feel less anxious.

In contrast, language anxiety could play a significant role in the performance of learners on an oral task. In this light, several studies have found that anxiety is a good predictor of second language proficiency (see MacIntyre & Gardner, 1991 for a review of the literature on language anxiety). Based on these findings, I expect that some subtypes of motivation (e.g. language anxiety and integrativeness) are more strongly related to oral proficiency than others.

Sub-question 2: Are some aspects of oral proficiency more strongly related to motivation than others?

Oral proficiency, as measured by a story-telling task, can be further divided in smaller aspects such as vocabulary, grammar, fluency, and functional adequacy. It might be that motivation is more strongly related to some aspects than others. For example, previous research has indicated that learners with a high level of language anxiety are less fluent in spoken L2 than learners with a lower level of language anxiety (MacIntyre & Gardner, 1994). Furthermore, Kleinmann (1977) found that highly anxious learners produced different grammatical constructions than less anxious learners (as cited in Horwitz, Horwitz & Cope, 1986). Based on these findings, I expect that some aspects of oral proficiency (particularly fluency and grammar) are more strongly related to motivation than others.

Sub-question 3: Are there differences in oral proficiency and motivation between learners from different educational levels?

In the present study, the oral proficiency and motivation of learners from different educational levels will be investigated. At the end of primary school, Dutch children have to complete a standardised test developed by the CITO (Centraal Instituut voor Toets Ontwikkeling/Central Institute for Test Development) comprised of questions covering multiple subjects such as geography, English, and history. Based on the student's score on this test and the primary school teacher's recommendation (independent of the test result), the student then moves on to one of the various educational levels in secondary school (see section 2.1). Therefore, it is relevant to investigate whether participants enrolled in higher educational levels outperform their peers enrolled in lower educational levels in the first year of secondary school and whether they differ with regard to motivation.

Given that part of the score on the CITO test is determined by a student's English skills, I expect that there will be differences in performance on an oral proficiency task between participants of different educational levels and, more specifically, that students enrolled in higher educational levels will outperform students enrolled in lower educational levels. Regarding motivation, it is difficult to provide a hypothesis because there is no research to date that investigates differences in motivation between students of different educational levels.

Sub-question 4: Are there differences between the ratings of teachers and the ratings of an independent researcher?

Two sets of oral proficiency ratings are compared in the present study: one by the secondary school English teachers, and one by an independent researcher (i.e. the author of this paper). This provides me with the opportunity to investigate whether the ratings of teachers are comparable to the ratings of an independent researcher. Upshur and Turner (1999) found teacher raters to be relatively consistent and comparable to other teacher raters.

However, less is known about the reliability between teacher ratings and those of an independent researcher. It is important to know whether teacher ratings are comparable to those of an independent researcher because many studies on L2 acquisition use teacher ratings as objective measures of L2 proficiency. In the present study, the teachers rated students they did not teach themselves and therefore their ratings might be more objective as opposed to teacher ratings of students that are taught by the teacher rater. Consequently, I expect to find no differences between the teacher ratings and the ratings of an independent researcher.

2. Methodology

2.1 Participants

For the current study, data on oral proficiency and motivation gathered as part of a longitudinal study conducted by Piggott (in preparation) was analysed. Piggott's study investigates the effects of delaying explicit focus on form for young L2 learners of English in the Netherlands. In Piggott's study, data on oral proficiency was gathered by administering a story-telling task to students in their first year of Dutch secondary school. Data on motivation was gathered by administering a translated version of the international form of the Attitude/Motivation Test Battery (AMTB) to the same group of Dutch students before Christmas break in their second year of secondary school (Gardner, 2005; Gardner, 2010). During this first stage of testing, participants were between 11 and 13 years old, went to the same school, and lived in the same province (Overijssel) in the Netherlands.

All participants (N=120) were enrolled in one of five educational levels: 1. Vocational education (VMBO-TL; N = 8) 2. Senior general secondary education (HAVO; N = 10) 3. VMBO-TL/HAVO (comprised of students both from VMBO-TL and HAVO; N = 34) 4. Pre-university education (VWO; N = 40) 5. HAVO/VWO (comprised of students both from HAVO and VWO; N = 28). These educational levels differ with regard to the extent to which

each track prepares a student for future education. For example, the VMBO-TL level generally aims to prepare students to enter vocational education after graduation while the VWO level aims to prepare students for a university education.

The parents of all students received a letter in which they were informed of the study and were able to object to the use of their child's data for research purposes; only the parents of one child decided to do so. Initially, 253 students participated in Piggott's study, however, some participants did not complete either the oral proficiency task or the AMTB, or their oral task was not recorded (due to technical problems), and therefore, in the current study, data of 120 students was analysed.

2.2 Materials and procedure

2.2.1 Oral proficiency

In Piggott's study, the participants were told they were going to be tested on their oral English skills. They were not told this was part of a study but rather that it counted towards their final grade for the English course. Thus, the participants were unaware of the fact that their scores would be used for a study. This was done to ensure that they were as motivated for this test as for any other regular test that counts towards their final grade.

As a measure of oral proficiency, a story-telling task was administered. In this narrative task, participants were presented with pictures from one of two picture books: *Frog, where are you?* (Mayer, 1969) and *A boy, a dog, a frog, and a friend* (Mercer & Mayer, 1993). A subset of pictures (N = 15 for *Frog, where are you?* and N = 13 for *A boy, a dog, a frog, and a friend*) was selected from these picture books while ensuring that the pictures still told a logical story. The performance of the participants on the narrative task was rated by an English teacher other than their regular English teacher (from the same secondary school). This was done to ensure that the ratings were as objective as possible.

Testing took place during regular school hours and participants were tested in a quiet room in pairs. Before the test session began, participants were told they were being recorded and received a paper with trivial questions they had to ask each other in English (e.g. ‘What do you like to do after school?’). This had the purpose of making the participants feel at ease with the situation as well as provide them with the opportunity to speak English before the actual assessment. This part was not taken into consideration for the grading of the task. After several minutes of conversation, participants were told the actual test would now begin. When the pair was presented with the pictures from the picture book, the participant that went first had some time to look at the pictures before he or she had to start. The only instruction given by the teacher was that the pictures formed a story and that he/she had to tell the story. While each participant was telling the story, the teacher sometimes provided encouraging remarks such as: ‘go on’, ‘very good’, etc. After the first participant, the second participant received the same instructions and had to tell the picture story based on the other picture book. Usually, testing took no longer than 20 minutes for each pair. The teacher that administered the test also rated the participants’ performance during the same test session.

Teachers graded the participants’ narratives by scoring their performance on four different aspects of oral proficiency: vocabulary, grammar, fluency, and functional adequacy. For each aspect, teachers gave a score between 1 and 5, with 1 being the lowest possible score and 5 the highest (see section 2.2.2 below). These scores will be used for the current study as a measure of oral proficiency. As part of a paid research assistantship, together with another research assistant, I anonymised all available data by giving a random number to each participant and deleting their names. Subsequently, I put all the results in an Excel file with participants being ordered by educational level. As previously mentioned, participants whose score on the oral task was not available, due to absence during testing, were excluded from the study.

2.2.2 Oral proficiency assessment

Since all participants were rated by different teachers, I rated all available oral proficiency tests as part of data analysis for this thesis. This had the benefit of providing me with the ability to calculate a measure of interrater reliability (see section 2.3.2 below). I followed the same procedure as in Piggott's (in preparation) study with regard to scoring the participants' narratives. I graded the narratives by giving a score between 1 and 5 on four different aspects of oral proficiency. These scores refer to the Common European Framework of Reference for Languages (CEF) levels, with a score of 1 indicating a CEF level of A1, and a score of 5 indicating a CEF level of B1. Moreover, the rating scale is also based on a rating scale for oral proficiency created by the Center for Applied Linguistics (see Appendix A). Parts of this scale were translated and some additional information was added to make the distinction between the levels clearer as can be seen in table 1.

The four different aspects of oral proficiency that were scored are vocabulary, grammar, fluency, and functional adequacy. As can be seen in table 1, the score for vocabulary indicates to what extent a participant is able to communicate in the target language, by using a wide variety of words, phrases, and chunks, without having to fall back on words from the L1. The grammar score indicates to what extent a participant is able to create correct grammatical sentences, use some forms of coordination and subordination, and attempt more complex L2 structures. Fluency refers to a participant's speech rate and his/her ability to complete an utterance without hesitations. The score for functional adequacy indicates to what extent a participant has successfully completed the task at hand (i.e. to what extent the participant is able to tell a story that accurately describes the pictures and whether the story is comprehensible to a native speaker).

First, I scored the performance of several participants and discussed my scores with Piggott. Since she also was one of the raters in her own study, she was able to guide me in the rating process. Then, after making sure I understood the rating procedure, I scored each narrative by carefully listening to the recordings (see Appendix B for an excerpt of speech taken from a recording and an explanation of the rating given to this particular participant's performance).

Table 1
Oral Rating Scale (Replicated with permission from Piggott (in preparation)).

Aspect	Score				
	1 (A1.1)	2 (A1+)	3 (A2)	4 (A2+)	5 (B1)
Vocabulary	Produces only isolated words combined with words from L1 and/or incorrect use of words. Unable to produce all words needed to communicate the entire message. Clearly struggles.	Produces mainly isolated words and phrases. Phrases and chunks can occur but are diverse and not always target like.	In addition to isolated words, uses phrases/chunks. Use is quite diverse but not always accurate or sophisticated. Able to communicate with words.	Is able to use a wide range of words, phrases, chunks to communicate. Predominantly accurate production, however, not always target like/sophisticated.	Uses a wide range of words, phrases, chunks to communicate. There is a balanced production of target like choice of words. Errors are due to attempts to produce more complex words or word combinations.
Grammar	Only uses short simple sentences. A lot of sentences are not target like structures.	Predominantly uses one (simple) sentence structure. Variation can be present but is unsuccessful. Grammar is not yet sufficiently sophisticated to convey the intended message correctly.	Can create correct grammatical sentences. Structures can be very diverse but not always successful. Structures are still predominantly short and simple.	Mainly creates correct grammatical sentences. Length of sentences is not only short and there are some forms of coordination and subordination.	Creates correct grammatical sentences and has a balanced use of different sentence types (diverse). Grammatical errors are due to attempts at more complex L2 structures.
Fluency	Speech rate is low with long pauses. Words, phrases or clauses are repeated without any modification.	Speech rate is low but there are few pauses and false starts. Is able to complete an utterance. Hesitations do occur frequently.	Speech rate varies but on average is high with few pauses or hesitations. Student is mainly able to complete a sentence in one run.	Speech rate is high. Pauses and hesitations rarely occur. Student sometimes reformulates what is said and adds new information when doing so.	A native speech rate. Rarely any pauses. Reformulations do occur but are a modification of what was said before.
Functional Adequacy	Overall the task was unsuccessfully completed. Content scarcely conveys the intended message.	Task was almost successfully completed. Content sometimes met the expectations. The general message is starting to be comprehensible.	Task was overall successfully completed. Content sometimes didn't meet the expectations. The general message is mainly comprehensible.	Task was successfully completed. Content met the expectations of the task and were comprehensible enough in the ears of a native speaker.	Task was very successfully completed. Content was relevant and met the expectations. Utterances were predominantly comprehensible.

2.2.3 *Motivation*

In Piggott's study, data on motivation was gathered by administering a translated version of the international version of the AMTB (Gardner, 2010). This questionnaire has been proven to be a reliable measure of various constructs regarding attitude towards language learning and motivation (Gardner, 2005; Gardner, 2010). The AMTB consists of six major constructs that reflect different aspects of attitude and motivation (see table 2). Each construct is assessed by several scales, consisting of multiple items. For example, the construct of integrativeness (i.e. wanting to identify with and be part of the L2 culture) is measured by three scales: integrative orientation, attitudes toward English speaking people, and interest in foreign languages. Each of these scales is made up of multiple items that are positively or negatively keyed (e.g. 'I would like to learn many different languages', and 'I do not like learning a foreign language', respectively). Thus, integrativeness is a construct of motivation which can be divided into three smaller variables that affect it.

In the AMTB, the construct of attitudes towards the learning situation is measured by taking into account attitudes towards the language teacher as well as attitudes towards the language course. However, in Piggott's study, items regarding attitudes towards the language teacher were excluded since she believed these results could be demotivating for the teachers that take part in the study. As Gardner (2010) recommends, I used the constructs as a whole, rather than the variables for further analyses of reliability. Therefore, I will only be referring to the six major constructs from here on.

Table 2
Description of the Various Constructs of Motivation Measured by the AMTB and the Number of Positive and Negative Questions per Construct and in Total.

Construct	Positively formulated questions	Negatively formulated questions	Description
Integrativeness (N = 22)	17	5	This construct considers the desire of a language learner to identify with the cultural group of the L2 language.
Attitude Toward the Learning Situation (N = 8)	5	3	This construct considers attitudes of the language learner towards the language teachers as well as attitudes towards the language course.
Motivation (N = 30)	15	15	This construct considers motivational intensity, desire to learn English, and attitudes toward learning English.
Language Anxiety (N = 20)	10	10	This construct considers the anxiety a language learner might experience when learning or using the L2 in the classroom and outside the classroom.
Instrumentality (N = 4)	4	0	This construct considers the material, or practical, gain that might motivate a language learner to learn the L2.
Parental Encouragement (N = 8)	8	0	This construct considers the role of the language learners' parents or caregivers.
Total (N = 92)	59	33	

The questionnaire was administered during school hours and made available to the participants on computers. The participants were asked to indicate to what extent they agreed with the statements by giving a score between 1 and 7 with 1 meaning complete disagreement and 7 meaning complete agreement. Some participants did not provide a score for every item on the questionnaire. These items were indicated by 0 and were excluded from further analysis. Because the order of the statements was randomised in the online questionnaire (i.e. virtually no statements belonging to the same construct occurred next to each other), I created

an Excel file where the results are ordered by construct for each participant. This allowed for a swift calculation of the internal consistency of each construct (see section 2.3.1 below).

2.3 Reliability

2.3.1 Internal consistency of the AMTB

Given that a translated version of the AMTB was used, it is important to know whether the internal consistency of the various constructs is high, as it is in the original AMTB (Gardner, 2010). First, to be able to calculate the internal consistency of the various constructs of the AMTB, I converted the scores of the negatively formulated statements to positive scores because a high score on a negatively keyed item represents low motivation while a high score on a positively keyed item represents high motivation. As mentioned above, Gardner (2010) recommends to calculate Cronbach's Alpha (CA) (i.e. internal consistency) for each construct as a whole, as opposed to calculating CA for each smaller scale separately. This means, for example for the construct of integrativeness, that I calculated CA by taking into account all items (N=22) that belong to this construct, even though the items themselves also belong to several different scales. All scales (that belong together) are thus expected to represent the same construct.

As can be seen in table 3, the internal consistency of most constructs is high (i.e. CA > 0.8), except for the construct of instrumentality (CA=0.64).

Table 3
Cronbach's Alpha (Internal Consistency) per Construct

Construct	Cronbach's Alpha
Integrativeness	0.88
Attitudes Toward the Learning Situation	0.88
Motivation	0.91
Language Anxiety	0.94
Instrumentality	0.64
Parental Encouragement	0.83

The low value of CA for instrumentality might be due to the fact that there are only four items in the questionnaire that measure instrumentality. Nevertheless, the value of CA for instrumentality is too low and, therefore, the construct of instrumentality was excluded from this study (Bland & Altman, 1997). Given that the internal consistency of the other constructs is high, I calculated mean scores for all constructs based on the individual items belonging to each construct. This eventually resulted in one mean score per construct per participant. These mean scores were used for further analyses in this study.

Furthermore, I calculated two different overall sum scores for motivation and attitude: overall P and overall N. In both cases, the overall sum score was calculated by adding up all the mean scores for the various constructs per participant. However, both scores differ with regard to how the mean score for language anxiety was calculated since previous research has shown that language anxiety could be considered beneficial as well as detrimental to oral proficiency. For overall P, it was assumed that a high score on language anxiety is beneficial to motivation and therefore, that highly anxious learners are more motivated than less anxious learners. In this case, the mean score was computed by converting the scores of the negatively formulated statements to positive scores because a low score on a negatively formulated language anxiety item represents high anxiety and thus high motivation. Conversely, for overall N, it was assumed that a low score on language anxiety is beneficial to motivation and therefore, that less anxious language learners are more motivated than less anxious learners. In this case, the mean score was computed by converting the scores of the positively formulated statements to negative scores (e.g. a 6 becomes a 1) because a high score on a positively formulated item represents low anxiety and thus low motivation. This allowed me to investigate whether there are differences between the correlations of the two overall scores with several measures of oral language proficiency.

2.3.2 Inter-rater reliability

As mentioned above, participants were rated on oral proficiency by a teacher other than their regular teacher. In total, there were five teachers that each rated a subset of the participants. To be able to calculate inter-rater reliability, I rated all of the narratives that were also rated by the teachers. The upside of this additional rating is that this allowed me to compare the ratings of teachers to the ratings of an independent researcher (i.e. the author of this paper), thereby enabling me to investigate whether teachers assess participants differently than an independent researcher. The downside is that I could not compare the teachers' ratings to each other since every teacher rated a different subset of participants.

I calculated inter-rater reliability by making use of the Spearman-Brown formula ($(k \cdot r / 1 + (k - 1) \cdot r)$ in which k = number of raters and r = Pearson correlation). The Spearman-Brown formula takes into account the correlation between the ratings of multiple raters as well as the number of raters. First, I calculated correlations between the scores of the teachers and the scores of the independent researcher by running a Bivariate Pearson Correlation (BPC) in SPSS (Version 22.0). As can be seen in table 4, not all correlations were significant.

Table 4
Pearson Correlations between Teachers' Rating and an Independent Rater's Score per Aspect.

Teacher	Aspect			
	Vocabulary	Grammar	Fluency	Functional Adequacy
1 (N = 13)	0.95***	0.85***	0.83***	0.95***
2 (N = 28)	0.62***	0.52**	0.59**	0.73***
3 (N = 37)	0.85***	0.75**	0.72***	0.85***
4 (N = 8)	0.57	0.76*	0.54	0.63
5 (N = 34)	0.55*	0.34	0.58***	0.49**

N = number of participants graded by this particular teacher.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

For teacher 4 especially, the ratings on three different aspects did not correlate significantly with my own ratings. This might be because this teacher rated only a very small subset of participants ($N = 8$). However, since the Spearman-Brown formula does not demand the correlation between raters to be significant, I could still calculate reliability for teacher 4 and teacher 5.

After calculating correlations between ratings, I used these correlations in the Spearman-Brown formula, which then enabled me to calculate a measure of inter-rater reliability. As can be seen in table 5, the inter-rater reliability is mostly high (i.e. > 0.7). In these cases I calculated a mean score for each aspect comprised of both the scores of the teacher and my scores.

Table 5
Inter-Rater Reliability per Teacher and per Aspect.

Teacher	Aspect			
	Vocabulary	Grammar	Fluency	Functional Adequacy
1 (N = 13)	0.98	0.92	0.91	0.97
2 (N = 28)	0.77	0.68*	0.74	0.85
3 (N = 37)	0.92	0.86	0.83	0.92
4 (N = 8)	0.72	0.86	0.70	0.77
5 (N = 34)	0.71	0.51*	0.74	0.66*

* Reliability is low (i.e. < 0.7)

In cases where the inter-rater reliability is not sufficient (i.e. < 0.7), I solely used my ratings for further analyses. Because my ratings correlated strongly (i.e. > 0.7) and significantly with the ratings of the most teachers on most aspects, I assumed my ratings to be more reliable than those of the teachers in those few cases where the ratings did not correlate strongly.

Overall, these results show that teacher ratings are comparable to the ratings of an independent researcher.

A correlational analysis between the various aspects of oral proficiency was also run to investigate whether it might be beneficial to compute an overall oral proficiency score comprised of the sum of the mean scores per aspect. These results are shown in table 6.

Table 6
Pearson Correlations between the Various Aspects of Oral Proficiency.

Aspect	Aspect			
	Vocabulary	Grammar	Fluency	Functional Adequacy
Vocabulary	1*	0.83*	0.82*	0.82*
Grammar	0.83*	1*	0.81*	0.83*
Fluency	0.82*	0.81*	1*	0.80*
Functional Adequacy	0.82*	0.83*	0.80*	1*

* $p < 0.001$

Given that these results indicate that all aspects of oral proficiency correlate significantly ($p < 0.01$) with each other and therefore cannot be easily distinguished, I computed an overall oral proficiency score comprised of the mean scores per aspect. This overall score will be used for further analyses.

2.4 Analysis

Given that the internal consistency of the AMTB is high and the inter-rater reliability is high, I was able to use the data for further analyses. First, I investigated whether there are differences in performance between participants of different levels on the oral proficiency task as a whole (i.e. the sum score of all aspects), and on each aspect separately, by running a One-Way Analysis of Variance (Anova). After that, I investigated whether there are differences in motivation between participants of different educational levels by running a One-Way Anova for each construct separately and for the sum scores of motivation as a whole.

Finally, I ran a correlational analysis to investigate whether there is a relation between motivation and oral proficiency. I did this by running multiple Bivariate Pearson Correlations in SPSS to calculate four sets of correlations: 1. the correlation between the overall sum score of oral proficiency and the overall sum score of motivation 2. the correlation between the overall sum score of oral proficiency and the mean score for each construct of motivation 3. the correlation between the mean score of each aspect of oral proficiency and the overall sum score of motivation 4. the correlation between the mean score of each aspect of oral proficiency and the mean score for each construct of motivation. This allowed me to investigate whether there is a relation between motivation and oral proficiency in general, whether some subtypes of motivation are more strongly related to oral proficiency than others, and whether some aspects of oral proficiency are more strongly related to motivation than others.

3. Results

3.1 Group differences

The participants' mean scores on the oral proficiency task are summarised in table 7 by aspect and educational level.

Table 7
Mean Scores (Standard Deviations) of Oral Proficiency by Educational Level and Aspect of Proficiency.

	VMBO- TL (N = 8)	HAVO (N = 10)	VMBO- TL & HAVO (N = 34)	VWO (N = 40)	HAVO & VWO (N = 28)
Aspect	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Vocabulary	2.1 (0.4)	2.1 (1.1)	2.1 (0.8)	2.6 (0.9)	2.4 (0.9)
Grammar	2.1 (0.6)	2.3 (1.0)	2.4 (0.9)	2.5 (0.9)	2.4 (0.8)
Fluency	2.1 (0.6)	2.8 (1.0)	2.5 (0.7)	2.7 (0.9)	2.7 (0.6)
Functional Adequacy	2.2 (0.7)	2.7 (1.3)	3.0 (0.9)	3.1 (0.9)	2.7 (0.8)
Overall	8.5 (1.9)	10.0 (4.3)	9.9 (2.9)	10.8 (3.4)	10.2 (3.1)

A one-way ANOVA was conducted to investigate whether there is an effect of educational level on the performance of the participants on an oral story-telling task. The results showed there were no differences in performance between participants of different educational levels on vocabulary ($F(4, 115) = 1.92$; $p = 0.112$), grammar ($F(4, 115) = 0.45$; $p = 0.772$), fluency ($F(4, 115) = 1.38$; $p = 0.246$), functional adequacy ($F(4, 115) = 2.14$; $p = 0.080$), and on the overall sum score of proficiency ($F(4, 115) = 1.10$; $p = 0.359$). These results suggest that all participants performed equally well on the oral proficiency task, regardless of their educational level.

Table 8
Mean Motivation Scores (Standard Deviations) by Educational Level and Construct.

Construct	VMBO- TL (N = 8)	HAVO (N = 10)	VMBO- TL & HAVO (N = 34)	VWO (N = 40)	HAVO & VWO (N = 28)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Integrativeness	3.9 (0.5)	3.9 (0.8)	4.0 (1.0)	4.2 (0.6)	4.2 (0.7)
Attitude Toward the Learning Situation	3.4 (0.7)	4.0 (1.1)	4.1 (1.1)	3.2 (1.1)	3.6 (1.0)
Motivation	4.1 (0.6)	4.4 (0.8)	4.2 (0.8)	4.0 (0.8)	4.0 (0.7)
Language Anxiety	3.2 (0.8)	2.7 (1.4)	2.8 (0.8)	3.1 (1.3)	3.1 (1.0)
Parental encouragement	4.3 (0.5)	4.2 (0.9)	4.1 (0.9)	4.0 (1.1)	4.2 (1.1)
Overall P	18.9 (1.7)	19.3 (2.6)	19.0 (2.9)	18.4 (3.1)	19.1 (2.4)
Overall N	19.8 (2.0)	20.4 (3.0)	20.3 (3.4)	19.2 (3.0)	20.0 (2.7)

Overall P = Sum score of the means of all constructs in which a high score on language anxiety is considered to indicate high motivation.

Overall N = Sum score of the means of all constructs in which a high score on language anxiety is considered to indicate low motivation.

Table 8 presents the mean AMTB scores by construct and educational level. A one-way ANOVA was conducted to investigate whether there is an effect of educational level on the scores of the participants on the AMTB questionnaire. The results showed there was no difference in scores on the AMTB between participants of different educational levels regarding integrativeness ($F(4, 115) = 0.78$; $p = 0.540$), motivation ($F(4, 115) = 0.72$; $p = 0.578$), language anxiety ($F(4, 115) = 0.63$; $p = 0.643$), parental encouragement ($F(4, 114) = 0.38$; $p = 0.823$), overall P ($F(4, 115) = 0.37$; $p = 0.827$), and overall N ($F(4, 115) = 0.71$; $p = 0.585$).

However, regarding the construct of attitude toward the language situation, a significant difference was found between groups ($F(4, 115) = 3.45$; $p = 0.011$). Post hoc

comparisons using the Bonferroni test indicated that the mean score of the VMBO-TL & HAVO group ($M = 4.1$; $SD = 1.1$) was significantly higher ($p = 0.007$) than the mean score of the VWO group ($M = 3.2$; $SD = 1.1$) on the construct attitude toward the language situation. These results indicate that the VMBO-TL & HAVO group have a more positive attitude toward the language learning situation than the VWO group as measured by the AMTB. Taken together, these results indicate that there are no differences in performance between groups on the oral proficiency task and that there are no differences in motivation between groups as measured by the AMTB, except for the construct of attitude toward the language situation.

3.2 Correlational analyses

The correlations between several constructs of motivation and several aspects of oral proficiency are shown in table 9. As can be seen in table 9 there are significant correlations between integrativeness and vocabulary, fluency, functional adequacy, and overall proficiency. Furthermore, there are significant negative correlations between language anxiety and vocabulary, grammar, fluency, functional adequacy, and overall proficiency. Lastly, a significant negative correlation between parental encouragement and functional adequacy was found. These results show that participants with a high integrative motivation performed better on the oral proficiency task than participants with a lower integrative motivation. Furthermore, participants with a higher level of language anxiety performed worse on the oral proficiency task than participants with a lower level of language anxiety.

Table 9
Pearson Correlations between Various Constructs of Motivation and Various Aspects of Oral Proficiency.

Construct	Aspect				
	Vocabulary	Grammar	Fluency	Functional Adequacy	Overall Proficiency
Integrativeness	0.26**	0.18	0.20*	0.22*	0.23*
Attitude Toward the Learning Situation	0.15	0.14	0.16	0.18	0.17
Motivation	0.11	0.06	0.11	0.11;	0.10
Language Anxiety	-0.35***	-0.34***	-0.31***	-0.32***	-0.35***
Parental Encouragement	-0.15	-0.16	-0.13	-0.18*	-0.17
Overall P	-0.04	-0.76	-0.03	-0.04	-0.05
Overall N	0.10	0.06	0.10	0.11	0.10

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

As mentioned previously, all aspects of oral proficiency correlate with each other significantly and it could therefore prove difficult to distinguish between the various aspects (see table 6). The results shown in table 9 seem to confirm that the various aspects of oral proficiency are related because every construct of motivation that correlates significantly with individual aspects of proficiency (i.e. integrativeness and language anxiety) also correlates significantly with the overall score of proficiency, except for parental encouragement.

4. Discussion

4.1 Motivation and oral proficiency

As mentioned above, a significant positive correlation was found between integrativeness and oral proficiency. This suggests that there is indeed a relation between motivation and second language oral proficiency. This finding is in line with previous research on the relation

between motivation and oral proficiency. For example, Hernández (2010) also found that learners with a high integrative motivation performed better on an oral proficiency task than learners with a lower integrative motivation. Therefore, the results seem to confirm that there is a relation between motivation and level of oral proficiency in Dutch learners of English as a foreign language.

This finding also suggests that participants who identify more with the L2 culture perform better on the oral proficiency task than participants who identify less with the L2 culture. This might mean that identifying with the L2 culture to some extent could be more beneficial to performance on an oral proficiency task than having no, or little affiliation with the L2 culture. It might also be the case that it works the other way around, namely that participants who are able to express themselves more efficiently orally, also identify more with the L2 cultural group because they are better able to immerse themselves in, and take part in the activities coming forth from the L2 culture (e.g. television shows, books, etc.) than participants who are less fluent orally. Since correlations do not provide any evidence for a causal relationship it is difficult to indicate which explanation is more plausible.

With regard to integrativeness, Dörnyei (2009) argues that English in its current state is very much a global language and therefore, the concept of integrativeness might not be relevant anymore. It might be unclear to English as a foreign language learners what the L2 community or culture is that they should identify with since the English community might be a global one without a specific culture attached to it. This could mean that the participants, also because of their relatively young age, are confused as to what the L2 cultural group is that they should identify with. Nevertheless, the results from this study suggest that the participants in this study have a clear idea which L2 community belongs to the English language. It seems unlikely that the degree of integrative motivation would have correlated significantly with oral proficiency if it was not clear to the participants what the L2

community is. Whether this is a global sense of community or a local one remains unclear. It would be interesting for future research to investigate which culture or community high school students in non-English speaking countries associate with the English language. Nevertheless, the results indicate that it might be beneficial for the language learners if the language teachers were to incorporate elements of the L2 culture in the curriculum of the language course.

A significant negative correlation was found between language anxiety and oral proficiency. Again, this finding suggests that a relation between motivation and second language proficiency exists. Furthermore, this result indicates that highly anxious language learners perform worse on an oral proficiency task than less anxious language learners. This is in line with previous research, which showed that anxiety has a negative influence on second language proficiency (e.g. MacIntyre & Gardner, 1991), and with findings by Papi and Teimouri (2014) who found that for two language learners groups who both displayed high overall motivation, the group with the lowest language anxiety showed higher L2 proficiency scores. Therefore, the results indicate that language anxiety has an important role in second language acquisition. Thus, it might be beneficial for language teachers to take into account these findings regarding language anxiety and strive to create a safe, anxiety-reducing environment in the language classroom.

Taken together, the findings regarding integrativeness and language anxiety, and the fact that no other correlations were found between the other constructs of motivation and oral proficiency, seem to confirm the hypothesis that there is a relation between motivation and second language oral proficiency, as well as the hypothesis that some constructs of motivation are more strongly related to oral proficiency than others. Also, the results seem to contradict statements that the socio-educational approach might not be relevant anymore (Dörnyei, 2009). Furthermore, the findings from this study seem to suggest that the AMTB is also

appropriate for use in non-bilingual contexts, and follows a long line of studies that suggest that motivation plays an important role in L2 proficiency (e.g. Gardner, 2010). Moreover, this study is one of only several studies that found a relation between motivation and a specific type of proficiency (i.e. oral proficiency), as well as the first study to find a relation between motivation and oral proficiency in Dutch second language learners of English.

4.2 Group differences regarding oral proficiency

Another interesting finding from this study is that the results indicate that there are no differences in performance on the oral proficiency task between participants of different educational levels. This means, for example, that participants enrolled in the highest educational level (i.e. VWO) did not perform better than participants enrolled in the lowest educational level (i.e. VMBO-TL). This is unexpected since part of the score on the CITO test that children have to complete in primary school to receive a recommendation for a secondary school educational level is determined by their English skills. Therefore, it would be reasonable to expect that children enrolled in a higher educational level would outperform children enrolled in a lower educational level on an English proficiency assessment.

However, oral proficiency is not part of the CITO test and therefore it might be that the score for English on the CITO test (which mainly consists of questions measuring reading proficiency) does not predict performance on an oral proficiency task. This would be in line with previous research by Geva (2006) who notes in her literature review that even though second language oral proficiency is significantly correlated with word reading skills, individual differences in other abilities, such as phonological processing and working memory, are stronger predictors of word reading skills. Thus, the assessment process for secondary school as it is now (via the CITO test) seems unsatisfactory with regard to English

oral proficiency skills; students enrolled in a higher educational level do not perform better than students enrolled in a lower educational level on an English oral proficiency task.

Given that the participants only just started secondary school at the moment of testing, it might also be that they had not yet received enough exposure to English in the language classroom for a difference to manifest itself. In primary school, most Dutch students receive roughly the same amount of English education. This means that, if any group differences would have been found, this would have probably been due to differences in general intelligence or learning style, among other things. However, it could have also been due to individual differences in amount of exposure to English outside of the classroom. Presumably, participants enrolled in the lower educational levels could have had more exposure to English outside of the language classroom. This could also explain why no differences were found in oral proficiency between participants enrolled in different educational levels. Unfortunately, no data on exposure to English outside of the classroom, and no data on the language background of the participants was gathered.

Taken together, the results indicate that a participant's educational level does not necessarily predict performance on an English oral proficiency task since no difference in performance was found between participants of different educational levels.

4.3 Group differences regarding attitude and motivation

The results showed there were no differences in attitude and motivation between participants of different educational levels except for the construct attitude toward the learning situation. Regarding this construct, the results showed that the VMBO-TL & HAVO group have a more positive attitude toward the learning situation than the VWO group. It is difficult to explain this finding since there is no literature that addresses motivational differences between different educational levels in the Netherlands. However, given that the construct attitude

toward the learning situation, in this study, consists of questions assessing the degree of positive attitude toward the English course, the results suggest that the VMBO-TL & HAVO group have had a more positive experience learning the L2 in the foreign language classroom.

Unfortunately, attitudes toward the foreign language teacher were excluded from this study since Piggott (in preparation) believed these results might be demotivating for the language teachers involved in the study. Therefore, it might be that the group difference found here is due to the groups having different foreign language teachers since teachers inevitably have an influence on a learner's experience in the foreign language classroom. This is also mentioned by Dörnyei (1994) who emphasises the role the language teacher has in motivating language learners. He argues that skilled language teachers often have a positive effect on the motivation and performance of language learners in language classrooms. Therefore, I believe it would be beneficial for future research to address attitudes toward the foreign language teacher in order to find out to what extent a particular teacher influences a learner's experience of the language classroom.

No further differences in motivation were found between groups for the other constructs. This seems to suggest that motivation, except attitude toward the language situation, is not related to educational level. Presumably, the participants had not been exposed to English for a sufficient amount of time for differences in motivation to become apparent. Therefore, differences in motivation may still emerge and it would be relevant for future research to carry out a longitudinal study, which investigates motivation of the same group of participants at different times.

4.4 The reliability of teacher ratings

As discussed in section 2.3.2 above, the results indicate that teacher ratings are comparable to the ratings of an independent researcher. However, in the present study the teachers rated

participants they did not teach themselves; this might have made their ratings more objective. Therefore, it would be interesting for future research into the reliability of teacher ratings to investigate whether teacher ratings of students, which are taught by the teacher, are also comparable with those of an independent researcher because teacher ratings are often used in L2 research.

4.5 Limitations and future research

Several limitations of this paper have to be taken into consideration. Firstly, the participants in this study all came from the same geographical area in the Netherlands and went to the same high school. Future researchers should try to find a more diverse population spread throughout the Netherlands. This would make the results more generalisable and reliable because, as previous research indicates, results regarding the relation between motivation and L2 proficiency may differ between language contexts (Gardner, 2010). Furthermore, there was a discrepancy between the time the AMTB was administered and the time the oral proficiency was assessed. The participants' oral proficiency was assessed at the end of their first year of secondary school while the AMTB was administered before Christmas break in their second year. This means that the participants' responses on the AMTB might have been influenced by their first year English grades. Also, for some groups (e.g. the VMBO-TL group), data of only a small number of participants was available for analysis due to technical issues. Future research would benefit from an even spread of participants enrolled in the various educational levels.

Second, the participants' oral skills were assessed in pairs. This might have influenced the results since the second participant of each pair was able to pick up words and phrases used by the first participant. Because the two picture stories had some overlapping story elements, it might have been the case that the second participant benefitted from the fact that

another participant told the story first. Also, a story telling task is one possible measure of oral proficiency. Future research should consist of multiple tasks that measure oral proficiency to be able to compare the results and provide a more comprehensive measure of oral proficiency.

Third, the AMTB is a long questionnaire consisting of over 100 statements. The teachers noted that some participants found the questionnaire to be too long and therefore started rushing through the questionnaire. It would perhaps be beneficial for future research to make use of a questionnaire consisting of less items, such as the mini-AMTB (Tennant & Gardner, 2004). Furthermore, the attitude toward the foreign language teacher was not taken into consideration in this study (see 4.3 above).

Fourth, the AMTB and the socio-educational model represent only one possible approach to attitude and motivation. There are other approaches that have also been empirically studied, such as the L2 motivational self system (Dörnyei, 2005; Dörnyei, 2009), which could also provide interesting new insights into the relation between motivation and proficiency. Therefore, it could prove useful for future research to investigate the comparability of the socio-educational approach and the L2 motivational self system, and perhaps make use of the questionnaires developed from both approaches.

Finally, a general point of criticism towards the socio-educational model was voiced recently by Ellis (2015). He argues that the socio-educational model takes no account of the fact that motivation is inherently not static and that motivation can also be drawn from achieving success and making progress in the L2 acquisition process. Therefore, it might be that the AMTB, which is based on the socio-educational model, is an outdated questionnaire which needs revising and perhaps more recent approaches are more relevant (e.g. the L2 motivational self system). Thus, it is important for future research to investigate whether the socio-educational approach is still relevant and whether other approaches to constructing motivation might be better for capturing the dynamic concept of motivation.

5. Conclusion

Overall, the finding that there is a relation between motivation and L2 proficiency is in line with findings of previous research on the relation between motivation and L2 proficiency.

This study adds to the existing body of research by showing that a relation also exists between motivation and L2 proficiency for Dutch high school students in the Netherlands, and that there is relation between motivation and a specific type of L2 proficiency, namely oral proficiency. More specifically, the results showed a negative correlation between language anxiety and oral proficiency, suggesting that highly anxious language learners performed worse on the oral proficiency task than less anxious learners. Furthermore, a positive correlation was found between integrativeness and oral proficiency, suggesting that language learners with a more positive attitude toward the L2 community outperformed the language learners with less positive attitude toward the L2 community on the oral proficiency task. These findings could prove useful for language teachers who might take these results into consideration and use them to improve the environment of the L2 classroom.

References

- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal*, *314*, 572.
- Crookes, G., & Schmidt, R. W. (1991). Motivation: reopening the research agenda. *Language Learning*, *41*(4), 469-512.
- Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, *78*(3), 273-284.
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*(3), 117-135.
- Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 9-42). Bristol, UK: Multilingual Matters.
- Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language and Communication Disorders*, *47*(5), 542-555.
- Ellis, R. (2015). *Understanding Second Language Acquisition*. Oxford, UK: Oxford University Press.
- Elzenga, A., & de Graaff, R. (2015). Motivatie voor Engels en Frans in het tweetalig onderwijs. *Levende Talen Tijdschrift*, *16*(4), 16-26.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London, UK: Edward Arnold Publishers.
- Gardner, R. C. (2005). *Integrative motivation and second language acquisition*. Canadian Association of Applied Linguistics/Canadian Linguistics Association Joint Plenary

- Talk. London and Ontario. Retrieved February 27, 2016 from <http://publish.uwo.ca/~gardner/>
- Gardner, R. C. (2007). Motivation and second language acquisition. *Porta Linguarum*, 8, 9-20.
- Gardner, R. C. (2010). *Motivation and Second Language Acquisition: the Socio-Educational Model*. New York, NY: Peter Lang Publishing.
- Gardner, R. C., Day, J. B., & MacIntyre, P. D. (1992). Integrative motivation, induced anxiety, and language learning in a controlled environment. *Studies in Second Language Acquisition*, 14(2), 197-214.
- Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second-language acquisition. *Canadian Journal of Psychology*, 13(4), 266-272.
- Gardner, R. C., Tremblay, P. F., & Masgoret, A. (1997). Towards a full model of second language learning: an empirical investigation. *The Modern Language Journal*, 81(3), 344-362.
- Geva, E. (2006). Second-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 1-16). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hernández, T. A. (2006). Integrative motivation as a predictor of success in the intermediate foreign language classroom. *Foreign Language Annals*, 39(4), 605-617.
- Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94(4), 600-617.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132.

- Iwashita, N. (2010). Features of oral proficiency in task performance by EFL and JFL Learners. In M. T. Prior et al. (Eds.), *Selected Proceedings of the 2008 Second Language Research Forum*, 32-47. Somerville, MA: Cascadilla Proceedings Project.
- Kleinginna Jr, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and emotion*, 5(3), 263-291.
- Lukmani, Y. M. (1972). Motivation to learn and language proficiency. *Language Learning*, 22(2), 261-273.
- MacIntyre, P. D., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: a review of the literature. *Language Learning*, 41(1), 85-117.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283-305.
- Mayer, M. (1969). *Frog, where are you?.* New York, NY: Dial Press.
- Mayer, M., & Mayer, M. (1971) *A boy, a dog, a frog, and a friend.* New York, NY: Dial Press.
- Papi, M., & Teimouri, Y. (2014). Language learner motivational types: a cluster analysis study. *Language Learning*, 64(3), 493-525.
- 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.
- Spolsky, B. (1969). Attitudinal aspects of second language learning. *Language Learning*, 19(3), 271-275.
- Tennant, J., & Gardner, R. C. (2004). The computerized mini-AMTB. *CALICO Journal*, 21(2), 245-263.
- Upshur, J. A., & Turner, C. E. (1999). Systematic effects in the rating of second-language speaking ability: test method and learner discourse. *Language Testing*, 16(1), 82-111.

Williams, M. (1994). Motivation in foreign and second language learning: An interactive perspective. *Educational and Child Psychology*, 11(2), 77-84.

Appendix A



- 4646 40th Street, NW • Washington, DC 20016-1859 • <http://www.cal.org>

CAL ORAL PROFICIENCY EXAM AND STUDENT ORAL PROFICIENCY ASSESSMENT RATING SCALE (COPE/SOPARS)

English Version © 2009 CAL

JR. NOVICE-LOW	JR. NOVICE-MID	JR. NOVICE-HIGH	JR. INTERMEDIATE-LOW	JR. INTERMEDIATE-MID	JR. INTERMEDIATE-HIGH	Jr. Advanced-Low	JR. ADVANCED-MID	JR. ADVANCED-HIGH
Oral Fluency								
-Produces only isolated words (i.e., single-word responses) and/or greetings and polite expressions such as <i>good morning</i> and <i>thank you</i> .	-In addition to isolated words, uses phrases of two or more words, and/or memorized phrases or sentences (e.g., <i>My name is...</i> , <i>I don't know</i>) in predictable topic areas. -May attempt to create sentences, but is not successful. -Long pauses are common.	-Uses memorized expressions with reasonable ease. -Shows emerging signs of creating with the language to communicate ideas. -Creates some sentences successfully, but cannot sustain sentence-level speech.	-Goes beyond memorized expressions to maintain simple conversations at the sentence level by creating with the language, although in a restrictive and reactive manner. -Handles a limited number of everyday social and academic interactions.	-Maintains simple sentence-level conversations. May initiate talk spontaneously without relying on questions or prompts. -Gives simple descriptions successfully. -May attempt longer, more complex sentences. Few, if any, connectors are used.	-Initiates and sustains conversations by using language creatively. -Shows emerging evidence of paragraph-like speech with some connected sentences (e.g., <i>then, so, that, etc.</i>) in descriptions and simple narratives, but has no actual paragraphs with a main idea, organization, and connection.	-Reports facts easily. Can discuss topics of personal interest and some academic topics at the paragraph level to satisfy school and everyday requirements. -Narrates and describes at the paragraph level also, although haltingly at times. -False starts are common.	-Handles with ease and confidence concrete topics of personal and general interest and a number of academic topics. -Narrates and describes smoothly in paragraphs having a main idea, organization, and a variety of sentence connectors (e.g., <i>first, next, finally; then, when, that, although, but, therefore, so, etc.</i>).	-Handles most social and academic requirements confidently, but may hesitate when responding to complex, formal tasks (Superior level). -Organizes and extends discourse (multiple paragraphs) in an emerging ability to hypothesize on abstract topics (<i>if-then</i>) and support opinions.
Grammar (Speaking)								
-May use greetings and polite expressions accurately. -Lacks an awareness of grammar and syntax.	-Memorized expressions with verbs and other short phrases may be accurate, but inaccuracies are common. -Does not successfully create at the sentence level with conjugated verbs.	-Creates some sentences with conjugated verbs, but in other attempts to create sentences, verbs may be lacking or are not conjugated. -Other grammatical inaccuracies are present.	-Uses a variety of common verbs in present tense (conjugations may be inaccurate) in sentences. -Other verb tenses/forms may appear in memorized language. -The listener may be confused by this speech due to the many grammatical inaccuracies.	-Uses an increasing number and variety of verbs. -Verbs are mostly in present tense although awareness of other verb tenses (future/past) and forms may be evident. -Many grammatical inaccuracies may be present.	-Uses a large variety of verbs well in present tense. Uses many verbs in the past tenses but lacks control of past. May use future and other verb forms. -Grammatical inaccuracies may still be present. Awareness of inaccuracies may be evident.	-Uses present, past, and future tenses. -May effectively self-correct when aware of grammatical inaccuracies. -Structures of native language may be evident (e.g., literal translation).	-Has good control of present, past, and future tenses. -Some inaccuracies may remain, but speech is readily understood by native speakers of the language. *In some cases, may use non-standard varieties of grammar.	-Uses all verb tenses accurately and sometimes uses complex grammatical structures, (e.g., <i>if...occurred, then...might also happen</i>). -Some patterns of error may persist, but they do not interfere with communication.

Vocabulary (Speaking)								
-Uses single words in very specific topic areas in predictable contexts. -May use greetings and polite expressions.	- Uses single words, short phrases, greetings, polite expressions, and other memorized expressions on a limited number of topics. -Frequent searches for words are common. May use native language or gestures when attempting to create with language.	-Uses vocabulary centering on basic objects, places, and common kinship terms, adequate for minimally elaborating utterances in predictable topic areas. -Use of native language and gestures is common to expand topics.	-Has basic vocabulary for making statements and asking questions to satisfy basic social and academic needs, but not for explaining or elaborating on them. -Use of some native language is common when vocabulary is lacking.	-Has basic vocabulary, permitting discussions of a personal nature and limited academic topics. Serious gaps exist for discussing topics of general interest. -If precise word is lacking, may use circumlocution ineffectively. May resort to native language.	-Has a broad enough vocabulary for discussing simple social and academic topics in generalities, but lacks detail. -Sometimes achieves successful circumlocution when precise word is lacking. May use native language occasionally.	-Vocabulary is primarily generic but is adequate for discussing concrete or factual topics of a personal nature, topics of general interest, and academic subjects. -May use circumlocution successfully when specific terms are lacking.	-Has adequate vocabulary for including detail when talking about concrete or factual topics of a personal nature, topics of general interest, and academic subjects. -Uses circumlocution effectively. Rarely uses native language.	-Uses precise vocabulary for discussing a wide variety of topics related to everyday social and academic situations. -Lack of vocabulary rarely interrupts the flow of speech.
Listening Comprehension								
-Recognizes single, isolated words, greetings and polite expressions.	-Understands predictable questions, statements, and commands in familiar topic areas (with strong contextual support), though at slower than normal rate of speech and/or with repetitions.	-Understands simple questions, statements, and commands in familiar topic areas, and some new sentences with strong contextual support. -May require repetition, slower speech, or rephrasing.	-Understands familiar and new sentence-level questions and commands in a limited number of content areas with strong contextual support for unfamiliar topics. -Follows conversation at a fairly normal rate.	-Understands sentence-level speech in new contexts at a normal rate of speech although slow-downs may be necessary for unfamiliar topics. -Carries out commands without prompting.	-Understands longer stretches of connected speech on a number of topics at a normal rate of speech. -Seldom has problems comprehending everyday topics. (Can request clarification verbally.)	-Understands main ideas and many details in connected speech on some academic topics and on topics of personal interest.	-Understands main ideas and most details in connected speech on a variety of topics, but may be unable to follow complicated speech. -May have difficulty with highly idiomatic speech.	-Understands complex academic discourse and highly idiomatic speech in conversation. -Confusion may occur due to socio-cultural nuances or unfamiliar topics.

* This feature may not appear, but if present in student speech, is acceptable at the Jr. Advanced-Mid level of proficiency.

The COPE/SOPA Rating Scale is based on the ACTFL Proficiency Guidelines, American Council on the Teaching of Foreign Languages (1986, 1999).

Contact Lynn Thompson lhompson@cal.org for more information.

Appendix B

Below are two transcripts of parts of the audio files of two different participants. Only the most essential information relevant for the rating is included here.

P = participant

T = teacher/rater

Numbers between brackets are pauses in seconds.

Example 1

P: Uh, I see a little boy

(1 sec)

P: He is uh

(2)

P: Fish.. uh

T: Yeah, yeah fishing

P: Fishing

(2)

P: I see a dog

P: He

(4)

P: Duwt

(2)

P: Hij, he

(3)

P: The boy ff

(7)

T: Falls or push

P: The dog push the boy int the water

(8)

P: The boy is swimming naar de wal

(7)

P: The dog have auw

T: Yeah, by who?

(2)

P: The

(2)

P: Animal bit he

(12)

P: The animal

(2)

P: duwt the dog in the water

As can be seen from this transcript, the participant is not very fluent. She takes long pauses between utterances, has trouble finishing an utterance in one go, and hesitations occur frequently. Therefore, this participant received a 1 for fluency. Regarding vocabulary, the participants uses many words from the L1. Also, she produces mainly isolated words or very limited phrases. Thus, this participant received a 1 for vocabulary. As for grammar, the participant uses only short simple sentences and a lot of the structures are not target like structures. Therefore, this participant received a 1 for grammar. Finally, because this participant used many Dutch words and non-target like grammatical structures, it was very

difficult to make sense of the story. Therefore, this participant also received a 1 for functional adequacy.

Example 2

P: Well, uh, I see his bedroom

P: And I see a little boy with a dog and a frog in a jar and it's night

(2)

P: And here I see the boy sleeping with his dog and the frog is trying to get out of the jar

T: Yeah

(2)

P: Uhm, here is the boy searching for the frog and

(2)

P: Here is the boy outside looking for the frog with his dog

(5)

P: The boy was looking here in the hole but he's not there and the dog is uh

(4)

P: Jumping to a uh

(3)

P: bijenkorf ik weet niet meer

T: Do you know a bij? (to other participant)

T: They call that a beehive

P: There is a mole I think there in the hole.

In comparison with the transcript from example 1 this participant is much more fluent. She takes less pauses and shorter pauses between utterances and she mainly finishes utterances in

one go. However, her speech rate is not very high. Therefore, this participant received a 3 for fluency. With regard to vocabulary, it is clear that this participant relies less on L1 words than the participant discussed above. However, the phrases and chunks she uses are not very diverse, and she still relied on Dutch to some extent. Thus, this participant received a 2 for vocabulary. As for grammar, this participant is able to create correct grammatical sentences. However, these structures are still predominantly short, simple, and not very diverse. Therefore, this participant received a 3 for grammar. With regard to functional adequacy, the participant is clearly able to tell the story. But, because she still relied on Dutch to some extent, this means that the general message might not always be comprehensible to an English native speaker. Thus, this participant received a 3 for functional adequacy.