MASTER'S THESIS

Norwegian sixth and seventh graders' metalinguistic awareness and its interrelatedness with their teacher-reported English competence: a pilot study

Marianne Bakken

Master's in Education with English Didactics
Department of Language, Literature, Mathematics and Interpreting
Supervisor: Christine Erna Elisabeth Möller-Omrani
15.05.2020

Summary

The study of metalinguistic awareness (MLA) has become increasingly popular in the last decades. Several studies have examined the relationship between MLA and second language acquisition (SLA) and have found positive correlations between the two. This study aims to investigate one 6th grade and one 7th grade's level of MLA and the interrelatedness with their teacher-reported English competence. The age-group was chosen to be able to compare the results with Tellier (2013), who found a significant difference in means between the two grades. The study developed an MLA-test which draws on the already developed MLA-test of Tellier (2013) and EVLANG (1997-2001). The data collection mainly consists of the test-results. Additionally, the study collected teacher reports on the participants' English competence. The test results were analyzed by a quantitative scoring-system as well as by a qualitative discussion to yield additional insight in the results. The findings show that there was a clear difference in means between grades 6 and 7. The study did not find any correlations between the participants' state of MLA and teacher-reported English competence.

Sammendrag

Forskning på metalinguistisk bevissthet (MLA) har økt i popularitet de siste tiårene. Flere studier har undersøkt forholdet mellom MLA og andrespråklig tilegnelse og har funnet positive korrelasjoner mellom de to. Denne studien har undersøkt MLA hos elever på et 6.trinn og et 7. trinn og videre vurdert sammenhengen mellom MLA og deltakernes engelskkompetanse. Engelskkompetansen til deltakerne er oppgitt av deres engelsklærere. Aldersgruppen ble valgt for å kunne sammenligne resultatene med Tellier (2013) som fant en betydelig forskjell i gjennomsnittsresultatet mellom de to trinnene. Studien utviklet en MLA-test som bygger på den allerede utviklede MLA-testen av Tellier (2013) og EVLANG (1997-2001). Deltakernes svar fra testen er studiens datainnsamling. I tillegg samlet studien lærerrapporter om deltakernes engelskkompetanse. Testresultatene ble analysert med et kvantitativt skåringssystem. Videre ble det gjennomført en kvalitativ diskusjon for å få en grundigere innsikt i resultatene. Funnene viser at det var en klar forskjell i gjennomsnittsresultat mellom 6. og 7. trinn. Studien fant ingen korrelasjoner mellom deltakernes MLA og deres lærerrapporterte engelskkompetanse.

Acknowledgments

I want to express my deepest gratitude to my supervisor, Christine Erna Elisabeth Möller-Omrani. Thank you for your guidance and support throughout this project. I appreciate your counseling and encouragement, as well as your careful readings through my many e-mails and numerous drafts.

I want to thank the school for allowing me to conduct the study there. And a much obliged thank you to the participants and their teachers who made this project possible.

I would also like to thank my family. Ottar Sverre, Rita-Karin and Ingvild, your support and encouraging words, as well as your readings through my drafts have been invaluable.

Lastly, thank you to my friends who have shown interest in this project. You have played a big role in making this process enjoyable.

Table of contents

List of figures

List of tables

1.0 Introduction

In the last decades, increasing emphasis has been put on metalinguistic awareness (MLA) for educational purposes. Following Tunmer, Pratt & Herriman (1984), the definition the present study draws on is the individual's ability to distance themselves from the language itself and to evaluate, analyze, and manipulate the linguistic features (p. 289). *Metalinguistic awareness* emerged from the more familiar term *language awareness,* and there has been conducted several studies investigating the possible positive relationship between MLA and language competence- and learning and reading comprehension, as well as the potential correlation with L2 (second language) acquisition (Alderson, 1997 and Elder, 1997 both in [1]Simard & Gutìerrez, 2018, pp. 209-210). Some studies have also compared the state of MLA in bilinguals and monolinguals, to be able to detect a potential advantage in bilinguals (Bialystok, 1999; 2001; 2004, Tellier 2013). The main aim of this study is to investigate the state of metalinguistic awareness in grade 6 and grade 7 in a primary school in Bergen. Additionally, the study's objective is to research whether there is a clear difference in means between grades 6 and 7. These two grades has been chosen to compare this study's results with Tellier (2013), who found a significant difference in means between grade 6 and 7 (ages 11 and 12). Lastly, a comparison between the participants' metalinguistic awareness and their teacher-reported English competence is included to be able to state whether there is a correlation between the two. Consequently, the study's research questions are:

1. What is the state of metalinguistic awareness in grade 6 in a primary school in Bergen?
2. What is the state of metalinguistic awareness in grade 7 in a primary school in Bergen?
3. Is there a correlation between the participants' MLA and their teacher-reported English competence?
4. Is there a clear difference in MLA scores between grade 6 and 7?

The participant's English competence will be assessed by their respective teachers. The participants' MLA will be investigated with the use of an MLA-test that has been developed for this purpose. The MLA test is a continuation of the already developed MLA-test by Tellier (2013) as well as two supplemental tasks taken from EVLANG (1997-2001). The thesis is divided into five chapters. Chapter 1 is the introduction of the thesis. Chapter 2 consists of the

---

[1] Due to current limitations in access to the original works, the author has included sources referenced in other research and review articles, even though she is aware that second-hand referencing should be avoided.

study's theoretical framework and will present a conceptual understanding of the term *metalinguistic awareness* before relevant research are elicited. Chapter 3 provides an explication of the methods and materials used in the study. The design of the MLA test, the task items, and a review of the scoring-system will be described in the respective chapter. Furthermore, chapter 4 consists of the results in grades 6 and 7, as well as a comparative qualitative discussion between the two grades. The teacher assessments of the participants' English competence will also be elicited in this chapter. Lastly, chapter 5 provides a general discussion of the findings as well as propose suggestions for further research.

2.0 Theoretical background

2.1 An historical overview of metalinguistic awareness

As mentioned in section 1.0, *metalinguistic awareness* emerged from the more familiar term *language awareness.* Language awareness can be described as an individual's consciousness and sensitivity towards language teaching, which was first introduced in the works of Wilhelm von Humboldt (1767-1835, cited in Jessner, 2008, p. 387). According to Jessner (2008), language awareness and metalinguistic awareness are used interchangeably. However, language awareness is more extensive in its application (p. 258). Lev Vygotsky addressed the relationship between cognitive development and metalinguistic abilities already in 1934, although the term *metalinguistic* was not utilized in his works (Sartori, Giolfo & Cassuto, 2016, p. 35-36). *Metalinguistic* refers to cognitive or linguistic activities regarding language (Gombert, 1993 cited in Simard & Gutìrrez, 2018, p. 205), and is often related to knowledge, ability, and awareness. Metalinguistic awareness, metalinguistic knowledge, and metalinguistic abilities are three entities that have gained more attention concerning second language acquisition (SLA). The term *metalinguistic awareness* has been redefined by researchers continuously. The concept will be explained based on theoretical perspectives before the study's interpretation of the *metalinguistic awareness* is presented in section 2.2.

2.2 Defining metalinguistic awareness

Metalinguistic awareness as a unique construct was first referenced by Cazden (1974 cited in Bialystok, 2004, pp. 121-122), who defined metalinguistic awareness as an "ability to make language forms opaque and attend to them in and for themselves". Stating that it is a "special kind of language performance, one which makes special cognitive demands, and seems to be

less easily and less universally acquired than the language performances of speaking and listening." In simple terms, metalinguistic awareness is an understanding that language is more than a way to communicate. One can distance oneself from the language and evaluate, analyze, and manipulate it. Tunmer, Pratt and Herriman (1984) discussed the relationship between metalinguistic awareness and literacy and defined metalinguistic awareness as the "ability to reflect upon language and manipulate its structural features, treating language itself as an object, as opposed to using it to understand and produce sentences" (p. 289). Metalinguistic awareness points to a combination of attention and explicit mental representations (Bialystok, 2004, pp. 126-127). In other words, metalinguistic awareness requires selective attention on different aspects of information in a language. The keyword *attention* is also found in Bouffard and Sarkar's definition of metalinguistic awareness, where attention to explicit knowledge of the language is emphasized (2008, cited in Simard & Gutìerrez, 2018, p. 205). Other researchers, such as Jessner (2008) defines metalinguistic awareness as an "ability to focus on linguistic form and to switch focus between form and meaning" (p. 277). These definitions are still applicable today and similar to the one provided for this study. Following Tunmer, Pratt & Herriman (1984), the definition the present study draws on is the individual's ability to distance themselves from the language itself and to evaluate, analyze, and manipulate the linguistic features (p. 289).

Two concepts, namely the process of analysis and the control of processing, are necessary to mention when addressing metalinguistic awareness. The process of analysis refers to the ability to systematize explicit knowledge and complex structure in a language (Roehr-Brackin, 2018, p. 20). Explicit knowledge is knowledge one has organized in known systems whether or not one is conscious about this knowledge. Control of processing refers to control to direct attention to specific aspects of, i.e. a linguistic problem, like ambiguity or conflict (Bialystok, 2004, p. 131-132). These two concepts are especially essential to address when dealing with an evaluation of metalinguistic tasks since such tasks require high levels of control and attention (Bowey, 1988; Lundberg & Torne´us, 1978; Tunmer & Herriman, 1984 cited in Bialystok, 2004, p. 132).

2.3 Previous Research

Metalinguistic awareness has been studied in the fields of linguistics, developmental psychologists, and educational psycholinguists (Pinto, Titone and Trusso, 1999, cited in Jessner, 2008, p. 258). Early studies focused on monolingual children, however more recent studies have studied MLA in bi- and multilinguals. This thesis focuses on metalinguistic awareness in children with English as their L2 and both mono- and bilingual children has taken part in the study. Research addressing MLA and metalinguistic knowledge and their correlation with language competence and L2 acquisition will be presented. Additionally, a presentation of research regarding possible advantages in MLA for bilinguals will be given.

Several studies and research projects have aimed at raising linguistic awareness in schools. One of them is the Evlang Programme (EVLANG). EVLANG was a European programme of pedagogical innovation and research. The project aimed to examine whether the activities of language awareness that was conducted in Primary School achieved the expected effects (Andrade, 2004, pp. 28-30). The project later became an organization that involved teachers, children (ages 10-12) and researchers in numerous European countries. The introductory objective was to expand the children's language awareness, which was done with, amongst other things, tasks revolving language knowledge (Roehr-Brackin, 2018, p. 47). According to Jones (2005, cited in Roehr-Brackin, 2018, p. 48), the project has been viewed as a success and has received positive feedback from the projects' participants.

In *Metalinguistic Awareness and Second Language Acquisition* (Roehr and Gánem-Gutiérrez, 2013), Karen Roehr and Gabriela Adela Gánem-Gutiérrez present studies that have investigated metalinguistic knowledge and metalinguistic awareness. The studies revolving cognitive approaches to the metalinguistic dimension and measures representing an educational perspective will be highlighted, with emphasis on Angela Tellier. Tellier (2013) investigated children's metalinguistic awareness who had English as their L1 (first language) and developed a paper-and-pencil test for this purpose. Some of the tasks are based on natural languages and some on artificial languages. The tasks differ in cognitive complexity, which means that the tasks vary in the number of mental manipulations required to be able to complete the task. The tasks require the participants to demonstrate a high degree of analysis and control in order to complete them successfully (pp. 24-25). The tasks are divided into 1) tasks based on European languages and 2) tasks based on an artificial language. The tasks draw on, i.e., lexical semantics, ambiguity, cognates, morphology and translations (Roehr-

Brackin, 2018, p. 128). According to Tellier (2013), the test had no noticeable ceiling effect (p. 28), meaning that the tasks were not too difficult. The test has been used in several studies and has demonstrated good reliability overall (Tellier, 2015, cited in Roehr-Brackin 2018, p. 128). The presented results demonstrated that there was a definite improvement of metalinguistic awareness over the years, which is consistent with earlier research that has found that MLA improves with age and cognitive development (Hakes, 1980; Bialystok, 1988, cited in Tellier, 2013, p. 23). However, a more significant difference in means was found between children between the ages 8 and 9, and again in children between the ages 11 and 12. A similar study was conducted by Simard, Foucambert and Labelle (2013). However, their approach was to find a potential correlation between metasyntactic abilities and syntactic abilities and reading comprehension. The study's participants were L1 (first language) and L2 (second language) speakers of French from the ages 8 to 12. With metasyntactic ability, Simard et al. refer to the ability to reflect on the syntactic aspect of languages such as words, grammatical classes and rules governing their combinations (Bowey, 1986, 2005, cited in Simard et al., 2013, p. 47). Accordingly, metasyntactic abilities are metalinguistic abilities associated with syntax. Metasyntactic abilities has been claimed to support reading comprehension because a fair level of metasyntactic abilities will allow one to predict the sequence of words in sentences as well as promote a reassembly of decoded words into phrases and thereby making recall easier (Tunmer et al., 1984, cited in Simard et al., 2013, p. 60). Syntactic ability is the ability to decode a language by pairing phrases and sentences. This processing of linguistic information is incremental as the individual must integrate incoming words syntactically and semantically in order to derive meaning (Koda, 2007, cited in Simard et al., 2013, p. 47). Syntactic abilities have shown to correlate with metasyntactic abilities and with reading skills among native and non-native speakers. Simard et al. (2013) aimed to examine if metasyntactic abilities syntactic abilities had a significant role in explaining reading comprehension skills in French-speaking native and non-native children. Moreover, they wanted to understand the relationship between metasyntactic abilities and syntactic abilities in the modeling of reading comprehension of these children. After 37 French-native speakers and 36 non-native speakers completed a series of measures, they found that metasyntactic abilities have a significant contribution to the construction of reading comprehension for both participant-groups. Concerning syntactic abilities, they found that it did not make a significant contribution to reading comprehension for either of the groups. However, they state that it contributes indirectly through metasyntactic abilities. The findings in the contributions of metasyntactic abilities in reading comprehension shows that an ability

to reflect on the components in a language such as words, grammatical classes and rules demonstrates that it plays an important role when learning to read. (Simard et al., 2013, pp. 45-70).

Lastly, Thepseenu and Roehr (2013) did a study that investigated students at a Thai university concerning their beliefs about metalinguistic knowledge. Also, they examined the students' metalinguistic performance to see if there were any correlations between their beliefs and their performance. The students had English as their L2. The participants answered an English placement test, a test of L2 metalinguistic knowledge and a questionnaire. The questionnaire was included to be able to determine the student's beliefs about metalinguistic knowledge. Finally, a sub-sample of the group (in total 64 participants) were interviewed to gather a deeper understanding of these beliefs. Thepseenu & Roehr (2013) found that all participants had some metalinguistic knowledge about the targeted grammar features. They also found that most of the students believed that grammar rules for writing and reading were useful, and that studying grammar is crucial to learn the language. However, some claimed that the grammar-rules sometimes interrupted the communication-flow. The focus on grammar while speaking could intrude the fluency of speech and affect the student's confidence. The findings showed a correlation between the student's beliefs and their metalinguistic knowledge. The findings demonstrated that the students with negative beliefs in the role of grammar also received less successful test scores. (Thepseenu & Roehr, 2013, pp. 90-109).

2.4 The relationship between metalinguistic awareness and language learning

Simard and Gutìerrez (2018) have presented an overview of research on metalinguistic constructs in second language acquisition (SLA) in *The Routledge Handbook of Language Awareness* (pp. 205-221). A short description of the studies deemed relevant for this study will be provided in this sub-chapter as one of this study's research questions investigates the correlation between metalinguistic awareness and the participants' English competence.

Alderson et al. (1997, cited in Simard & Gutìerrez, 2018, p. 209), investigated how metalinguistic knowledge related to language competence. They operationalized metalinguistic knowledge as identification of speech parts, correction and explanation of errors. They found that there was a significant correlation between error correction and grammar, reading and writing. A different study done by Elder et al. (1997, cited in Simard & Gutìerrez, 2018, p. 210) used the same tests as Alderson et al. (1997). Their results showed a

moderate to a weak correlation between metalinguistic knowledge and L2 achievement. Later, Elder and Manwaring (2004, cited in Simard & Gutièrrez, 2018, p. 210) measured metalinguistic knowledge using equivalent tests as in Alderson et al. (1997) and Elder et al. (1997) to investigate pupils' L2 learning success. The pupils were divided into groups who received different types of instruction. Their results showed that the pupils who had received explicit instruction had a higher success rate in the measurements of metalinguistic knowledge than the communicative groups. Moreover, Masny (1987, cited in Simard & Gutièrrez, 2018, p. 210) found positive correlations between metalinguistic awareness and L2 learning. She found that the ability to identify grammatical sentences was related to language aptitude, reading competence and L2 classroom achievement. Lastly, Renou (2001, cited in Simard & Gutièrrez, 2018, p. 210) examined metalinguistic awareness in tasks that investigated identification, correction and explanation of errors. She found a positive correlation between metalinguistic awareness and global L2 competency for the pupils that had received explicit instruction but not for the communicative group. However, both groups showed a positive correlation in the tasks that addressed grammar, structure and targeted vocabulary (Simard & Gutièrrez, 2018, p. 210-211).

2.5 Metalinguistic awareness for educational purposes

According to Theepsenau and Roehr (2013), several studies support the statement that a high level of learner awareness is associated with improved performance in the learner's L2 (p. 91). This statement is also supported by Simard and Gutièrrez (2018) who claim that development in children's metalinguistic awareness is beneficial for some aspects of second language acquisition (p. 216). Recent studies have investigated how to improve children's metalinguistic awareness. Some approaches have been a focus of error analysis, contrasting children's L1 and L2 and teaching children a *starter language* before their L2 learning (Roehr-Brackin, 2019, p. 53). Children at ages 2-3 already start to experiment with language and can draw attention to specific morphological characteristics of a word. Word-games and riddles are excellent mediums for learning how to manipulate the linguistic features of a language. They can be encouraged to play with the language in using minimal-pairs or by inventing new compounds, e.g. *rainbrella* for *umbrella* (Birdsong, 1989 cited in Roehr-Brackin, 2018, p.6). In riddles and jokes, the children will have to be able to detect the grammatical feature that makes the jokes or riddle humorous. There might be a hidden double-meaning, a metaphor, an idiom or ambiguity. In other words, to be able to understand

riddles and jokes, one must have a certain level of metalinguistic awareness and ability (Shade, 1991, cited in Zipke, 2008, pp. 130-131). According to the Norwegian Directorate for Education and Training (NDET), the teachers are required to support the pupils in their development of meta-reflection of their own learning and metacognitive learning strategies (NDET, 2019, p. 2). Also, the English subject curriculum states that the pupils should have knowledge of how word, sentence and text structures are used and how the English language is structured after grade 7 (NDET, 2019, p. 2). Moreover, in the core elements of the English course, it is stated that "language learning involves developing language awareness" and "knowledge of English as a system and being able to use language-learning strategies" (NDET, 2017, p. 2). As stated in section 2.1, language awareness and metalinguistic awareness are sometimes used interchangeably. Language awareness is, however, a more collective term than metalinguistic awareness and can also include, i.e., metalinguistic knowledge.

2.6 The "bilingual advantage"

One should not underestimate the complexity of defining bilingualism. In research, bilingualism is not a categorical variable that can be compared to the variables grade, age or gender. According to Bialystok (2001, p.8), bilingualism is, at best, a scale where on one side of the scale, one will have nearly no awareness that other languages exist. On the other side of the scale, one will have fluent proficiency in two languages. In which Bialystok questions when a person can claim to be bilingual. In this thesis, bilingualism is defined as the ability to have functional fluency in two, or more, languages. This definition corresponds to Bialystok: "[…] their functional proficiency in the two languages is equivalent – they can carry on conversations and engage in the same kinds of activities" (2001, pp. 19-20). There has been a dispute in whether bilinguals have an advantage in acquiring new languages or not.

Early studies in the 1950s and 1960s indicated that bilinguals experienced a cognitive deficit. Though, studies in the 1970s disproved these findings, pointing to a positive correlation between cognitive development and linguistic advantages to bilingualism (Bialystok, 1991, p. 92). Vygotsky (1986 cited in Jessner, 2008, p. 358) stated that by learning a foreign language, a child could further develop s/he's native language. There have been many studies on the positive relationship between cognitive development and bilingualism (Jessner, 2008, p. 358). However, according to Bialystok (1991, p. 91) some of the studies might have been too enthusiastic. It is important to note that it is not a universal pattern that points to a clear

bilingual advantage overall when it comes to language awareness. Still, Bialystok states that the *control over attention* may develop more advantageously in bilingual children (2004, p. 248), which is an important component in metalinguistic awareness. Cummins (1991, cited in Jessner 1999, p. 202) explains the bilingual system using a metaphor: an iceberg. He brought attention to a common underlying proficiency in languages in bilinguals. This proficiency is not part of the speaker's L1 or L2, but a linguistic container that can be used in both languages and helps compare languages more easily. Bilingual children probably develop *control of processing* to a higher degree than monolingual children, especially if the children's L2 is acquired in an academic environment (Bialystok 1986, cited in Bialystok 1991, p. 92).

In short, several studies investigating metalinguistic awareness- and knowledge found positive correlations with, inter alia, SLA acquisition, reading comprehension and error correction. Following Simard and Gutièrrez (2018) and Theepsenau and Roehr (2013), a positive development in children's metalinguistic awareness is advantageous for some aspects of SLA. On the issue of bilingualism, Bialystok (1991, 2004) found that there might be a more favorable development of *control over attention* and *control of processing* in bilingual children, which is essential components in metalinguistic awareness.

3.0 Method
3.1 Research Project

The thesis aims to investigate the state of MLA in one 6th grade and one 7th grade. The two grades will be compared to investigate whether there is a clear difference in means between the two. The age group has been chosen due to Tellier (2013), who found a significant difference in means between grades 6 and 7. The increase in means was not found between the ages 10 and 11, which is why the statistical significance between grades 6 and 7 drew this study's attention. An MLA-test has been developed to be able to investigate the participants state of MLA. The MLA test draws on the already developed MLA-test of Tellier (2013) and EVLANG (1997-2001). The test has been adapted to suit the participant group, which are pupils from 11 to 12 years with English as their L2. This chapter provides a presentation of the methods and materials utilized. The chapter also describes the tasks and analysis, as well as how the test was conducted.

3.2 Theoretical drive

The theoretical drive is the conceptual direction of the project overall. This study has a quantitatively driven, mixed-method design where the core component is a measurement (Morce & Niehaus, 2009, p. 9-11). The study collects and analyzes data both quantitively and qualitatively. The main part of the analysis emanates from a quantitative scoring-system. In this study, one of the aims is to test a possible theory, which is whether there is a clear difference in MLA in grades 6 and 7 (ages 11 and 12). The study's main aim is to examine the relationship between variables that can be measured on instruments. Resulting in analyzing the numbered data using statistical procedures (Creswell, 2015 and Newman & Benz 1998 cited in Creswell & Creswell, 2018, p. 4). The supplemental qualitative analysis provides a fuller understanding of the research question and collected data. The qualitative analysis is also implemented to shed light on reoccurring tendencies in the participants' answers. Qualitative research is usually exploratory and detail-oriented, where the process of research involves emerging questions and in-depth investigation (Creswell, 2015 and Newman & Benz 1998 cited in Creswell & Creswell, 2018, p. 4). In this study, the researcher will make interpretations of the meaning of the data which is a typical qualitative process. Mixed methods research integrates qualitative and quantitative research. The core assumption in this method is that the integration of qualitative and quantitative data yields additional insight beyond the information provided by either the quantitative or qualitative data alone. The combination of the two can be done in several ways, but one of the methods tend to be more dominant than the other. This design relies on the scoring-system to a great extent and is more quantitative than qualitative.

3.3 Participants

Forty seven (47) participants in grades 6 and 7 were invited to join the research project. Of these, 40 participants took part in the project. Specifically, 17 participants in grade 7 and 23 participants in grade 6. The participants were 11 and 12 years old except one pupil who was 10 years old. The categorization in age is an indispensable step toward making empirical conclusions (Bialystok, 2001, p.8). The participants were L2 learners various L1's. Two participants in grade 7 had a different language than Norwegian as their L1. In grade 6, nine participants had a different language than Norwegian as their L1. The school was chosen due to the researcher's prior knowledge of the school and the pupils' English teachers, which is a typical convenience sample. The group of participants was chosen due to the claimed difference in means between the two age groups. As this is a small-scale study, the sample of

participants is not representative for a generalization of the results. Results may vary from school to school due to, amongst other things, sociodemographic variables. However, the sample is big enough to state indications that may apply to other pupils of the same ages.

3.3.1 The "bilingual advantage"

This thesis required the participants to self-report their fluency in other languages. The categories were *fluent, good* and *some* proficiency in oral and in written form of the language(s). They were also questioned about their language habits at home, with friends, when playing video games and watching TV. These questions were included on the last two pages of the MLA-test and is depicted below (Figure 1 and 2). The researcher can use this information when discussing any apparent differences in results between those who claim to be bilingual and those who claim to be monolingual. This study does not consider participants who reported fluent proficiency in Norwegian and English as bilinguals since they are L2 learners.

Noen spørsmål om deg:

Hvor gammel er du?

Jeg er _____

1. Er du gutt, jente eller identifiserer du deg som noe annet?

Gutt ☐   Jente ☐   Annet ☐

2. Hvor lenge har du bodd i Norge?

Hele livet ☐   Mindre enn to år ☐   Mindre enn ti år ☐

3. Hvilke språk snakker du?
Sett et kryss under "muntlig" og et kryss under "skriftlig" for hvert av språkene.

Språk	Muntlig			Skriftlig		
	Litt	Bra	Flytende	Litt	Bra	Flytende

Figure 1: Language background, p. 1

4. Hvilke språk bruker du oftest i følgende situasjoner:

Når du forteller moren din om dagen på skolen:

_____

Når du leker med søskene dine hjemme:

_____

Hvis du vil kjøpe noe og må be faren din betale for det:

_____

Når du forteller vennene dine om et program du har sett på TV:

_____

Når du spiller dataspill:

_____

Når du ser på TV eller YouTube:

_____

Figure 2: Language background, p. 2

3.3.1.2 Bilingualism in grade 6

There were ten participants in grade 6 that reported fluent proficiency in another language than Norwegian. These languages were English, Turkic, Arabic, Swedish and Somali. Five of these participants reported that they spoke an alternative language to Norwegian to at least one of their parents. These languages were Somali, Turkic, Arabic and Danish.

3.3.1.3 Bilingualism in grade 7

In grade 7, six participants reported fluent proficiency in another language than Norwegian. These languages were Turkic, Arabic, Swedish and Somali. Three of these participants reported that they spoke an alternative language to Norwegian to at least one of their parents. These languages were Arabic and English.

3.4 Teacher Assessment

The teachers were asked to assess the participants' English competence on a scale from 1-6. Both teachers received the same instructions, which were to assess their writing- and reading competence and comprehension in the course. This was done to be able to answer whether there is a correlation between the participants' MLA and their teacher-reported English competence. The study must take reservations from making distinct conclusions regarding the relationship between MLA and English competence as the two teachers may have interpreted the instructions differently or assessed the participants on unequal terms despite the given instructions. For the assessment, the teachers were given a class chart with numbers on each desk. The numbers on the class chart corresponded with the numbers labeled on the desks and the tests. A list of numbers was enumerated below the class chart, allowing the teachers to write the number (1-6) alongside the participant-number. In this manner, the teachers could assess the participants without disrupting the participants' anonymity. The teachers were informed about the assessment a week before the tests were conducted to ensure that the teachers were prepared to give the participants a reflected assessment. The teacher assessment and a comparison of the participants' MLA scores are presented in section 4.8 and followed by a discussion in section 5.0.

3.5 Data Collection

3.5.1 Pilot test

Five weeks before the study was conducted, a pilot test was carried out. The pilot test was conducted on 10 participants in the same age group as the participants in the study. Following Mackey and Gass (2016), it is essential for researchers to distribute time for conducting pilot tests (p. 53). The pilot test was carried out to uncover any problems with the instructions or tasks, and to address these issues before the main study was carried out. The participants' given timeframe was equivalent to the time frame the main study's participants were given, as described in section 3.5.2. The researcher made notes during and after the test. Afterward, the questions that emerged from the participants were evaluated to ensure a precise wording in the task instructions. The results were analyzed primarily to prevent any floor effect or ceiling effect. One speaks of a floor-effect if the tests' questions or tasks are too difficult, resulting in most of the participants scoring near to the bottom. Likewise, a ceiling-effect is when the questions or tasks are too easy for the participants, resulting in most of the participants receiving perfect scores (McCabe, 2017, p. 2). Floor-and ceiling effects are essential to avoid

to secure differentiating results. The analysis assured that the tasks varied in difficulty, without any floor- or ceiling effects, as earlier mentioned. Consequently, no tasks were excluded. However, the wording in the task instructions was modified to be more precise.

3.5.2 Conducting the test

The test was conducted in the first two lessons of the day in both classes. The desks were labeled with a number from 1-23, corresponding with the numbers written on the front page of the tests. The participants were first given general information about the project and also instructed that they were not to turn the page before they were instructed to. The participants were given a time warning when it was approximately one minute left before moving to the next task. There was a timer set on five minutes per task, but the exact time varied on each task to ensure a reasonable pace. If a participant was done before the five minutes, they could use the time in looking through their answers for the respective task. After completing a task, the pupils could not go back to the task at a later time. The participants were informed that if they had questions, they could ask the researcher. However, the researcher had to consider what questions could be answered concerning the tests' validity. Therefore, if they had formal questions such as where to write the correct answer or if they were supposed to answer in the artificial language, they were given the answer. Questions based on giving them hints to the correct answer were not answered. The participants were not allowed to talk to each other. In grade 7, there were no issues in this regard. However, there were some disturbance in grade 6. Whether this threatens the validity of the answers will be discussed in section 5.1.

3.6 The MLA-test

There were several issues to address before the design of the test was finalized. One of them were if the test should be language-specific or not. By language-specific, it is meant whether the test should keep to one specific language in all tasks. In this case, in Norwegian or English. However, the participants varied in Norwegian competence as some did not have Norwegian as L1. This could have affected the results and possibly interfered with the MLA-scores and is one of the reasons why the MLA-test is language independent. The test draws on different natural languages as well as artificial languages, which is illustrated in appendix 1. This way, the participants' competence in e.g., Norwegian would not interfere with the results as it could have if the test tasks were in Norwegian. The tasks in the test build on the test developed by Tellier (2013), which were developed for participants with English as their L1

aged 8-11. Also, two tasks from EVLANGs' (1997-2001) MLA test were adapted and included in the developed test. The constructs investigate the awareness of crosslinguistic metalinguistic awareness, inflectional morphology- and word formation, and word-order. There is included one simple and one complex task per construct. The more complex task requires the participants to utilize several mental manipulations (Stankov, 2003, cited in Tellier, 2013, p. 24). Measures of metalinguistic awareness that are intended for young learners tend to make high demands on both analysis of knowledge and analysis of processing (Roehr-Brackin, 2018, p. 127), which both Tellier's MLA-tasks, as well as this study's tasks, have implemented. According to Bialystok (2001), analyzing the grammatical acceptability of a sentence is presumably the archetypical metalinguistic task (p. 139). This sub-chapter will explain each task as well as describe the adaptations. The quantitative scoring system will also be explained for each task. Each task explanation starts with a picture of the task-component. Thereafter, the test will be depicted as a whole in section 3.6.8 to provide the reader with a clear overview of the test.

ITEM	MLA task (operationalization)	Degree of complexity	MLA	Language	Construct investigated: awareness of….	Linguistic domain
Tellier item 1	cognates	simple	Crosslinguistic awareness	**Natural languages** Spanish Portuguese Italian Dutch Romanian Esperanto English Norwegian	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 2	cognates	complex	Crosslinguistic awareness	**Natural languages** Italian German Esperanto	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 3	Plural marking	simple	General MLA	**Artificial languages**	Systematicity in inflectional morphology	inflectional morphology
Tellier item 7	Case marking	complex	General MLA	**Artificial languages**	Systematicity in inflectional morphology; additional element: what constitutes a sentence (word order)	inflectional morphology & syntax
EVLANG item 1.5	Chinese numbers	Simple	General MLA	**Natural language** Chinese	Systematicity in word formation	Word formation (compounding)

| EVLANG item 2.5 | Sentence in unknown L | complex | General MLA | **Artificial language** | Systematicity in what constitutes a sentence (word order) | Syntax |

Appendix 1: Table of tasks and its components

3.6.1 The MLA-tasks

3.6.2 Task 1 Sentence pairing and translation

1:

Hvilke setninger betyr det samme?

1. El gato es blanco.
2. O elefante é grande.
3. La rana è verde.
4. De olifant is groot.
5. Tigrul este negru şi verde.
6. La rano estas verda.
7. The cat is white.

Par:

☐ and ☐

☐ and ☐

☐ and ☐

Setningen som er igjen betyr:

Norsk: _____

Engelsk: _____

Figure 3: Task 1

Task 1 is the simple task component of the investigation of awareness of lexical relationships between languages. Task 1.1 ask the participants to pair words from seven sentences in different languages. The languages are Spanish, Portuguese, Italian, Dutch, Romanian, Esperanto and English. They are supposed to find three word pairs. The second component of the task (1.2) asked the children to translate the sentence that was not matched with another sentence. They were first asked to translate this sentence in Norwegian, and then in English. This task seeks to investigate the participants' crosslinguistic metalinguistic awareness.

The tasks' sentences have remained the same as in Tellier's' (2013) test. However, the second part of the task has been extended to ask the participants to translate the sentence in both English and Norwegian.

In 1.1, the participants get one point per correct sentence-pair matched (maximum 3 points). The correct sentence-pairs are 1 and 7, 2 and 4, and 3 and 6. The remaining sentence is, therefore, sentence number 5: *Tigrul este negru si verde.* In task 1.2, the participants can receive a maximum score of 12 points. 6 points for the English translation and 6 points for the Norwegian translation. The participants should be left with the sentence: *Tirgul este negru si verde* which translates to *The tiger is black and green/Tigeren er svart og grønn*. The participants received one point for each word correctly translated. The participants received 2 points for translating *tigrul* to *tigeren,* as the *-en* suffix indicates that there is a definite article. The maximum score for task 1 was 15. 3 points for task 1.1 and 12 points for task 1.2.

3.6.3 Task 2 Translation IT, GE, ESP

2. Benytt alle språkene du kan for å hjelpe deg med å finne ut hva setningene nedenfor setningene betyr. Skriv hele setningen, eller så mange ord som du kan.

Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato.

_____

_____

Der Hund, die Katze und die Maus sind sehr gute Freunde.

_____

_____

Estas bela tago do mi volas vanilan glaciaĵon kaj limonadon

_____

_____

Figure 4: Task 2

The second task is the second component of crosslinguistic awareness, which is the complex task. The task presents three sentences in three different languages (Italian, German and Esperanto) and asks the participants to translate them to Norwegian. They are told to try to write the whole sentence or as many words they can. The instruction emphasizes that it is fine to guess the answer if they are unsure. This task seeks to investigate the participants' awareness of lexical relationships between languages.

3.6.3.1 Adaptation

The study's participants speak Norwegian. Therefore, the translation tasks were examined considering the association with the Norwegian words. The second sentence had the German *Freunde* in it. The word *Freunde* draw on the word *friend* in English but is not comparable with the Norwegian *venn.* The German *Freunde* remained in the sentence as the English *friend* is a high-frequency word that one expects the participants to know at this age. The test excluded the Esperanto *aĉeti* in the third sentence since it was deemed unnecessary for the sentence's meaning and it does not draw any association to the Norwegian language. As for the scoring, some of the scoring-criteria was excluded in this study. In Tellier (2013), the participants could receive an additional point if "the answer, whether or not it is an accurate translation, is written as a complete sentence" which this study chose to discard as the instructions in the test did not require the participants to write the whole sentence. The same applied to the additional point given if "the answer displays sensible meaning, whether or not it is an accurate translation", as this is a subjective assessment that could affect the overall fairness of the scoring (Tellier, 2013, p. 4). However, the participants who translated *caffè con latte* as *kaffe latte* got two points. One point for the keyword *kaffe* and one point for the keyword *latte.* The decision to give the participants a point for l*atte* (milk) is to ensure a fair scoring as *kaffe latte* is a common phrase in Norwegian. This exception was applied to all participants in grades 6 and 7.

3.6.3.2 Scoring

The participants can receive a maximum of 7 points in each sentence. There are 6 keywords in each sentence. The participants receive one point each keyword correctly translated. In addition to this, the participants receive an additional point if they translate at least one word other than the keywords. In task 2.1, the participants are introduced with a sentence in Italian: *Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato.,* which in English

translates to: *I have a cup of [white / milky] coffee with milk and sugar, and a chocolate cake,* and in Norwegian: *Jeg har en kopp med kaffe og melk [kaffe latte] med melk og sukker, og en sjokoladekake.* The keywords in Norwegian are *kopp, kaffe, melk/latte, sukker, kake* and *sjokolade.* In task 2.2, the participants are required to translate a sentence in German: *Der Hund, die Katze und die Maus sind sehr gute Freunde,* which in English translates to: *The dog, the cat and the mouse are very good friends,* and in Norwegian: *Hunden, katten og musen er veldig gode venner.* The keywords in Norwegian are: *hund, katt, mus, veldig, gode* and *venner.*In task 2.3, the participant are asked to translate a sentence in Esperanto: *Estas bela tago do mi volas aĉeti vanilan glaciaĵon kaj limonadon,* which in English translates to: *It's a beautiful day so I want a vanilla ice-cream, and lemonade,* and in Norwegian: *Det er en nydelig/vakker dag så jeg har lyst på en vanilje-is[krem], og en limonade.* The keywords in Norwegian are: *vakker/nydelig, dag, vil, vanilje, is[krem]* and *lemonade.* The maximum score for task 2 is 21 points.

3.6.4 Task 3 Plural and singular nouns
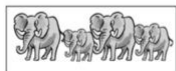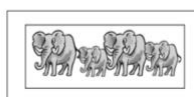
Figure 5: Task 3.1

7. rahaki

8. samini

Figure 6: Task 3.1

....

Nedenfor står et ord fra det samme språket som ordet "samini" er
fra. Arkari er ordet for **en.** Hva sier du om det er **mer enn en?**
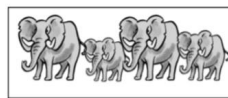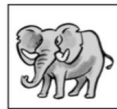
Arkari

Figure 7: Task 3.2

Task 3 investigates the awareness of systematicity in inflectional morphology. The task is
divided into two components (3.1 and 3.2) where 3.1 presents 8 words in an artificial
language. 4 of the words are singular nouns, whereas 4 of the words are plural nouns. The
participants are asked to pair the singular nouns in a box to the left and the plural nouns in the
box to the right. The participants are also given an example before the task instruction. Task
3.2 is a production task, where the participants must produce the plural noun of a word in an
artificial language. They are told the word is in the same language that the word *saminii* is,
which they worked on in task 3.1.

3.6.4.1 Adaptation

In Tellier (2013), task 3.1 utilized the noun *elephant* in four different languages. This study
produced four artificial nouns (in total eight nouns if both plural and singular form is
included). The nouns that belonged together had the same root-word but different suffixes.
This study decided to produce artificial nouns that were more similar to each other to simplify
task 3.1 before introducing them to the complex task (3.2). This study produced task 3.2,
which was incorporated to test if the participants were able to replicate the grammatical rule
and produce an inflection for the plural noun.

In 3.1, the participants are given one point per correctly identified noun. The order of the nouns does not matter. The singular nouns are: *oligare, garbatan, samini* and *rahak.* The plural nouns are *oligares, garbatans, saminii,* and *rahaki.* In task 3.2, the participants are required to add the correct suffix to the noun *Arkari* to mark it as a plural noun. The plural form of *Arkari* is *Arkarii.* The participants are given 2 points for the correct suffix of the target language. If the participant has added a non-target-like inflection (-s, -es, -í, and so forth), s/he is given 1 point. The participant is given zero points if there is no suffix in their answer.

3.6.5 Task 4 Case marker
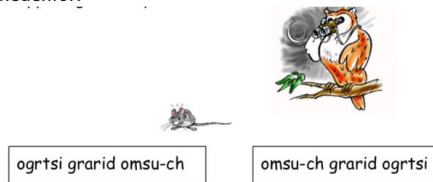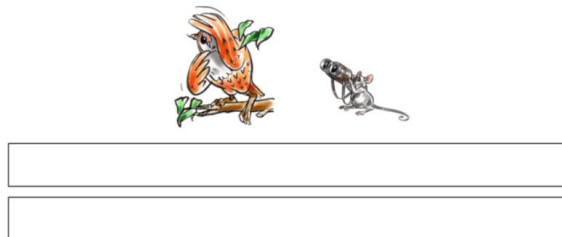
Figure 8: Task 4

Task 4 is a complex task that investigates the awareness of systematicity in inflectional morphology. In addition to this, it also investigates awareness of what constitutes a sentence (word order). The task presents a picture of an owl and a mouse and introduced an artificial language that describes the owl's and mouse's action (the owl looks at the mouse). However,

the sentence was written in two ways. *Owl* was written as the initial word in the first sentence, whereas *mouse* was written as the initial word in the second sentence (SVO-word structure and OVS word-structure). The task draws on the awareness of inflectional morphology, where the task requires the participants to apply a morphological rule. In this case, the object is marked with a *-ch*-suffix. In this language, the change of word-order will not change the meaning of the sentence. Then, the participants are shown a picture of the same action, with reversed roles from the example, and are asked to describe the picture in two ways.

3.6.5.1 Adaptation

There was no adaptation for task 4, other than the change of language and wording in the task instruction.

3.6.5.2 Scoring

In task 4, the owl (ogrtsi) is the object in the sentence. Therefore, the owl should obtain the -*ch*-suffix: *Ogrtsi-ch grarid omsu* and *omsu grarid ogrtsi-ch*. If the participant produces two sentences with the correct addition of the *-ch* suffix and in the correct word order, they are given 2 points. If one of the sentences are produced correctly, they receive 1 point. If none of the two sentences has the correct suffix, the participants are given zero points.

3.6.6 Task 5 Chinese numbers

5. Nedenfor ser du tall som er skrevet på kinesisk.

1	yi
2	er
3	san
4	si
5	wu
10	shi
11	shiyi
12	shier
20	ershi
23	ersisan

Bruk listen med tall som hjelp og eksempelet som modell for å skrive numrene til de kinesiske tallene nedenfor.

Eksempel: shisan: 13

shisi: _____

ershiwu: _____

sanshisan: _____

Figure 9: Task 5.1

Nå er det din tur. Bruk eksempelet nedenfor som modell og skriv følgende tall på kinesisk:

Eksempel: 13: shisan

15: _____

24: _____

....

Figure 10: Task 5.2

Task 5 is the first task from EVLANG and is the simple task-component of the investigation of systematicity in word formation. This task is divided into task 5.1 and 5.2. The task instruction has included a list of numbers with the Chinese words for each number to the right. In 5.1, they are given three Chinese words and are asked to write down which numbers these words represent. In 5.2, the participants have to write down the Chinese words corresponding to two numbers they are given.

3.6.6.1 Adaptation

There were no adaptations for task 5, other than the change of language and wording in the task instruction.

3.6.6.2 Scoring

The maximum score for task 5 is 5 points. The participants are given 1 point per correctly identified number. The maximum score for task 5.1 is, therefore, 3 points. The three numbers are 14 (*shisi*), 25 (*ershiwu*) and 33 *(sanshisan).* Both ciphers in each answer need to be correct to receive a point. The same applies to 5.2, where two numbers needs to be written in Chinese. These are *shiwu* (15) and *ershisi* (24). Accordingly, the maximum score for task 5.2 is 2 points.

3.6.7 Task 6 Translation and word order in *Krarid*

6. Her er tre setninger på språket "krarid". Du finner den norske oversetningen nedenfor:

En	inu	betsiaki	*Jeg har sett jaguaren*
Jeg	*jaguaren*	*har sett*	

En	baka	betsiaki	*Jeg har sett fisken*
Jeg	*fisken*	*har sett*	

Min	baka	betsiaki	*Du har sett fisken*
Du	*fisken*	*har sett*	

Se nærmere på hvordan setningene er laget og prøv å skriv setningen nedenfor på språket "krarid".

Du har sett jaguaren

_____

Figure 11: Task 6

The last task in the test is also taken from EVLANG. This is a production-task where the participants have to produce a sentence from an artificial language *Kraris*. They are given three sets of words, with the Norwegian translation underneath the word from the artificial language. The participants are asked to produce the sentence "du har sett jaguaren" in the artificial language. Task 6 is a complex task that investigates the awareness of systematicity in what constitutes a sentence.

3.6.7.1 Adaptation

There were no adaptations for task 6, other than the change of language and wording in the task instruction.

3.6.7.2 Scoring

In 6.1, the participants are given 1 point for each correctly produced word, which will give them a maximum of 3 points. The participants are supposed to translate the sentence: *Du har sett jaguaren* to the artificial language *Krarid.* The sentence in *Krarid* is: *Min baka betsiaki.* In 6.2, the participants receive 1 point if the sentence is written in the correct word-order. The correct word order is in a SOV-structure.

3.6.8 The MLA-test in its entirety

The MLA test is depicted in its entirety below to give an overview of the test. All images are included in the appendices in a bigger image size.

1:

Hvilke setninger betyr det samme?

1. El gato es blanco.
2. O elefante é grande.
3. La rana è verde.
4. De olifant is groot.
5. Tigrul este negru şi verde.
6. La rano estas verda.
7. The cat is white.

Par:

☐ and ☐

☐ and ☐

☐ and ☐

Setningen som er igjen betyr:

Norsk: _____

Engelsk: _____

Figure 12: Task 1

2. Benytt alle språkene du kan for å hjelpe deg med å finne ut hva setningene nedenfor setningene betyr. Skriv hele setningen, eller så mange ord som du kan.

Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato.

_____

_____

Der Hund, die Katze und die Maus sind sehr gute Freunde.

_____

_____

Estas bela tago do mi volas vanilan glaciaĵon kaj limonadon

_____

_____

Figure 13: Task 2

3. Se på dette ordparet.

en *elefant* ➡ *elefanto*

Mer enn en elefant ➡ elefantoj

....

Prøv å finne ordpar fra listen nedenfor. Et av ordene i ordparet skal bety **en** og et av ordene i ordparet skal bety **mer enn en.** Skriv numrene på ordparene i boksene nedenfor.

1. oligare

2. garbatáns

3. saminii

4. rahak

5. garbatan

6. oligares
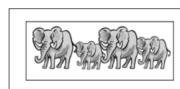
☐ og ☐

☐ og ☐

☐ og ☐

☐ og ☐

Figure 14: Task 3.1

7. rahaki

8. samini

....

Nedenfor står et ord fra det samme språket som ordet "samini" er fra. Arkari er ordet for **en.** Hva sier du om det er **mer enn en?**

Arkari	
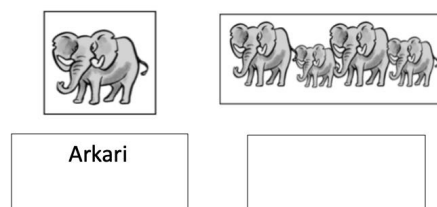
Figure 16: Task 3.2

4. Nå skal vi jobbe med et språk som kalles "ograma". Se på navnene på disse to dyrene.

ogrtsi

omsu

....

Det finnes to ulike måter på dette språket å beskrive hva som skjer på bildet nedenfor.

ogrtsi grarid omsu-ch   omsu-ch grarid ogrtsi

....
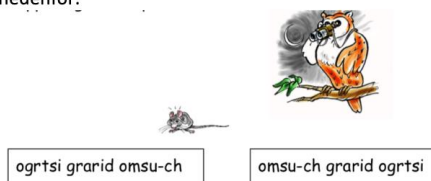
Hvordan vil du beskrive det som skjer på bildet nedenfor? Skriv det på språket "ograma" på to ulike måter.
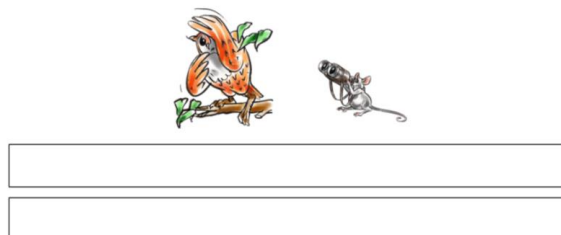
Figure 17: Task 4

5. Nedenfor ser du tall som er skrevet på kinesisk.

1	yi
2	er
3	san
4	si
5	wu

10	shi
11	shiyi
12	shier

20	ershi
23	ersisan

Bruk listen med tall som hjelp og eksempelet som modell for å skrive numrene til de kinesiske tallene nedenfor.

Eksempel: shisan: 13

shisi: _____

ershiwu: _____

sanshisan: _____

Figure 18: Task 5.1

Nå er det din tur. Bruk eksempelet nedenfor som modell og skriv følgende tall på kinesisk:

Eksempel: 13: shisan

15: _____

24: _____

....

Figure 19: Task 5.2

6. Her er tre setninger på språket "krarid". Du finner den norske oversetningen nedenfor:

En	inu	betsiaki	*Jeg har sett jaguaren*
Jeg	*jaguaren*	*har sett*	

En	baka	betsiaki	*Jeg har sett fisken*
Jeg	*fisken*	*har sett*	

Min	baka	betsiaki	*Du har sett fisken*
Du	*fisken*	*har sett*	

Se nærmere på hvordan setningene er laget og prøv å skriv setningen nedenfor på språket "krarid".

Du har sett jaguaren

_____

Figure 20: Task 6

3.7 Qualitative analysis

The results depend on the quantitative scoring-system. However, a qualitative discussion is incorporated in the result-chapter to be able to gather a complete understanding of the collected data. In some instances, incorrect answers may display a certain level of metalinguistic awareness. The researcher will analyze the answers qualitatively to pick up reoccurring patterns in the participants' answers. Additionally, the qualitative analysis makes it possible to discuss whether some of the answers show a level of metalinguistic awareness despite the answer being wrong. Furthermore, the provided discussion can examine the tasks' different components and shed light on plausible thought processes made by the participants, like i.e., in task 6.2. In task 6.2, a great number of participants received zero points. However, by analyzing the answers qualitatively, the researcher found that most of these participants wrote the sentence in a Norwegian word order (SVO) instead of the *Krarid* SOV-strucuture. These answers show that the participants have transferred their knowledge from their source language as a resource to facilitate the target language, which points to a certain level of MLA (further discussed in section 4.6.2).

3.8 Ethics

The participants are aged 11 to 12. The test is a paper-and-pencil test, which means that no IP-addresses were registered when the data was collected. The test is completely anonymous and did not ask about names nor birth date. A consent form with explicit information about the project, as well as an explanation of their rights was handed out six weeks before the study. The consent form is included below (Figure 21). The parents, together with their child, decided whether they wanted to be a part of the study. The consent form stated that they can withdraw from the project without giving any reason and that this decision had no consequences for the child. In this research project, all participants except one accepted to be a part of the study. The participant was given a choice to join the teachers' assistant to play an English-game in a different room or to answer the test without including the test in the study. The participant chose to answer the test. This test was marked and excluded in the study. There were no further issues that had to be addressed since there were no audio- or video recording of the study. The two teachers were fully informed about the study beforehand, as this did not affect the results of the study in any regard.

Figure 21: Consent Form in Norwegian

3.9 Validity and Reliability

This study will assess the study's reliability in light of the measure's and instrument's credibility and accuracy following Krogtoft & Sjøvoll (2018, pp. 99-100). According to Krogtoft & Sjøvoll (2018, pp. 99-100), high reliability is met if other researchers can verify the methods. Simply put, if other researchers use the same methods and get the same results. The study will also have high reliability if accidental errors are small. Reliability also depends on how questionnaires, and in this case, task instructions are worded. Regarding the study's

validity, it will be emphasized whether the measures have measured what has been intended and whether the measures can answer the study's research questions. The degree of correspondence between what has been investigated and what has been aimed to measure is referred to as theoretical validity. The result is theoretically valid if the result builds on theoretical framework and can be explained based on this theoretical framework.

To ensure the study's validity, the study builds on an established theoretical framework about metalinguistic awareness. The study utilized and adapted a scoring-system from pre-existing scoring-systems made of EVLANG (1997-2001) and Tellier (2013). An essential factor is that the study investigates what it wants to investigate and nothing more (Thurén, 2009, pp. 31-32). Therefore, the test was made as language independent as possible. The study was not interested in measuring the participants' English competence nor their Norwegian competence. Consequently, the test depended on various fictitious and natural languages to eliminate the risk of measuring the participant's competence in e.g. English. If the test was made up of English tasks that dealt with the participants' detecting plural and singular nouns and translating sentences from English to Norwegian, the study would not investigate the participants' MLA. The study would likely have established what level of English competence the participants had. The test deals with linguistic constructs like, for example, systematicity in inflectional morphology and systematicity in word order in different (presumably) unfamiliar languages. Therefore, one is able to investigate whether the participants have a metalinguistic awareness since they have to analyze the components in the given language and establish and apply a pattern. To ensure that the study captures all elements/features of the answers, the tests are analyzed both quantitatively and qualitatively, which enhance the study's reliability. Consequently, the quantitative analysis ensures that all answers are treated equally in what concerns the scoring. The established scoring-system can be used by other researchers on a later time. The qualitative discussion was implemented to be able to analyze the test results and discuss whether the answers showed a certain degree of MLA regardless of the answer being correct or incorrect. Additionally, it was to investigate possible patterns in the participants' answers and what that could entail regarding their MLA. The mixed methods ensure the study's validity since it enables the researcher to examine the results from different perspectives. Lastly, the pilot-test was helpful in optimizing the test-instructions. The wording in the test-instruction needs to be as precise and accurate as possible, to avoid any misunderstandings that could affect the participants' answers.

The validity would have been strengthened if the sample included stratification. This means that characteristics such as e.g., income- and education levels of the family members were included (Krogtoft & Sjøvoll, 2018, p. 98). Stratification has not been done in this study because of ethical issues. Such sensitive information could have made the sample of participants smaller or made some of the participants uncomfortable. This is something to be aware of when dealing with participants of this age and is why the study chose to exclude stratification of the sample.

4.0 Results

In the following, the studies' results will be presented, drawing on diagrams and tables as illustrations. The chapter is built up with a presentation of the results task by task. Grades 6 and 7 are presented separately per task. Afterward, a comparative qualitative discussion between the two grades is included, before the next task is elicited. The tables present the number of participants (N) and the mean score, which point to the average score on the respective task. Please note that the mean score is only applicable to those tasks that have a higher maximum score than 1. Those tasks that are categorized as *correct* or *incorrect*, i.e., tasks with a maximum possible score of 1 will be presented with a percentage of pupils who answered the task correctly. Some of the tables will contain the abbreviation N.A, which refers to *not applicable.* This will occur in tasks that have one sub-task presented with a mean score and another sub-task presented with a percentage of pupils who answered the sub-task correctly. The *min.* and *max.* represent the minimum and maximum score a participant received on the respective task. *Range* refers to the difference between the participant who received the lowest score and the participant who received the highest score on the task. Lastly, the *median* is the middle value from the least to greatest score that is given on the task concerned. Also, this study has not differentiated between *wrong answer* and *no answer* in the results. Participants who do not answer the task or answer the task incorrectly will receive zero points.

4.1 Task 1 Sentence pairing and translation

4.1.1 Quantitative analysis

Task 1 builds on crosslinguistic awareness, in investigating whether the participants can identify cognates of English- or Norwegian words in the given sentences. The participants are asked to pair sentences with the same meaning in task 1.1. Task 1.2 NO asked the participants

to translate the remaining sentence in task 1.1 to Norwegian, and task 1.2 EN asked the participants to translate the sentence to English.

4.1.1.1 Grade 6

As Table 1 shows, the mean score for task 1.1 is 2,69 out of the maximum possible score of 3. However, the second part of the task, which asked the participants to translate the remaining sentence in task 1.1 to Norwegian and English, is left with a lower mean score. As Table 1 presents below, the mean score in 1.2.1 (NO) is 1,91 out of a maximum possible score of 6 points. In 1.2.2 (EN), the mean score is 1,34 out of 6 possible points. It is important to note that some of the participants translated a different sentence than *tigrul este negru si verde.* The scoring-system would have to be adapted to catch these translations if some of the translations in English or Norwegian were correct. However, in this case, they were not. The mean score in task 1 is 5,95 out of 15 possible points, with a range of 10 points.

Task	N	Mean	Min.	Max.	Maximum possible score	Range	Median
1.1	23	2,69	1	3	3	2	3
1.2.1 NO	23	1,91	0	5	6	5	2
1.2.2 EN	23	1,34	0	5	6	5	0
Total 1.2	23	4,95	0	10	12	10	2
Total item 1	23	5,95	0	13	15	11	5

Table 1: Descriptive statistics for task 1 in grade 6.
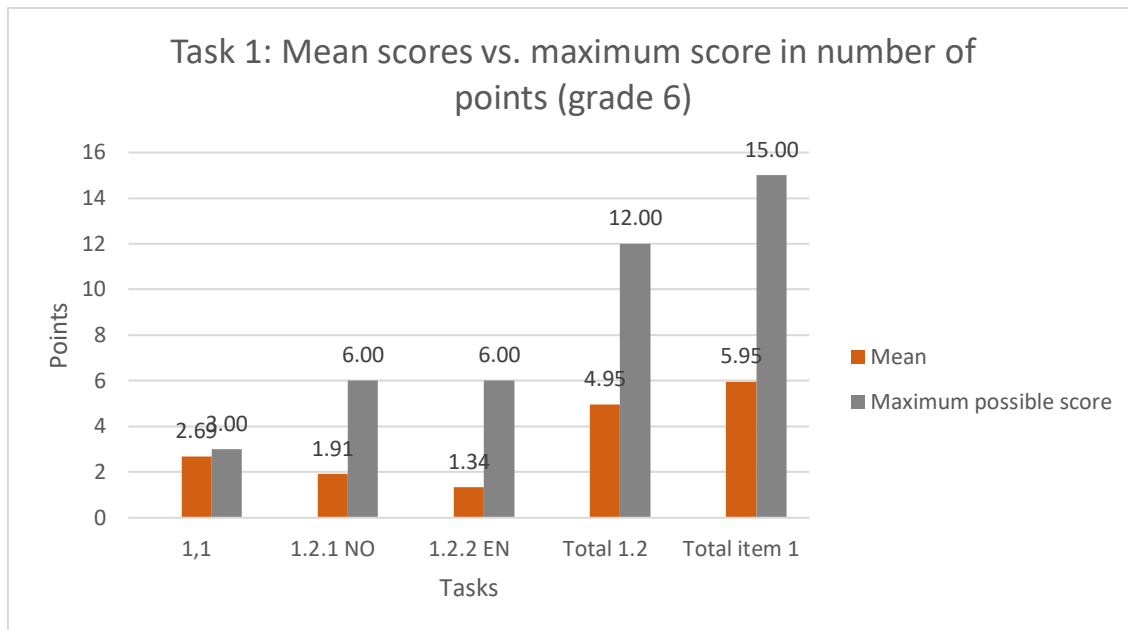
Figure 22: Mean scores vs. maximum score in number of points for task 1 in grade 6.

4.1.1.2 Grade 7

In grade 7, the overall mean score is 0,61 points lower than in grade 6. As Table 2 presents, the mean score in task 1 in total is 5,35 out of a possible score of 15 points. Task 1.1, where the participants are asked to pair sentences that mean the same, has a mean score of 2,35 out of the maximum possible score of 3 points. The range is 3, where some of the participants received zero points, and some received 3 points. The median, however, is 3 points. Task 1.2.1 (NO) has a mean score of 1,58 out of 6 possible points. Furthermore, task 1.2.2 (EN) has a mean score of 1,41 out of 6 possible points. The range is these two sub-tasks (1.2.1 and 1.2.2) is 4, with the minimum score was zero, and the maximum score received was 4 points.

Task	N	Mean	Min.	Max.	Maximum possible score	Range	Median
1.1	17	2,35	0	3	3	3	3
1.2.1 NO	17	1,58	0	4	6	4	1
1.2.2 EN	17	1,41	0	4	6	4	0,5
Total 1.2	17	3	0	8	12	8	2,5
Total item 1	17	5,35	0	11	15	11	4,5

Table 2: Descriptive statistics for task 1 in grade 7.
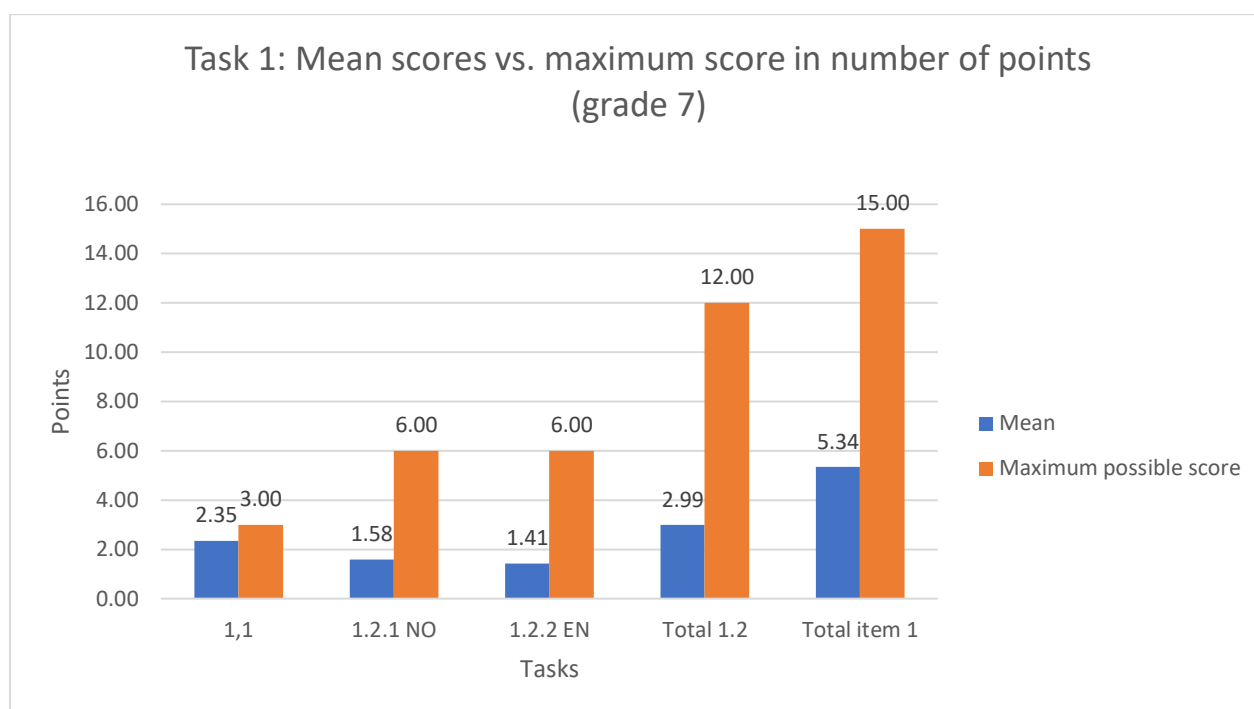
Figure 23: Mean scores vs. maximum score in number of points for task 1 in grade 7.

4.1.2 Qualitative analysis and discussion between grades 6 and 7

As Figure 24 presents (as well as Table 1 and 2), grade 6 received a better mean score than grade 7 in 2 out of the 3 sub-tasks. As earlier mentioned, task 1.2.1 NO and task 1.2.2 EN showed a lower mean in both grades compared to the tasks' maximum score. In both grades, *tirgul* was frequently translated to *cat,* which is noteworthy. *Tigrul* has a closer word-association to *tiger* than *katt* in Norwegian. In both grades, some of the participants translated the sentence to *Tigeren er svart og oransje* (*The tiger is black and orange*). This might indicate that the participants have understood the meaning of the sentence (a description of the tiger), resulting in them writing that the tiger is black and orange by familiarity.

Some of the participants translated other sentences than *tigrul este negru si verde.* One of them (P7-7C) translated the sentence *El gato es blanco* (*The cat is white*) to *Gatene er tomme* or, in English *The roads are empty.* The translation was given a zero-mark, but does report to a degree of metalinguistic awareness since the participant drew on word-association in i.e. the word *gato. Gato,* which is the English *cat* is similar to the Norwegian *gate* (street). The same applies to the Spanish *blanco* which, in English, means *white*. However, the Spanish word is similar to the Norwegian *blank*. The participant has translated the word to *tomme* (*empty*), which is a close synonym to the word *blank.*

43

A different participant (P15-6A) identified the correct remaining sentence in task 1.1. However, s/he translated *tigrul este negru si verde* to *katten er hvit* (The cat is white), which is the correct translation of one of the other sentences in 1.1 (El gato es blanco). The participant did not state that s/he had any form of Spanish proficiency, but that s/he had little to some proficiency in German and French. Translations like these, although given a zero mark, show a level of metalinguistic awareness as they report that the pupil understands the lexical relationship between languages.

To summarize, grade 6 received a higher mean in total in task 1. Grade 7 received a better mean score in one of the sub-tasks (1.2.2 EN), with 0,1 points.
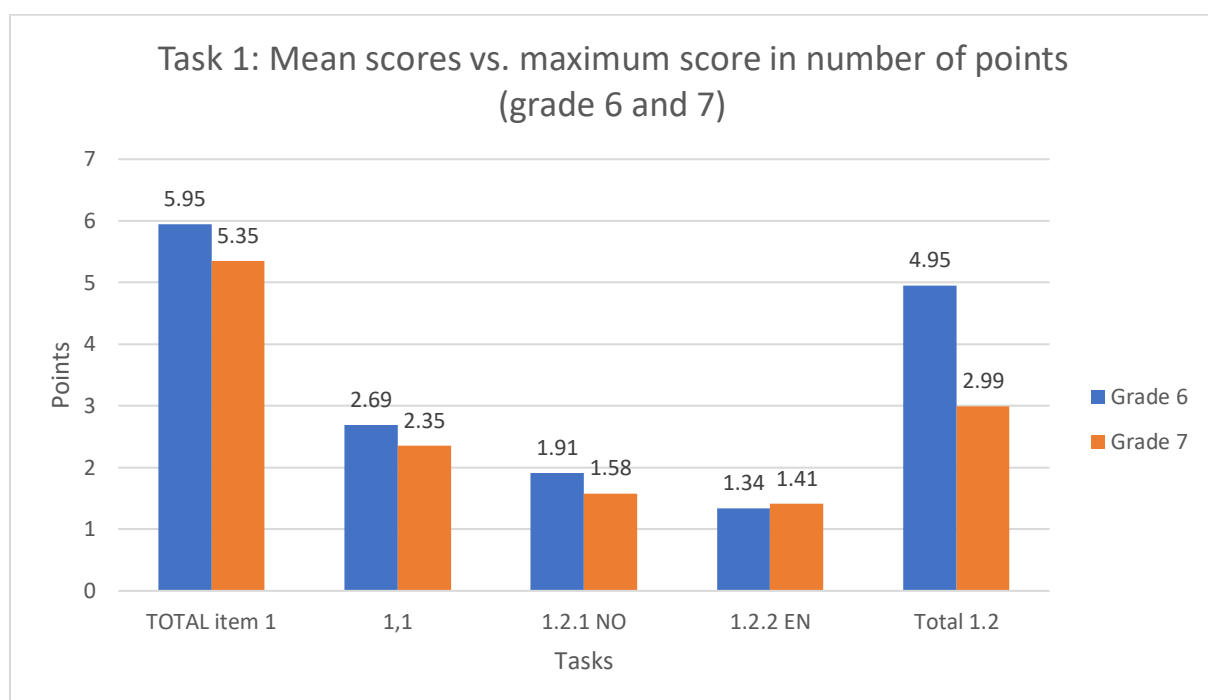
Figure 24: Mean scores vs. maximum score in number of points for task 1 in grades 6 and 7.

4.2 Task 2 Translation IT, GE, ESP

4.2.1 Quantitative analysis

Task 2 is divided in task 2.1 IT. (Italian), 2.2 GE (German), and 2.3 ESP (Esperanto) and required the participants to translate each sentence's keywords. Also, they could receive 1 point for translating one of the other words in the sentences other than the keywords.

Task 2.1 IT requires the participants to translate the sentence *Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato* (I have a cup of white/milky coffee with milk and sugar, and a chocolate cake/Jeg har en kopp med kaffe [latte] med melk og sukker, og en sjokoladekake. The sentence in task 2.2 GE is *Der Hund, die Katze und die Maus sind sehr*

gute Freunde (the dog, the cat and the mouse are very good friends/Hunden, katten og musen er veldig gode venner). Lastly, task 2.3 ESP asks the participants to translate *Estas bela tago do mi volas vanilan glaciaĵon kaj limonadon* (It's a beautiful day so I want a vanilla ice-cream, and a limonade/Det er en fin/nydelig dag så jeg vil ha en vanilje is[krem], og en limonade).

4.2.1.1 Grade 6

As shown in Table 3, there is considerable variations in means between task 2.1 IT 2.2 GE, and 2.3 ESP. The mean score in task 2.1 IT is 2,21 out of the maximum possible score of 7 points. As Table 3 presents, some of the participants received zero points whereas some received 5 points, resulting in a range of 5 points. The mean score in task 2.2 GE is 1,52 out of the maximum score of 7 points. This task has the widest range with 6 points, ranging from zero to 6 points received. Lastly, the Esperanto sentence (task 2.3) has the lowest mean score of 1,21 points out of the possible 7 points. The maximum point given was 4 and the minimum was zero, resulting in a range of 4 points. The overall mean for task 2 is 4,95 out of a possible maximum score of 21 points. The points range from zero to 15 points. The identification of keywords in tasks 1.1, 1.2 and 1.2 is explained in more detail in sections 4.2.1.1.1-4.2.1.1.3.

Task	N	Mean	Min.	Max.	Maximum possible score	Range	Median
2.1 IT	23	2,21	0	5	7	5	2
2.2 GE	23	1,52	0	6	7	6	0,5
2.3 ESP	23	1,21	0	4	7	4	0
Total item 2	23	4,95	0	15	21	15	4

Table 3: Descriptive statistics in task 2 in grade 6.
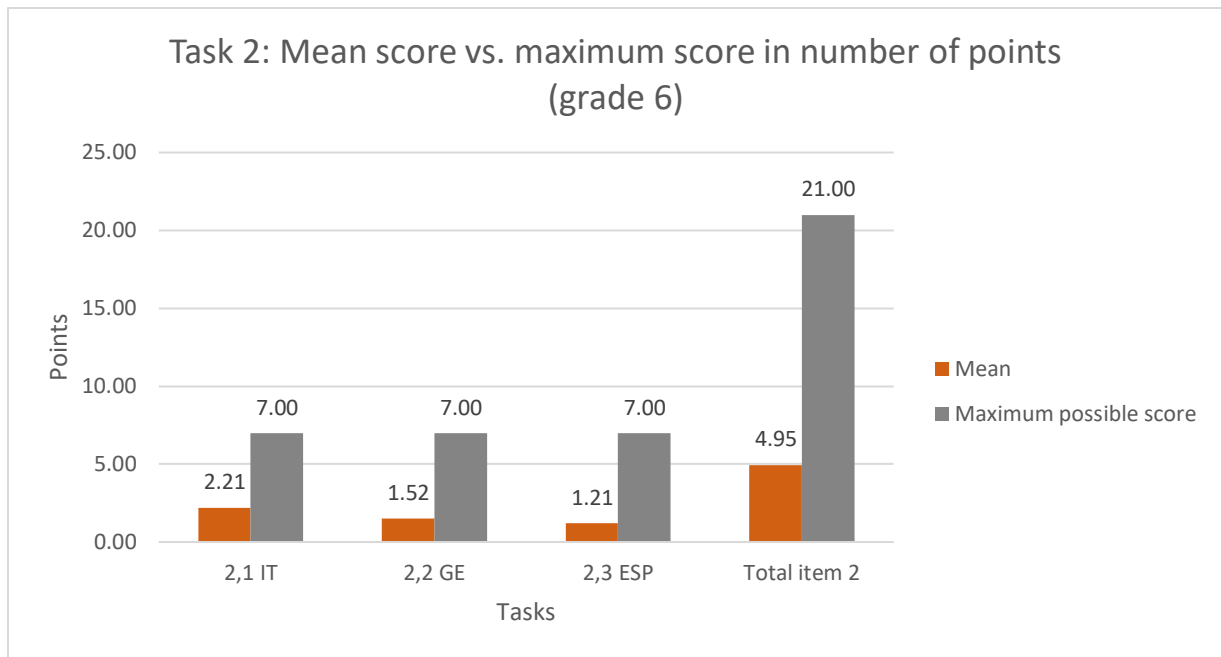
Figure 25: Mean score vs. maximum score in number of points in task 2 in grade 6.

4.2.1.1.1 Task 2.1 IT Keywords

The most common word identified was *caffè* (coffee/kaffe) which was identified by 15 out of the 23 participants. The word *latte* (latte/melk) followed closely with 11 out of the 23 participants correctly translating the word. *Cioccolato* (chocolate/sjokolade) was correctly translated by 11 out of the 23 participants, and *zucchero* (sugar/sukker) by 4 out of the 23 participants.
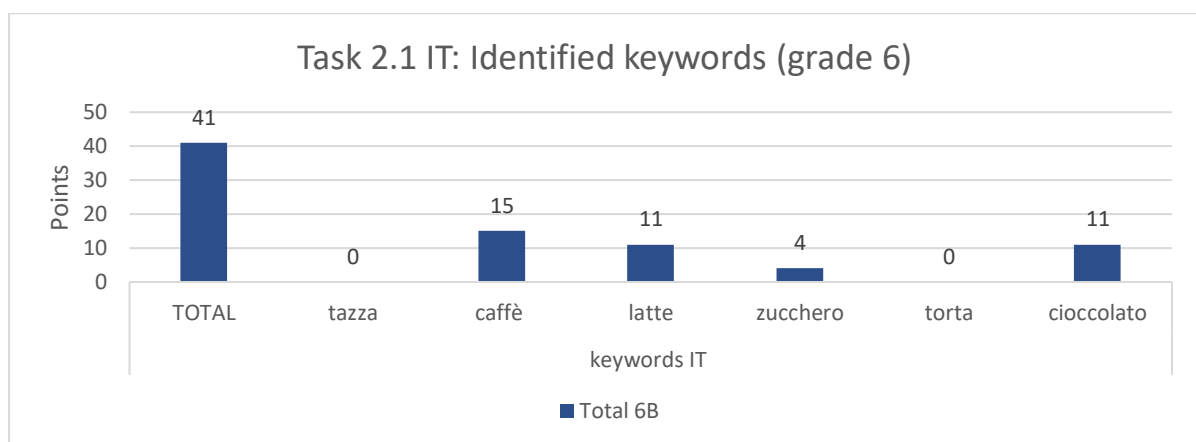
Figure 26: Number of correctly identified keywords in task 2.1 IT in grade 6.

4.2.1.1.2 Task 2.2 GE Keywords

In task 2.2 GE, the most frequent identified word was *Hund* (dog/hund), which 12 out of the 23 participants identified. The second most frequent word was *Maus* (mouse/mus) with 7 out of 23 participants correctly distinguished the word. *Katze* (cat/katt) was correctly translated by 6 out of the 23 participants. Lastly, *gute* (good/gode) and *Freunde* (friends/venner) was both identified by 1 participant.
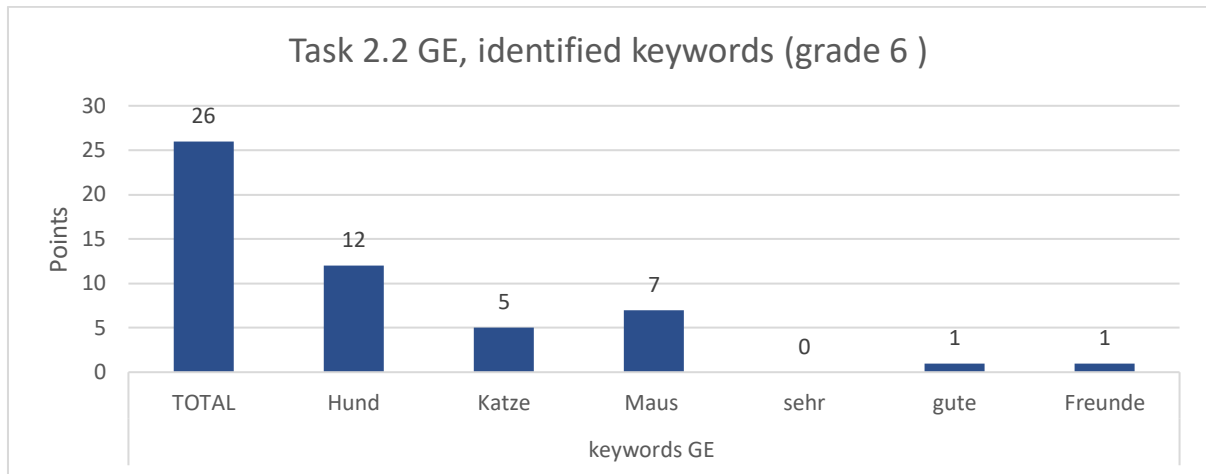
Figure 27: Number of correctly identified keywords in task 2.2 GE in grade 6.

4.2.1.1.3 Task 2.3 ESP Keywords

The most frequent word correctly identified was *vanilan* (vanilla/vanilje), which 10 of the 23 participants identified. *Limonadon* (lemonade/limonade) was correctly translated by 8 out of the 23 participants. In grade 6, no participants were able to correctly identify the words *bela* or t*ago*. 3 of the participants translated the word *glaciaĵon* (ice-cream/iskrem) correctly.
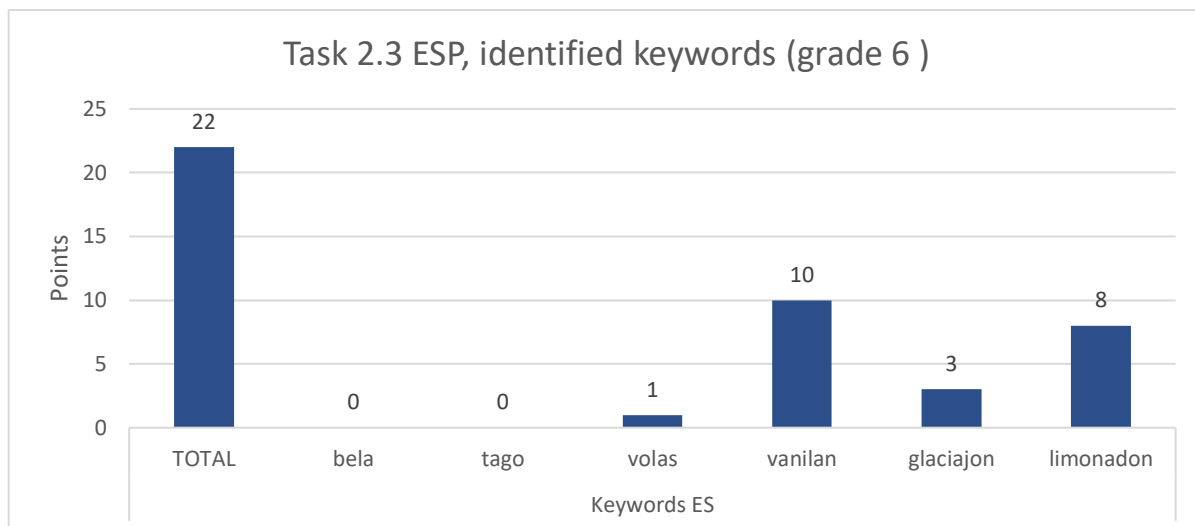
Figure 28: Number of correctly identified keywords in task 2.3 ESP in grade 6.

4.2.1.2 Grade 7

As Table 4 presents, the overall mean score in task 2 is 8,11 out of possible 21 points. The highest mean score is found in task 2.1 IT, with 3,35 out of the possible 7 points. The points range from zero to 5 points but has a median of 4 points. Task 2.2 GE has a mean score of 2,76 points out of 7 possible points. The lowest score given is zero, wheras the highest score a participant received was 6 points. However, the median is 2 points on this task, implying that a more significant part of the group received a lower score on task 1.2 GE than on task 1.1 IT. Lastly, task 2.3 ESP has the lowest mean score of 2 out of the 7 possible points with the same range as task 2.2 GE (6). Task 2.3 ESP also has the same median score as task 2.2 GE (2). The identification of keywords in tasks 2.1, 2.2 and 2.3 is explained in more detail in sections 4.2.1.2.1-4.2.1.2.3.

Task	N	Mean	Min.	Max.	Maximum possible score	Range	Median
2.1 IT	17	3,35	0	5	7	5	4
2.2 GE	17	2,76	0	6	7	6	2
2.3 ESP	17	2	0	6	7	6	2
Total item 2	17	8,11	2	13	21	11	9

Table 4: Descriptive statistics in task 2 in grade 7.
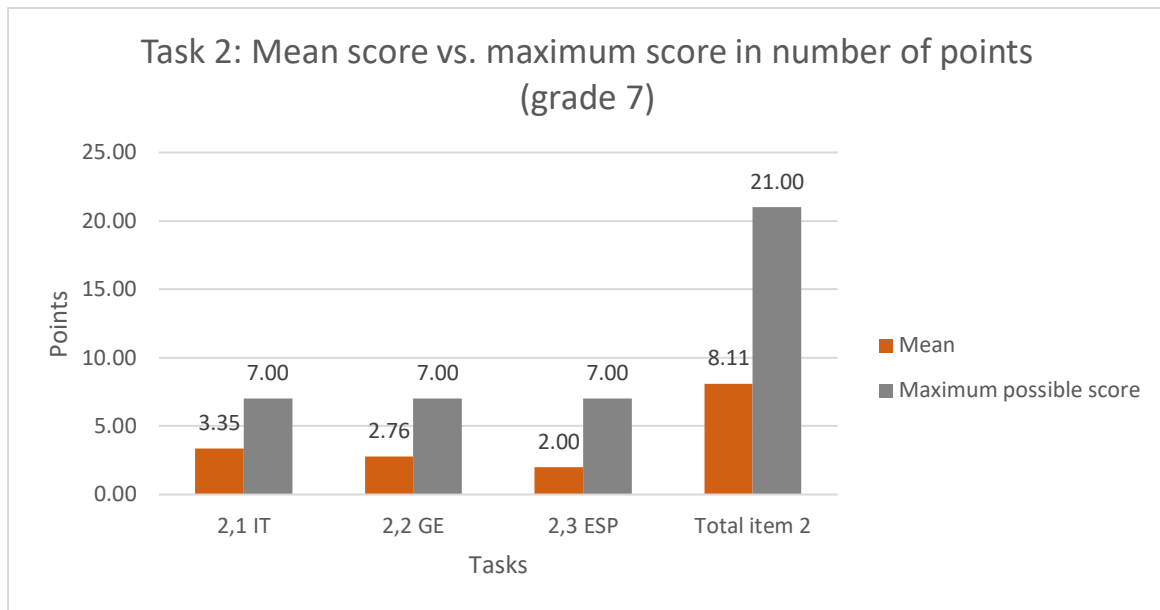
Figure 29: Mean score vs. maximum score in number of points in task 2 in grade 7.

4.2.1.2. 1 Task 2.1 IT Keywords

In 2.1 IT, the most frequent word identified was *caffè* (coffee/kaffe), which was identified by 15 out of the 17 participants. The word *latte* (latte/melk) was correctly translated by with 13 out of the 17 participants. *Cioccolato* (chocolate/sjokolade) was identified by 11 out of the 17 participants. Lastly, *zucchero* (sugar/sucker) was correctly identified by 6 out of the 17 participants.
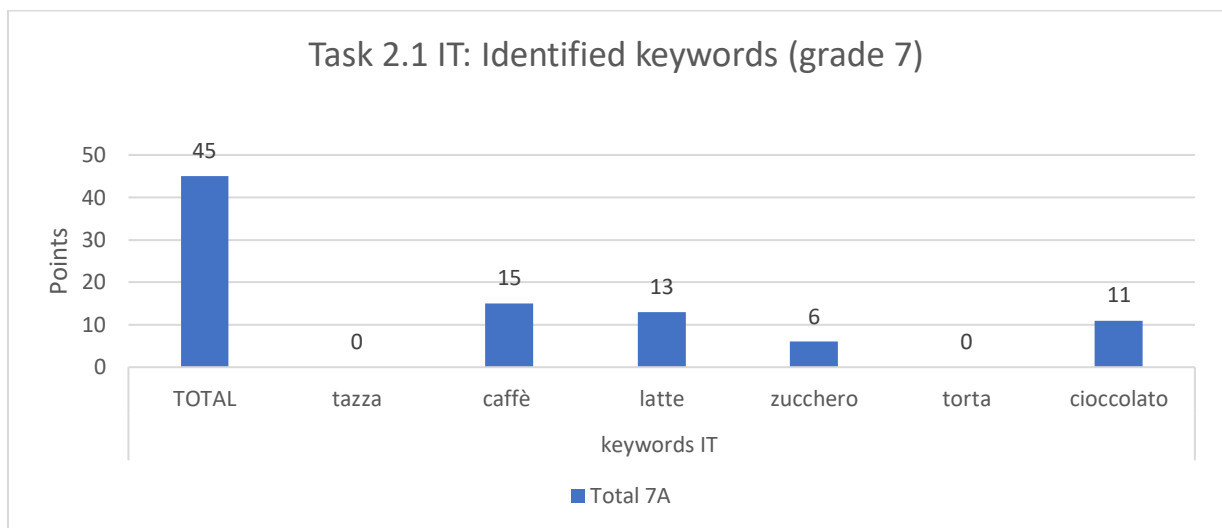
Figure 30: Number of correctly identified keywords in task 2.1 IT in grade 7.

In task 2.2 GE, the most frequent word translated was *Hund* (dog/hund), which 12 out of the 17 participants identified. The second most frequent word was *Maus* (mouse/mus) with 10 out of the 17 participants correctly translating the word. *Freunde* (friends/venner) was correctly translated by 4 participants and *gute* (good/gode) was correctly translated by 3 participants. None of the participants translated the word *sehr* (very/veldig) correctly.
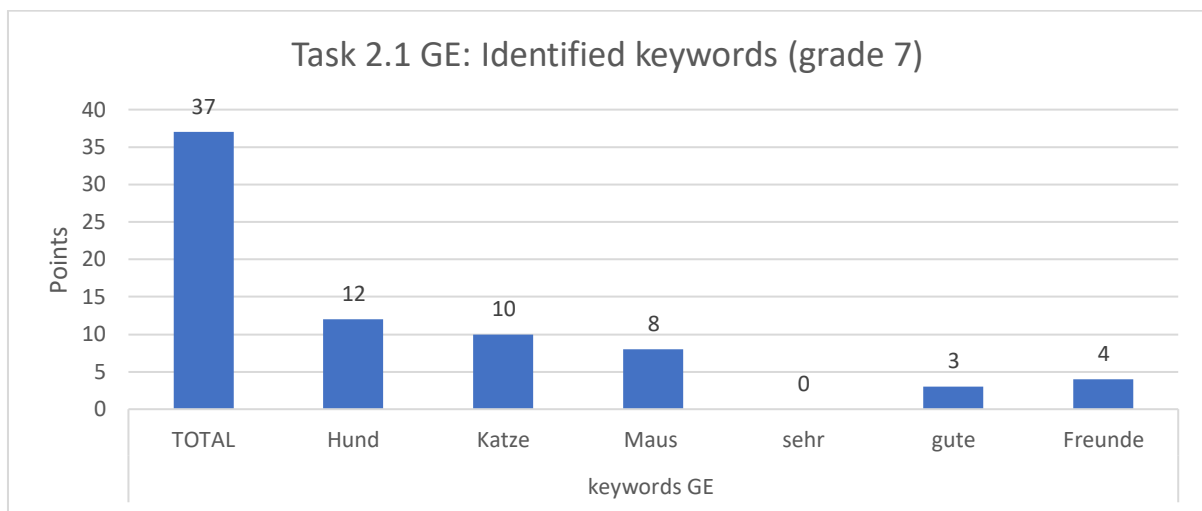
Figure 31: Number of correctly identified keywords in task 2.2 GE in grade 7.

4.2.1.2.3 Task 2.3 ESP Keywords

In task 2.3 ESP, the most frequently identified word was *limonadon* (lemonade/limonade) with 9 out of the 17 participants identifying the word. 8 of the participants correctly translated *vanilan* (vanilla/vanilje) and 6 of the participants identified the word *glaciaĵon* (ice-cream/is[krem]). 2 participants correctly idenitified the *volas* (want/vil). Lastly, 1 participant identified *bela* (beautiful/vakker).
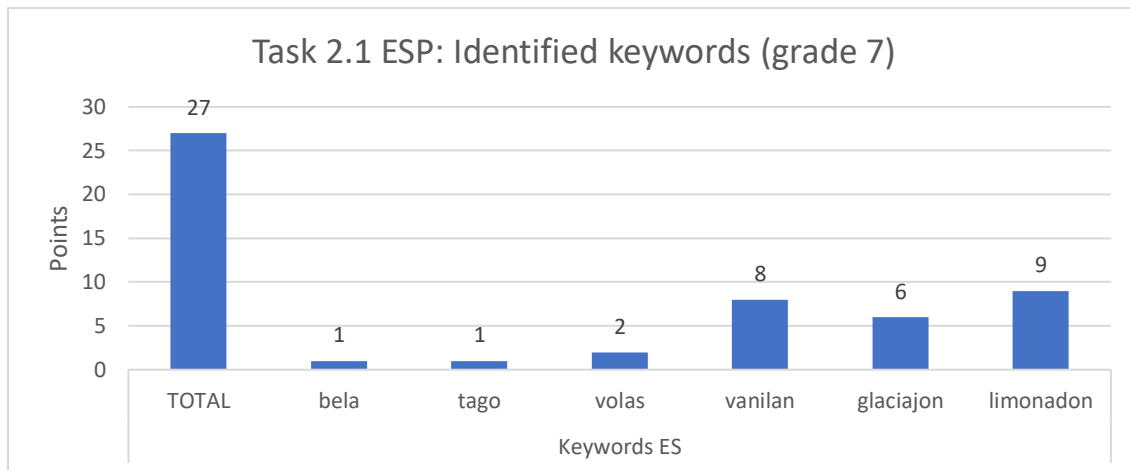
Figure 32: Number of correctly identified keywords in task 2.3 ESP in grade 7.

4.2.2 Comparative qualitative analysis and discussion between grades 6 and 7

As Figure 33 presents, the maximum possible score for task 2.1 was 21 points. There is a higher mean in grade 7 than in grade 6 in total. The mean score in grade 7 is 8,11, while grade 6 has a mean score of 4,91. Section 4.2.3.1 to 4.2.3.3 focuses on a qualitative approach of analysis, as well as a discussion of conceivable cause possibilities for the participants' answers considering metalinguistic awareness.
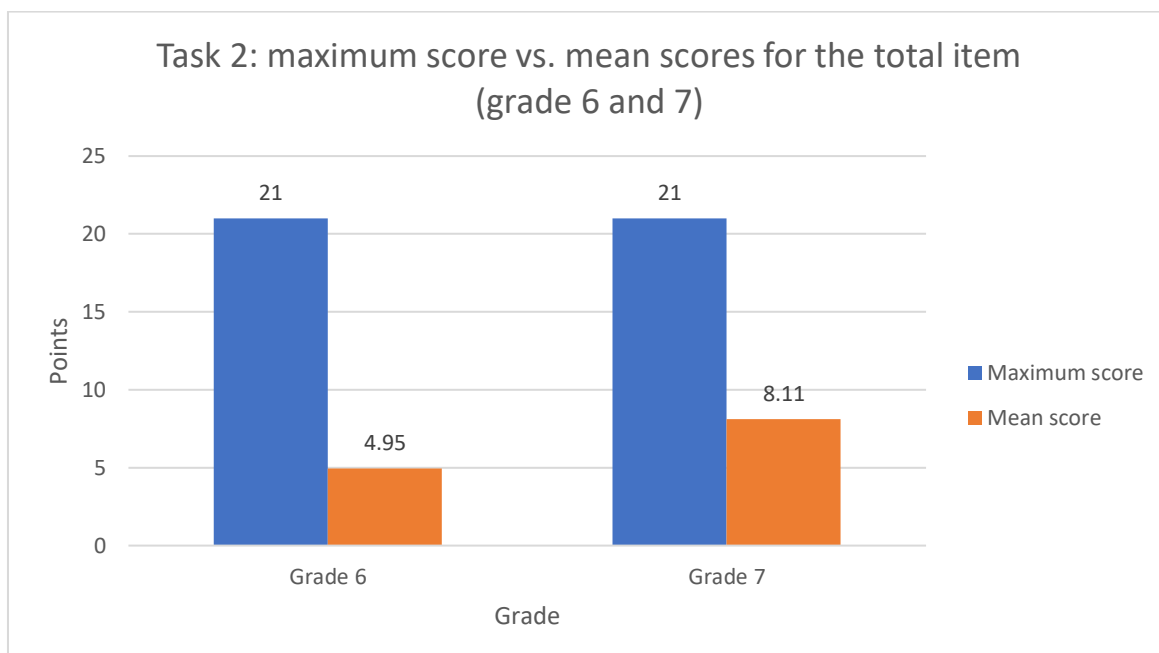
Figure 33: Maximum score vs. mean scores for task 2 in total in grades 6 and 7.

4.2.2.1 Task 2.1 IT

Task 2.1 asked the participants to translate an Italian sentence (see section 4.2.1.1). 15 (65,21%) participants in grade 6 (23) translated *caffé* correctly, along with 11(47,82%) participants that translated *latte* accurately. These are the two words most similar to the words in Norwegian. These two words were also the words that were the most translated in grade 7 (N=17). 15 (88,23%) of the participants translated *caffè* correctly along with 13 (76,47%) participants for *latte.* As presented, a higher percentage of the participants in grade 7 identified these two words than in grade 6. The same applies to *cioccolato,* which was identified by 11 (47,82%) participants in grade 6 and 11(64,70%) in grade 7. In some cases, the phrase was translated to *varm sjokolade* (hot chocolate), which is the same finding Tellier (2013) found in her study (p. 30). This did not interrupt the scoring since the translation includes *sjokolade* (cioccolato). However, in other cases, the phrase was translated to *kakao* or *varm kakao* which does not include the keyword *sjokolade* and did not qualify for a point. In other words, although those who wrote *varm sjokolade*, *kakao* and *varm kakao* presumably aimed for the same meaning, only those who wrote *varm sjokolade* receved a point. *Zucchero* was identified by 4 participants in grade 6 and 6 participants in grade 7. Zero participants identified *torta.* As Figure 34 presents, the maximum score, excluding the additional 1 point, was 6 points. The participants in grade 7 show a higher mean in all the identified keywords and has a mean score of 2,64 for task 2. The participants in grade 6 has a mean score of 1,78 points.
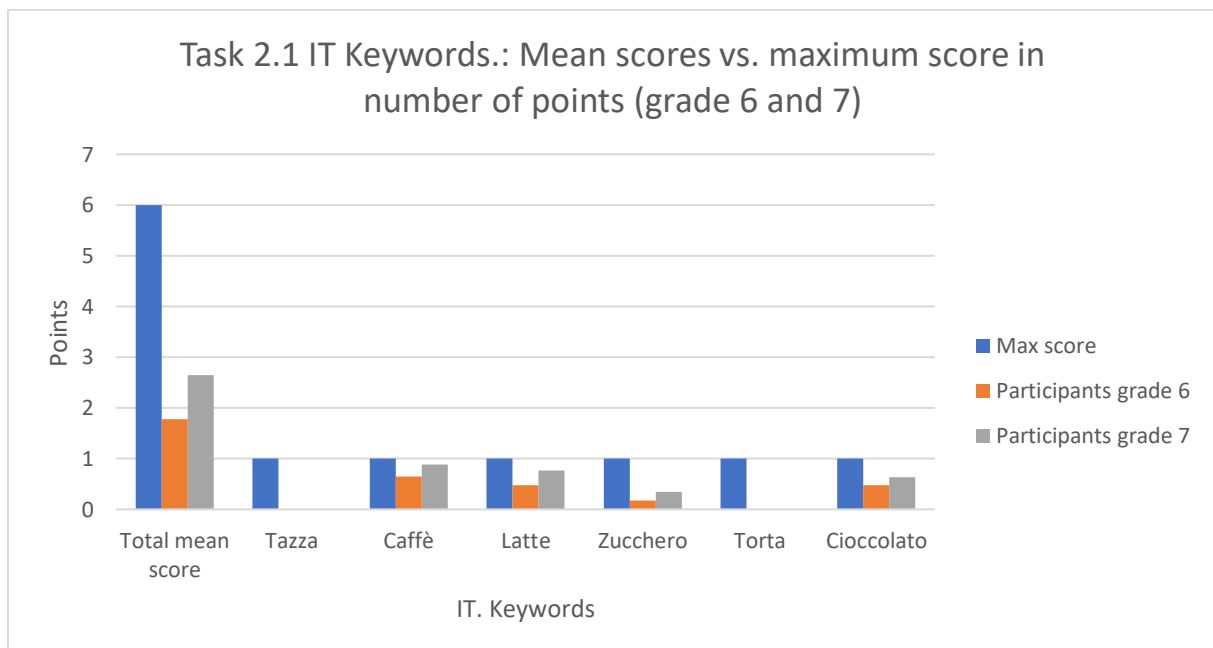
Figure 34: Mean scores and maximum score in task 2.1 IT in number of points in grades 6 and 7.

In task 2.2 GE, the participants were to translate a German sentence (see section 4.2.1.2). *Hund* (dog/hund) and *Katze* (mouse/mus) were the two most frequently translated words in both grades 6 and 7. In grade 6, 12 (52%) participants identified *Hund*, while there were 12 (70,58%) participants in grade 7 who did the same. For *Maus*, 7 (30,43%) participants in grades 6 and 8 (47,05%) in grade 7 that translated the word correctly. Both words are similar to the Norwegian spelling, which might have helped the participants in their translations. One could argue that the percentage of the translation for *Hund* should have been higher, since the word is identical in Norwegian. A different phenomenon that occurred for *Hund* was that some of the participants translated the lexical item to *Herr Hund*, as one can see in participant P5-7C: "Herr hund, og Mr. Katt og Mr. Mus er gode venner.". This may have occurred since the noun *hund* starts with a capital letter, like proper nouns do. Some of the participants also started the sentence with a greeting, like participant P12-7C: "Kjære Hund, Katt og Mus kan dere være snille.". This corresponds to Tellier's (2013) Study D, which found that some of the pupils wrote *Dear Dog*. Tellier explains this with the close similarity between the German *Der* and English *Dear*. Interestingly, the same phenomena occurred in the Norwegian translations even though the English *Dear* and Norwegian *Kjære* is quite different. Furthermore, none of the participants received a point for *sehr,* however, 1 (4,34%) (17,64%) participant in grade 6 and 3 participants in grade 7 translated *gute* correctly. Lastly, 1 (4,34%) participant in grade 6 and 4 (23,52%) participants in grade 7 translated *Freunde* correctly. As Figure 35 presents, grade 7 has an overall better mean score than grade 6. The maximum score in task 2.2 GE, excluding the additional 1 point, was 6 points. Grade 7 got a mean score of 2,17 while grade 6 got a mean score of 1,13. Grade 7 showed a higher mean in each keyword.
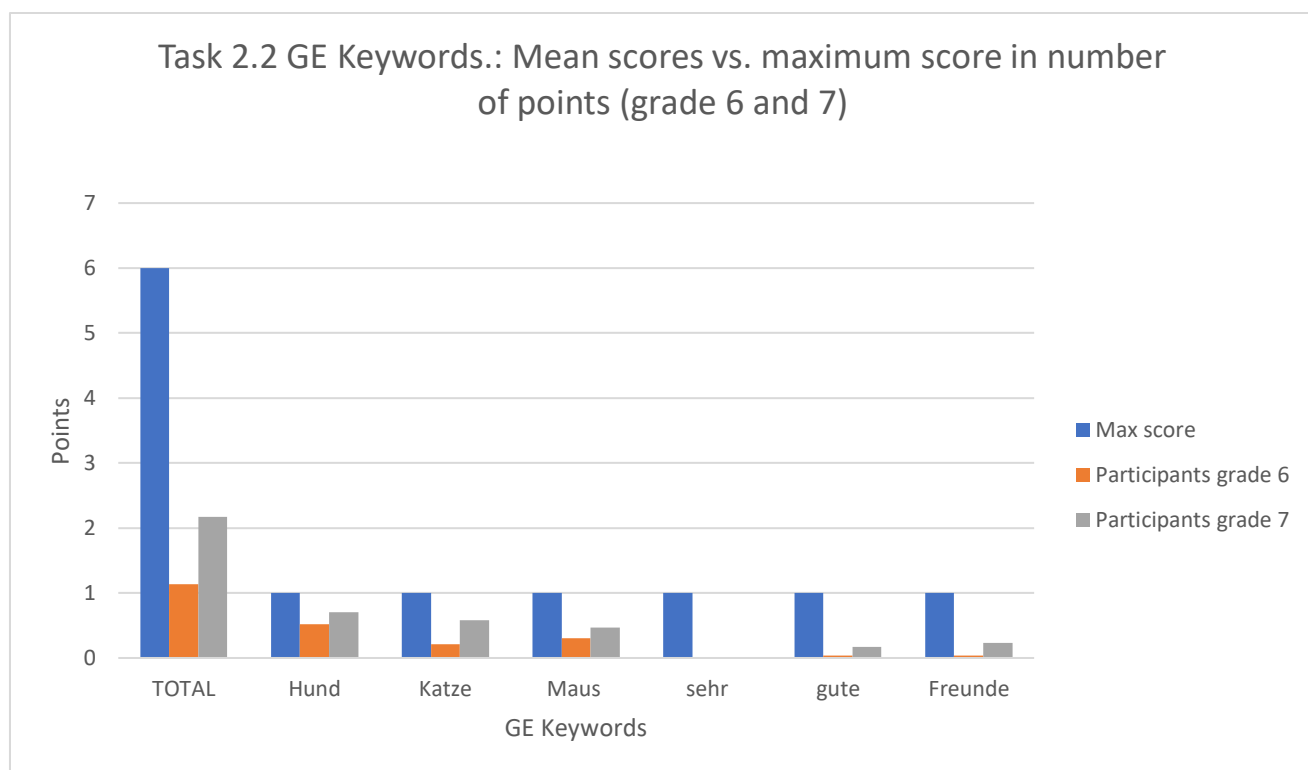
Figure 35: Mean scores and maximum score in task 2.2 GE in number of points in grades 6 and 7.

4.2.2.3 Task 2.3 ESP

Task 2.3 asked the participants to translate a sentence in Esperanto (see section 4.2.1.3). The words *vanilan* (vanilla/vanilje) and *limonadon* (lemonade/limonade) were the two most frequent correctly translated words. There were 10 (43,47%) participants in grade 6 and 8 (47,05%) participants in grade 7 who translated the word *vanilan* correctly. The correct translation of *limonadon* was conducted by 8 (34,78%) participants in grade 6 and 9 (52,94%) participants in grade 7. *Volas* were translated correctly by 1 (4,34%) participant in grade 6 and 2 (8,69%) participants in grade 7. *glaciaĵon* by 2 (8,69%) participants in grade 6 and 6 (35,29%) participants in grade 7, *bela* by zero (0%) participants in grade 6 and 1 (5,88%) participant in grade 7. Lastly, *tago* was correctly translated by zero (0%) participants in grade 6 and 1 (5,88%) participant in grade 7. One of the reoccurring phenomena was the translation of *tago* (day/dag) which was translated to *taco*, like in P7-7C: "Jeg vil ha taco til middag med en kopp limonade juice" (I want tacos for dinner with a cup of lemonade juice). These translations are presumably occurring due to the similarities in the Esperanto *tago* and the Norwegian *taco*. Another example, taken from participant P15-7C, is interesting in that the participant clearly show that the sentence deals with an ice-cream. However, s/he does not write the word "ice-cream" and is therefore not scored on the sentence: "Jeg vil ha to kuler en

vanilje og en sitron"[sic]. As figure 36 presents, grade 7 has an overall better mean score than grade 6. The maximum score in task 2.3 ESP, excluding the additional 1 point, was 6 points. Grade 7 got a mean score of 1,58, while grade 6 got a mean score of 0,95. Grade 7 showed a higher mean in each keyword.
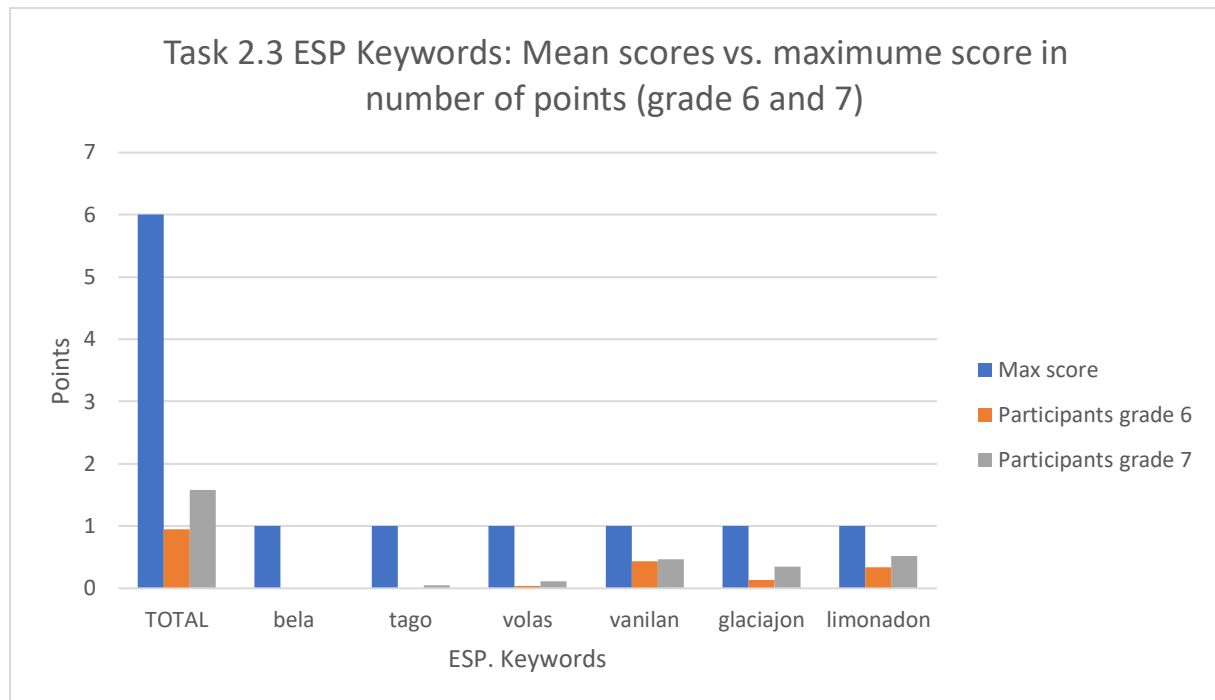
Figure 36: Mean scores and maximum score in task 2.3 ESP in number of points in grades 6 and 7.

4.3 Task 3 plural and singular nouns

4.3.1 Quantitative analysis

Task 3.1 required the participants to categorize plural and singular nouns in an artificial language. Task 3.2 required the participants to add a suffix on a noun in an artificial language to mark that the noun was in plural. As Table 5 presents, the participants could receive 8 points in task 3.1. The participants could receive 2 points if the correct suffix was added and 1 point if any suffix was added in task 3.2. The mean score and median are not applicable (NA) for task 3.2 since the answers are categorical (incorrect/correct suffix/any suffix). Therefore, the table presents the percentage of participants who got 1 or 2 points for task 3.2.

4.3.1.1 Grade 6

Table 5 reports a mean score of 5,86 out of 8 possible points for task 3.1. The answers range from zero to 8 points however, the median is higher than the mean score. The median is 7,

which means that a tail of lower scores has pulled the mean down more than the median. As Figure 37 demonstrates, 11 (27,82%) out of the 23 participants added the correct suffix to the noun and received the maximum score (2 points). 5 (21,73%) of the participants added a non-target like suffix and received 1 point for the task.

Task	N	Mean	% of participants	Min.	Max.	Maximum possible score	Range	Median
3.1	23	5,86	NA	0	8	8	8	7
3.2-correct suffix	23	NA	47,82%	0	2	2	2	NA
3.2- any suffix	23	NA	21,73%	0	1	1	1	NA
Total item 3.2	23	NA	69,56%	0	2	2	2	NA
Total item 3	23	NA	NA	0	10	10	10	NA

Table 5:Descriptive statistics of results in task 3 in grade 6.
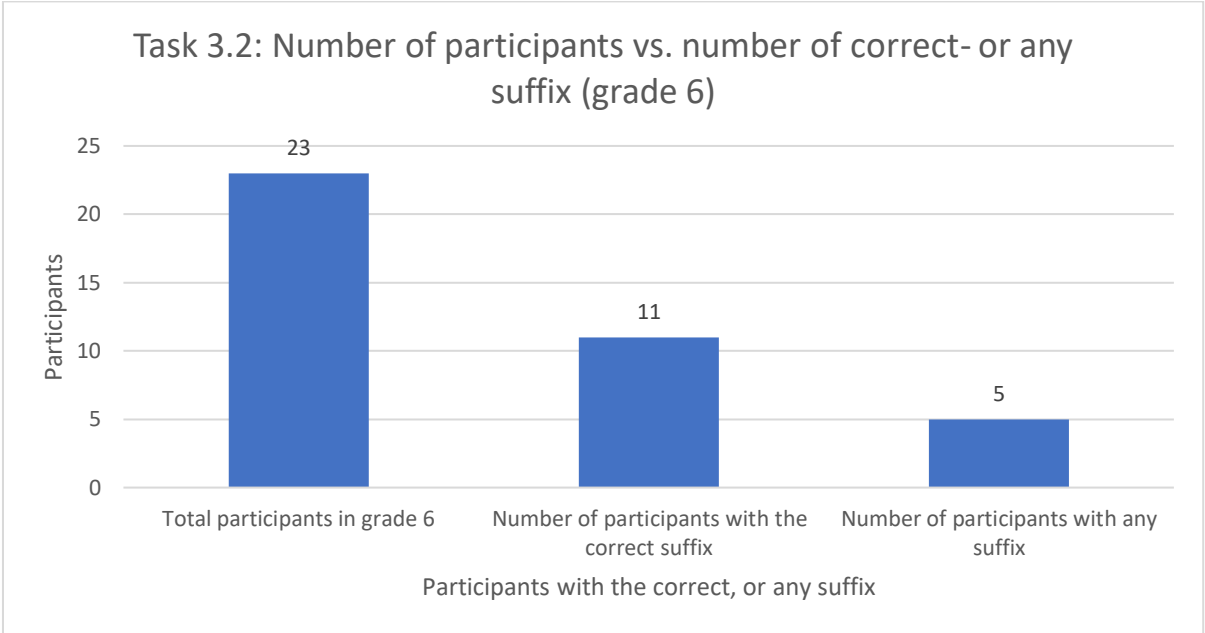
Figure 37: Number of participants vs. number of correct- or any suffix in task 3.2 in grade 6.

Table 6 reports a mean score of 7,41 out of 8 possible points for task 3.1 The answers range from 4 to 8 points, and the median for task 3.1 is 8. As Table 6, as well as Figure 38 shows, 13 (76,47%) out of the 17 participants added the correct suffix to the noun and received the maximum score (2 points). 3 (17,64%) of the participants added a non-target like suffix and received 1 point for the task.

Task	N	Mean	% of participants	Min.	Max.	Maximum possible score	Range	Median
3.1	17	7,41	NA	4	8	8	4	8
3.2- correct suffix	17	NA	76,47%	0	2	2	2	NA
3.2 – any suffix	17	NA	17,64%	0	1	1	1	NA
Total item 3.2	17	NA	94,11%	0	2	2	2	NA
Total item 3	17	NA	NA	5	10	10	5	NA

Table 6: Descriptive statistics of results in task 3 in grade 7.
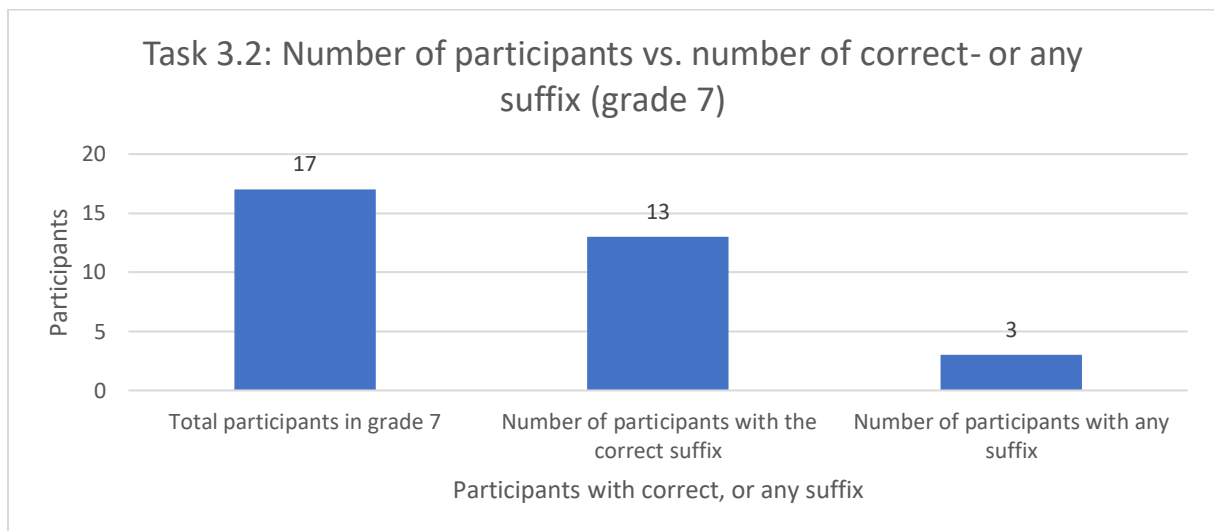
Figure 38: Number of participants vs. number of correct- or any suffix in task 3.2 in grade 7.

4.3.2 Comparative qualitative analysis and discussion between grades 6 and 7

Task 3 showed to be one of the tasks with the highest means and percentage of participants succeeding in the tasks for both grades. Grade 7 had a mean score of 7,41, and grade 6 had a mean of 5,86 points in task 3.1, pointing to a difference of 1,55 in means between the grades. 94,11% of the participants in grade 7 added a suffix to the noun in task 3.2. 69,56% of the participants in grade 6 did the same. Although both grades achieved a great score in task 3, grade 7 excelled in both task 3.1 and task 3.2. It is plausible that the artificial words assisted the participants in their answers to a considerable extent since the lexical items that were supposed to be from the same artificial language were similar to each other (*oligare-oligares*) and that this made it easier to place the singular nouns in one column and the plural nouns in the other. Those participants who analyzed the suffix in *samini* could simply apply the same pattern to task 3.2. Still, the large number of participants who added the correct suffix is noteworthy.

4.4 Task 4 Case marker
4.4.1 Quantitative analysis

Task 4 draws on the awareness of inflectional morphology, where the task requires the participants to apply a morphological rule. In this case, the participants are supposed to mark the object with an *ch*-suffix in both sub-tasks. The participants can receive a maximum of 2 points, 1 point in task 4.1 and 1 point in task 4.2.

4.4.1.1 Grade 6

As Table 7 reports, 26,08 percent of the participants added the correct suffix in task 4.1. 21,73% percent of the participants did the same in task 4.1. In both sub-tasks, the points received ranged from zero to 2 points. The average percentage of participants succeeding in task 4 was 23,91%.

Task	N	% of participants	Min.	Max.	Maximum possible score	Range
4.1	23	26,08%	0	1	1	1
4.2	23	21,73%	0	1	1	1
Total item 4	23	23,91%	0	2	2	2

Table 7: Descriptive statistics of results in task 4 in grade 6.
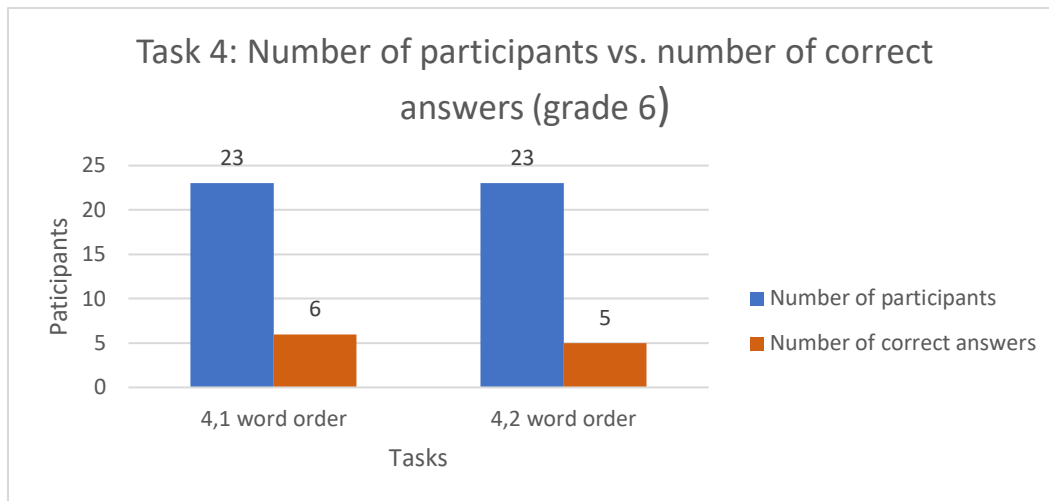
Figure 39: Number of participants vs. Number of correct answers in task 4 in grade 6.

4.4.1.2 Grade 7

As Table 8 shows, tasks 4.1 and 4.2 were answered correctly by the same number of participants. 8 (47,05%) of the 17 participants added the correct suffix in task 4.1 and task 4.2. The points ranged from zero to 2 points in both sub-tasks. The average percentage of participants succeeding in task 4 was 47,05%.

Task	N	% of participants	Min.	Max.	Maximum possible score	Range
4.1	17	47,05%	0	1	1	1
4.1	17	47,05%	0	1	1	1
Total item 4	17	47,05%	0	2	2	2

Table 8: Descriptive statistics of results in task 4 in grade 7.
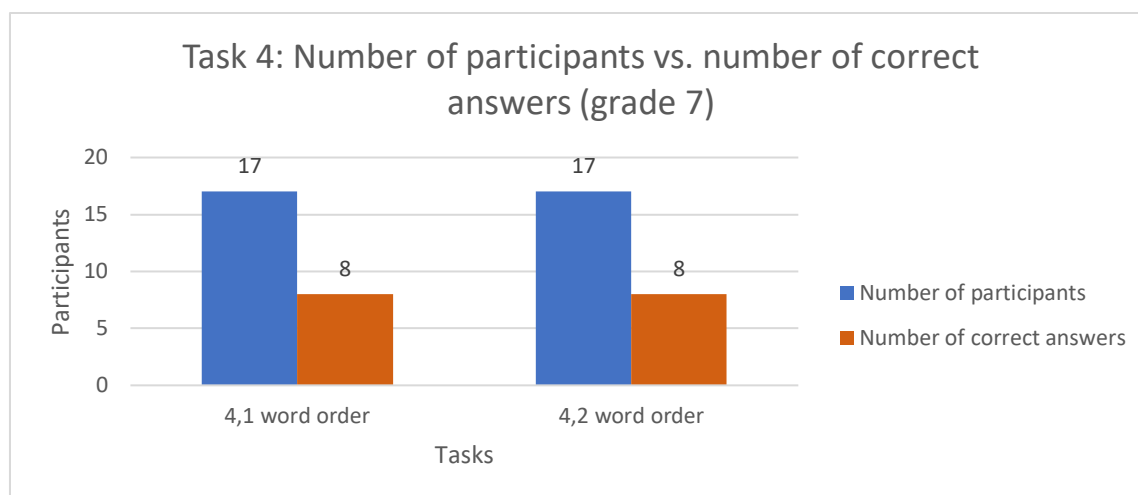
Figure 40: Number of participants vs. number of correct answers in task 4 in grade 7.

4.4.2 Comparative qualitative analysis between grades 6 and 7

In task 4, grade 7 got a considerably better score than grade 6. 23,91% of the participants in grade 6 applied the correct morphological rule, whereas 47,05% of the participants in grade 7 did the same. However, the percentages indicate that task 4 was a difficult task for both grades since less than 50% of the participants received points in both grades. This is a complex task where the participants must read an artificial language as well as produce sentences in this language themselves. One of the possible reasons for the tasks' difficulty can also be that English as a language has mostly lost its inflected case system. However, German and Turkic languages have extensive case systems. Therefore, it is interesting to take a closer look on the tests from participants who have some proficiency in Turkish or German.

One of the participants in grade 6 (P23-6C) claimed a fluent proficiency in both written and oral Turkic. S/he answered task 4.1: "omsu-ch grarid ogrtsi". The answer report that the accusative case marker has not been identified and applied as a rule in the production task. A different participant (P18-7A) claimed "good" to "some" proficiency in German and Turkic. This participant also failed to identify and produce the morphological rule in task 4.1 and 4.2. There is too little data to draw any connection between the participants' language proficiency and task results. However, this test did not find any bilingual benefits for this task.

Lastly, one participant did show a level of MLA even though s/he did not answer the tasks correctly. P7-7C handed in a piece of paper in addition to the test. The participant had reorganized the letters in the words *omsu* and *ogrtsi. Omsu* was reorganized to *mous* and the reorganized word for *ogrtsi* was not identified as it had been erased. Regardless of whether

this task answered the task, the participant shows an understanding that words can have a different meaning if they are organized in different sequences.
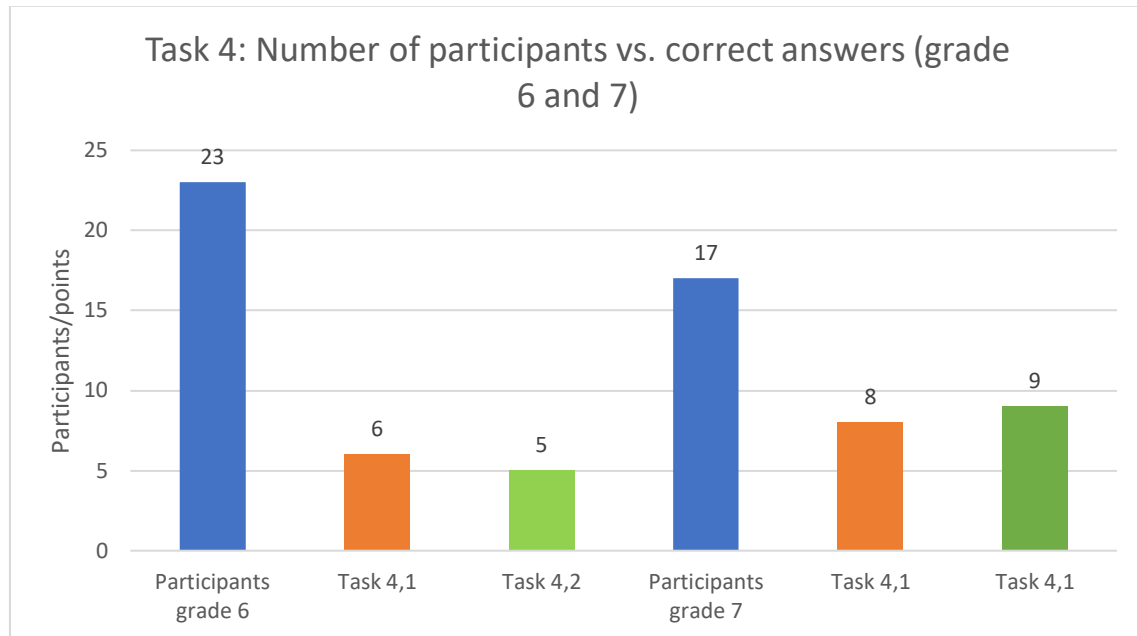
Figure 41: Number of participants vs. correct answers in task 4 in grades 6 and 7.

4.5 Task 5 Chinese numbers

4.5.1 Quantitative analysis

Task 5 is divided into 5.1 and 5.2. Task 5.1 requires the participants to translate different numbers written in Chinese to numbers. Three different Chinese words needs to be translated. In 5.2, the instruction is reversed. The participants are asked to write two different numbers in Chinese. The participants can receive 1 point per sub-task. They receive 1 point for the correct answer and zero points for incorrect answers.

4.5.1.1 Grade 6

As Table 9 presents, task 5.1.1 has the highest percentage of participants who received a point with 52,17%. Meaning that 12 out of the 23 participants received 1 point (see figure 42). 7 (30,43%) of the participants received 1 point in task 5.1.2 and 25 (21,73%) participants received 1 point in task 5.1.3. Meaning that the average percentage of participants who received 1 point in the sub-tasks of task 5.1 was 34,78%. Tasks 5.1.1-5.1.3 ranged in points given from zero to the 1 per task. There was a considerable difference in percentage for the two tasks were the participants were supposed to write numbers in Chinese. 10 (43,47%) of

the participants answered 5.2.1 correctly, whereas 4 (17,39%) of the participants received 1 point for task 5.2.2. The given points ranged from zero to 1 point per task. The average percentage of correct answers in task 5 is 30,4%.

Task	N	% of participants	Min.	Max.	Maximum possible score	Range
5.1.1	23	52,17%	0	1	1	1
5.1.2	23	30,43%	0	1	1	1
5.1.3	23	21,73%	0	1	1	1
Total 5.1	23	34,78%	0	3	3	3
5.2.1	23	43,47%	0	1	1	1
5.2.2	23	17,39%	0	1	1	1
Total 5.2	23	30,43%	0	2	2	2
Total item 5	23	30,4%	0	5	5	5

Table 9: Descriptive statistics for results in task 5 in grade 6.
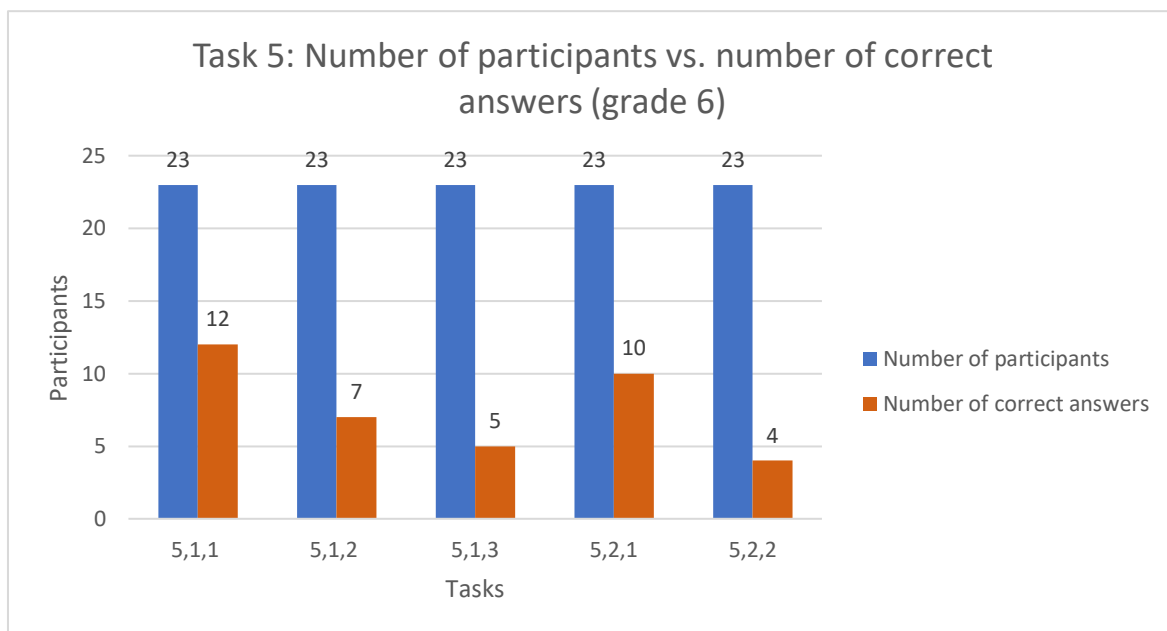
Figure 42: Number of participants vs. number of correct answers in task 5 in grade 6.

4.5.1.2 Grade 7

As Figure 43 presents, 16 (94,11%) out of the 17 participants in grade 7 received 1 point in task 5.1.1. Furthermore, 10 (58,82%) participants in task 5.1.2 and 4 (23,52)% participants in task 5.1.3 achieved the correct answers. This results in an average percentage of participants

who received 1 point in tasks 5.1.1-5.1.3 to 58,82%. As Table 10 presents, all tasks in item 5 ranged from zero to 1 point received. In task 5.2.1, 10 (58,82%) of the participants received 1 point, and 9 (55,88%) participants received 1 point in 5.2.2. This leads to an average percentage of participants who received 1 point in task 5.2 to 55,88%. The average percentage of correct answers in task 5 is 57,64%.

| Task	N	% of participants	Min.	Max.	Maximum possible score	Range
5.1.1	17	94,11%	0	1	1	1
5.1.2	17	58,82%	0	1	1	1
5.1.3	17	23,52%	0	1	1	1
Total 5.1	17	58,82%	0	3	3	3
5.2.1	17	58,82%	0	1	1	1
5.2.2	17	52,94%	0	1	1	1
Total 5.2	17	55,88%	0	2	2	2
Total item 5	17	57,64%	0	5	5	5

Table 10: Descriptive statistics in task 5 in grade 7.

Figure 43: Number of participants vs. number of correct answers in task 5 in grade 7.

4.5.2 Comparative qualitative analysis and discussion between grades 6 and 7

Task 5.1.1 was the task that most participants answered correctly in both grades 6 and 7. Though, grade 7 did noticeably better with a percentage of 94,11% against grade 6 with 52,17%. Grade 7 performed better than grade 6 in all sub-tasks. The average percentage in task 5.1 for grade 6 was 34,78% while grade 7 completed the task with 58,82% of the participants with correct answers. For 5.2, the average in grade 6 was 30,43%, whereas the average in grade 7 was 55,88%. Finally, the average percentage of successful answers in grade 6 was 30,4%, while it was 57,64% in grade 7. As Tables 9 and 10 demonstrates, task 5.1.3 appeared to be one of the most difficult tasks. The correct answer for task 5.1.3 *sanshisan* was 33. One of the reoccurring answers in 5.1.3 was the number 313. With san meaning 3 and *shi* meaning 10, there is logical reasoning behind these answers since the three-word combination might have sparked the belief of there being a three-digit answer. Task 5.2.2 was the task that the least of grade 6 answered correctly. Similarly, it was the task that grade 7 got the second-least correct answers in. This task asked the participants to write the number *24* in Chinese. The correct answer for the task was *ershisi.* The study will use participant P22-6A to elicit one of the most occurring answers in task 5.2.2. S/he answered *erssi.* Here, we can try to trace the logical process of the answer. The Chinese *er* is the number *2*, while the Chinese *si* is the number *4*. It is plausible that the participant(s) has added the numbers *2* and *4* to reach number *24* instead of looking at the word for *20*, which is *ershi.* This example shows that several participants were close to the correct answer and were close in the thought process when they answered the task. Despite that, these participants cannot receive 1 point since this is only assumptions made by the researcher and the answers were incorrect.
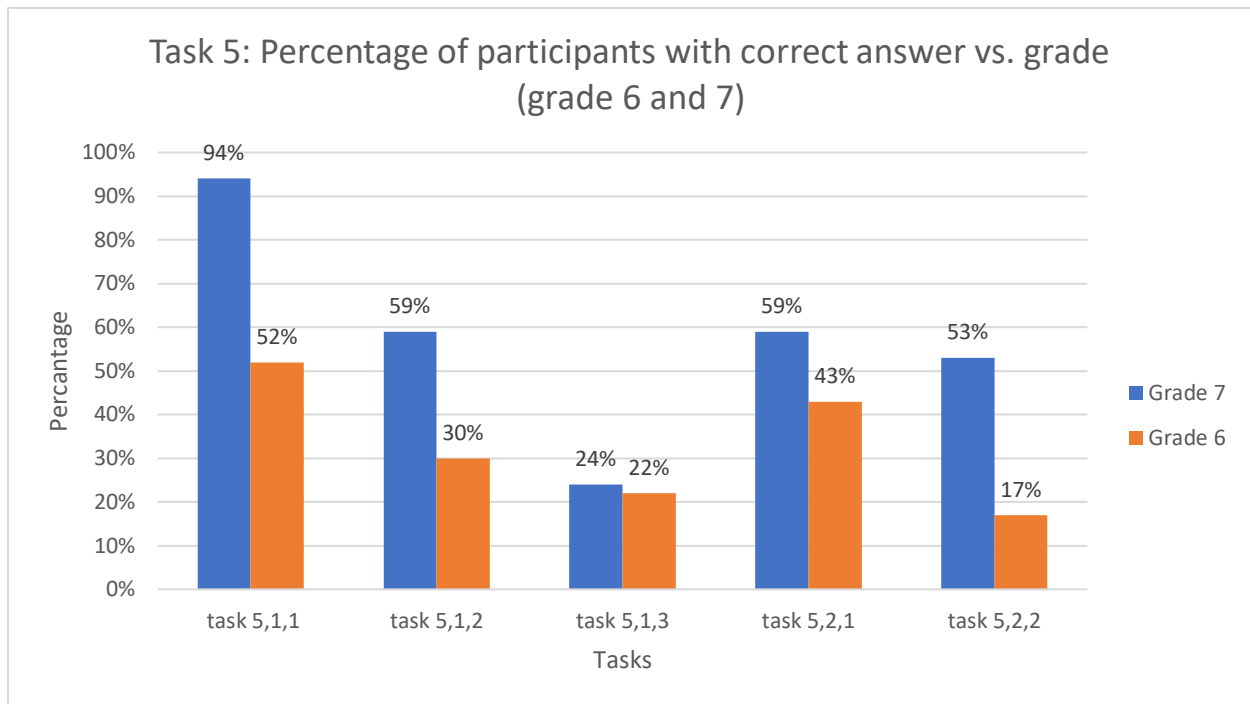
Figure 44: Percentage of participants with correct answers vs. grade in task 5.

4.6 Task 6 Translation and word order in *Krarid*

4.6.1 Quantitative analysis

Task 6 is divided into task 6.1 (translation) and task 6.2 (word order). The participants are asked to translate a sentence to the artificial language *Krarid*. They were scored on how many words they were able to translate in task 6.1 In task 6.2, they received 1 point if the word order corresponded with the word order in *Krarid*. As Table 11 and 12 presents, the results in task 6.1 is presented with mean and median scores. Task 6.2 is presented with the percentage of participants who constructed the correct word order and therefore received 1 point.

4.6.1.1 Grade 6

Task 6.1 had a maximum possible score of 3 points. As Table 11 presents, the task range was 3 points, meaning that some of the participants received 1 point for the respective task while some received 3 points. The median was 3, meaning that 3 points are the center of the distribution. In task 6.2, 9 out of the 23 participants received 1 point. Which, as we can see in Table 11, 39,13% of the participants answered task 6.2 correctly.

Task	N	Mean	% of participants	Min.	Max.	Maximum possible score	Range	Median
6.1	23	2	NA	0	3	3	3	3
6.2	23	NA	39,13%	0	1	1	1	NA
Total item 6	23	NA	NA	0	4	4	4	NA

Table 11: Descriptive statistics in task 6 in grade 6.
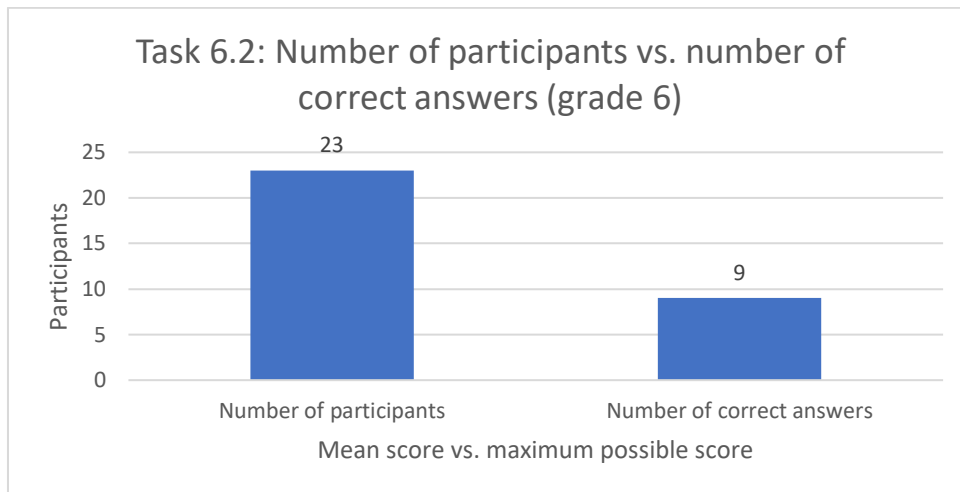
Figure 45: Number of participants vs. number of correct answers in task 6.2 in grade 6.

4.6.1.2 Grade 7

With a maximum possible score of 3 points, grade 7 had a mean score of 2,88 points in task 6.1 (Table 12). The median for the respective task was 3. The lowest score in task 6.1 was 2 points, while the highest score was 3 points. These results show a range of 1 point in 6.1. In task 6.2, 7 out of the 17 participants had the correct word order. Accordingly, 41,11% out of the participants received 1 point in task 62, as Table 12 demonstrates.

Task	N	Mean	% of participants	Min.	Max.	Maximum possible score	Range	Median
6.1	17	2,88	NA	2	3	3	1	3
6.2	17	NA	41,11%	0	1	1	1	NA
Total item 6	17	NA	NA	2	4	4	2	NA

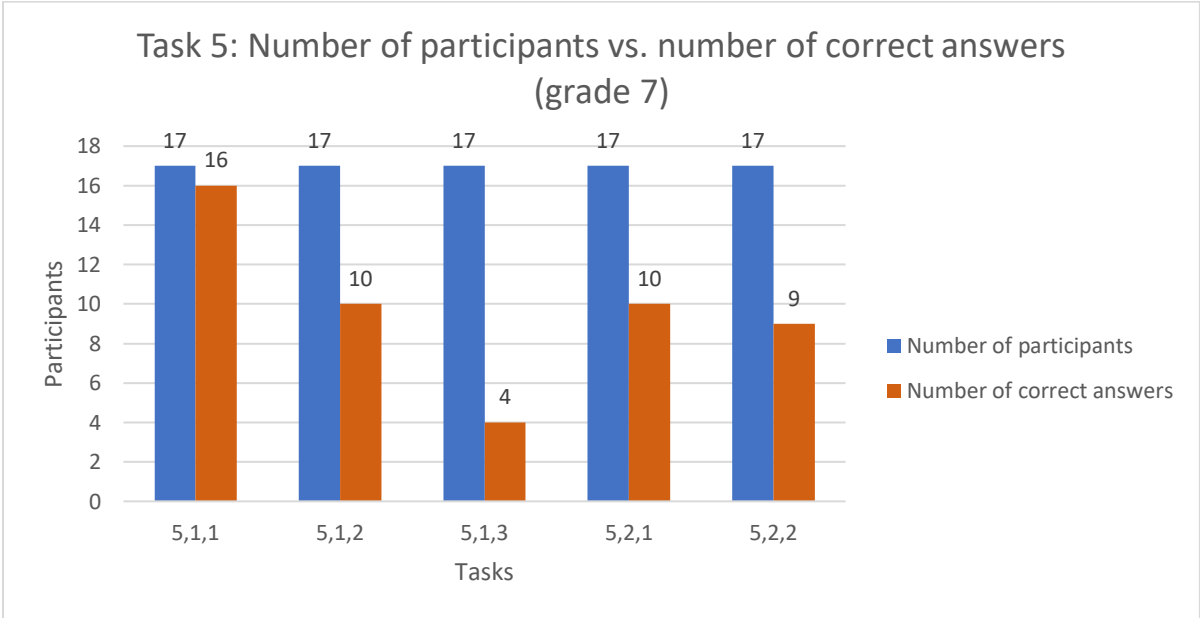*Table 12: Descriptive statistics in task 6 for grade 6.*
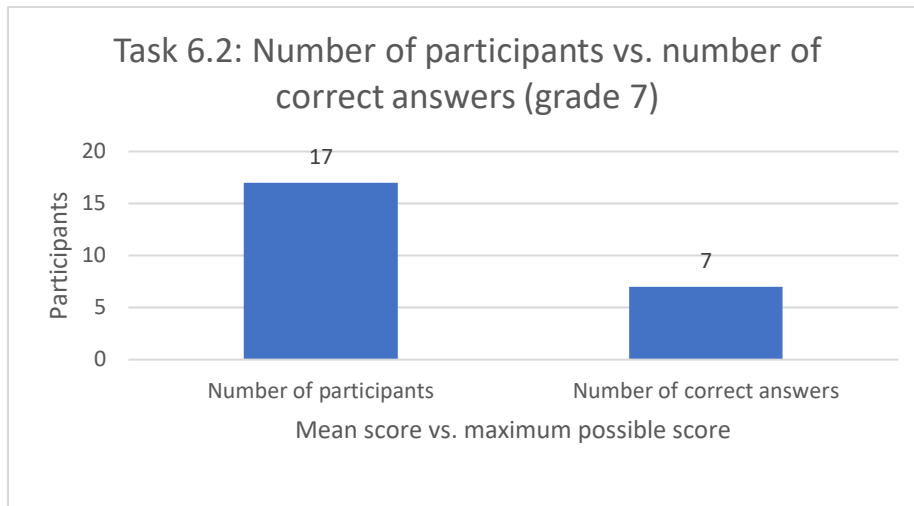
Figure 46: Number of participants vs. number of correct answers in task 6.2 in grade 7.

4.6.2 Comparative qualitative analysis and discussion between grades 6 and 7

As Table 13 presents, the median score in task 6.1 is 3 in both grades 6 and 7. However, the range in points is higher in grade 6 with 3 points against grade 7 with 1 point in the range. As Table 13 presents, the mean score is higher in grade 7 with 2,88 points against grade 6 with a mean score of 2 points. In task 6.2, the results in grades 6 and 7 are more homogenous. There were 39,13% of the participants in grade 6 who got the correct answer while there were 41,11% in grade 7. Resulting in a difference of only 1,95%. There was a noteworthy tendency in task 6.2 for the participants who received zero points. The participants that translated the words correctly reorganized the words to suit the Norwegian SVO-structure instead of the *Krarid* SOV-structure. These answers indicate that the participants have transferred their knowledge from their source language as a resource to facilitate the target language. As stated by Zhang, Zhang and Koda (2014 cited in Zhang and Chin, 2017, p. 6, p. 23), transfer facilitation from the native language can occur in a foreign language context and show a level of phonological awareness.

Grade	Task	N	Mean	% of participants	Min.	Max.	Maximum possible score	Range	Median
6	6.1	23	2	NA	0	3	3	3	3
7	6.1	17	2,88	NA	2	3	3	1	3
6	6.2	23	NA	39,13%	0	1	1	1	NA
7	6.2	17	NA	41,11%	0	1	1	1	NA

Table 13: Descriptive statistics comparing grade 6 and 7 in task 6 (grade 6 highlighted in blue).
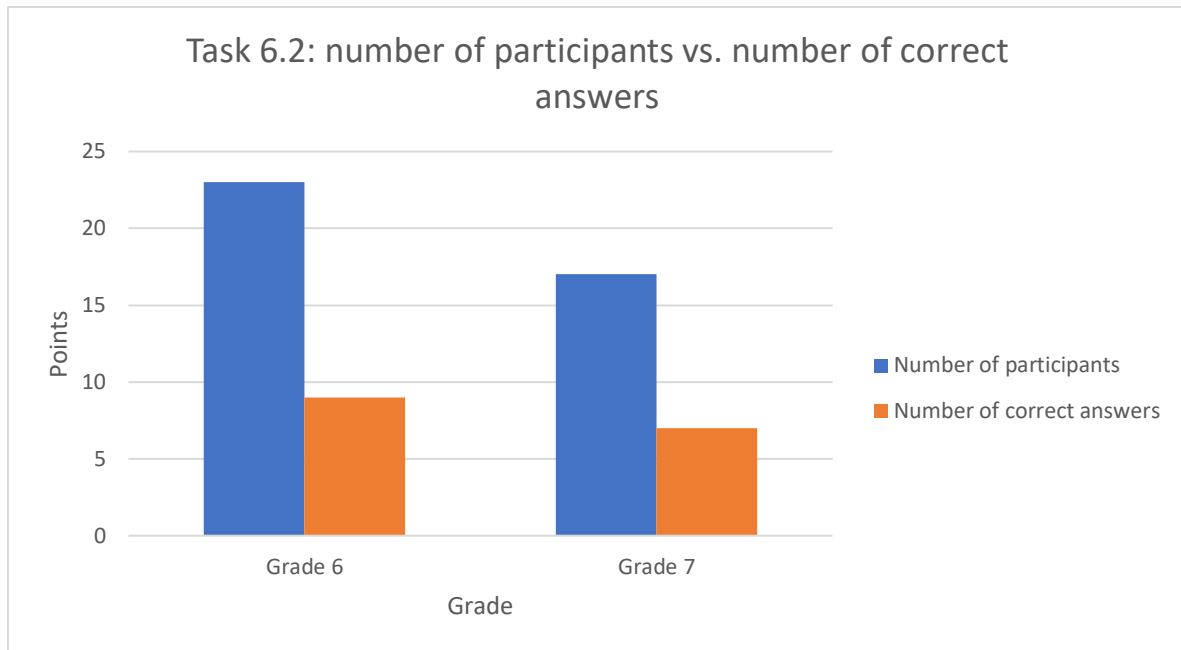
Figure 47: Number of participants vs. number of correct answers in task 6.2 in grades 6 and 7.

4.7 Overall MLA score in grade 6 and grade 7

4.7.1 Quantitative analysis

As table 14 presents, the mean score for the MLA-test in grade 6 is 22,47 points out of the maximum possible score of 57 points. In grade 7, the mean score is 29,70 which concludes a difference in means of 7,23 points. It must be noted that the mean score in grade 6 may be skewed due to two participants scoring well under the average (3 and 5 points). Which also contributed to a wide range. The range in grade 6 was 45 points, with the lowest total score of 3 points and the highest total score of 48 points. In grade 7, the range was 30 points. The lowest score was 12 points while the highest score was 42 points. There is a difference of 5 points between the participant who scored the lowest and the participants who scored the second lowest. However, the scores are substantially higher than in grade 6. Accordingly, no evident outliers were found in grade 7 as it was in grade 6. As Table 14 reports, the median was lower in grade 6 than in grade 7. The median in grade 6 was 21 points while the median in grade 7 was 30 points. It is evident that the results are more homogenous in grade 7 than in grade 6 when referring to the minimum and maximum score as well as the range in the two grades.

Grade	Task	N	Mean	Min.	Max.	Maximum possible score	Range	Median
6	Total MLA score	23	22,47	3	48	57	45	23
7	Total MLA score	17	29,70	12	42	57	30	30,5

Table 14: Descriptive statistics of overall MLA score in grades 6 and 7.

4.7.2. Qualitative analysis and discussion

The results show that grade 7 received an overall better mean score in MLA than grade 6 did. However, the participant that received the highest MLA score was a participant in grade 6 with 48 out of 57 possible points. The comparison of MLA between grades 6 and 7 will be discussed in section 5.1 since one of the aims of this study was to investigate whether there was a clear difference in means between grades 6 and 7. Though, one of the occurring tendencies that might have affected the results will be pointed out in this section. There was a more frequent tendency of participants in grade 6 to either erase or scratch their answers than in grade 7. This might indicate that a more significant proportion of the participants in grade 6 gave up instead of retrying the task. However, the study cannot make a claim about numbers of attempts since there is no collected data revealing this. Regardless, the collected data endorse the claim that there were more tasks in grade 6 than in grade 7 that was left unanswered. This may point to a roof-effect for some of the participants in grade 6. It should be noted that it occurred in grade 7 as well, even though not as frequent.

4. 8 Teacher assessment compared to MLA score

One of the aims of this study was to investigate whether there was a correlation between the participants' MLA and their L2 proficiency. Table 15 and 16 presents the teacher's assessment of the participants' academic achievement in the English course for grades 6 and 7. The teachers were given the same instructions and were asked to assess the participants from a scale from 1 to 6. The tables are structured in ascending order from the lowest to the highest assessment.

As demonstrated by Table 15, there is little correspondence between the teacher assessment and their given MLA score. Those participants that were given 1 in assessment vary in scores from 30 to 48 points in MLA. One of the participants who were given 1 in L2 proficiency was P22-6A who received the highest MLA score in grade 6 with 48 points. The participants who received 2 in assessment had a much lower MLA score than those who received 1. There were six participants that received 2 in teacher assessment. The MLA scores varied from 5 to 39 points, but three out of the six participants received an MLA score of fewer than 10 points. There was one participant given a 6 in teacher assessment. As presented by Table 15, this participant (P1-6A) received a total MLA score of 17 points. In comparison, P22-6A was assessed with 1 and had a total MLA score of 48 points. As presented in section 2.3.1, previous studies have pointed to a positive correlation between MLA and L2 competency (Simard & Gutièrrez, 2018, pp. 210). A discussion of the relationship between the teacher assessment and MLA score will, therefore, be addressed in section 5.2.

MLA score compared to teacher assessment, grade 6

Teacher assessment	MLA score	Participant
1	30	P14-6A
1	35	P18-6A
1	48	P22-6A
2	5	P2-6A
2	7	P4-6A
2	17	P11-6A
2	9	P13-6A
2	24	P16-6A
2	39	P17-6A
3	3	P5-6A
3	15	P6-6A
3	22	P7-6A
3	26	P9-6A
3	24	P25-6A
3	29	P10-6A
4	8	P20-6A
4	14	P23-6A
4	20	P3-6A
5	38	P8-6A

5	16	P15-6A
5	30	P19-6A
5	41	P21-6A
6	17	P1-6A

Table 15: MLA score compared to teacher-reported English competence in grade 6.

4.8.2 Grade 7

There were no participants who were given an assessment of *1* in grade 7 (Table 16). Two of the 17 participants were given *2* in assessment and received MLA scores of 12 (P1-7C) and 28 points (P4-7C). The participants that were given a *5* in English competence ranged in MLA scores from 16 to 41 points. The participant with the highest MLA score (42 points) was P17-7C who received a *3* in the teacher assessment. Though, the participant that received the lowest MLA score (12 points) was P1-7Cm who was assessed with the lowest number on the scale utilized by the teacher, namely *2*. Still, there is little indications between a relationship with the participants assessments and MLA scores, which will be discussed in section 5.2.

MLA score compared to teacher assessment, grade 7

Teacher assessment	MLA score	Participant
2	12	P1-7C
2	28	P4-7C
3	29	P10-7C
3	42	P17-7C
4	31	P6-7C
4	32	P11-7C
4	39	P12-7C
4	39	P13-8C
4	40	P14-7C
4	30	P15-7C
4	17	P16-7C
5	31	P5-7C
5	32	P8-7C
5	21	P18-7C

6	27	P7-7C
6	25	P9-7C

Table 16: MLA score compared to teacher-reported English competence in grade 7.

5.0 General discussion and conclusion

The aim for this study was to investigate the state of MLA of one 6th- and one 7th grade in a school in Bergen. Furthermore, one of the study's purposes was to examine whether there was a relationship between the participants' MLA and teacher-reported English competence. Lastly, it aimed to see if there was a clear difference in mean scores between grades 6 and 7 and to compare these findings with Tellier's (2013) results. Thus, the study aimed to shed light on the following questions:

1. What is the state of metalinguistic awareness in grade 6 in a primary school in Bergen?
2. What is the state of metalinguistic awareness in grade 7 in a primary school in Bergen?
3. Is there a correlation between the participants' MLA and their teacher-reported English competence?
4. Is there a clear difference in MLA scores between grade 6 and 7?

5.1 The state of MLA in grades 6 and 7

As shown in Table 15 and 16, the results differed between 3 to 48 points in grade 6 and 12 and 42 points in grade 7, which indicates that the tasks differentiated well and that there were no roof-effect in the test. Grade 6 had a range of 40 points in the test-scores, which points to a clear difference in MLA amongst the group of participants. Five out of the 23 participants received an MLA score beneath 10 points out of the possible 57 points. The mean score of 22,47 points is 7,23 points below grade 7 with a mean score of 29,70 points. The range of test scores in grade 6 is also higher than in grade 7. There were no participants in grade 7 that received a score lower than 10 points. The lowest score in grade 7 was 12 points. This points to an upward shift in the minimum and maximum scores between the two grades. The results demonstrated in table 7.1 indicates a clear improvement in means between grades 6 and 7. This improvement could also be statistically significant and should be tested in the future. However, as stated in section 2.3, an improvement in MLA between grade 6 and 7 is expected as children develop their MLA with age (Tellier, 2013, p. 36). It would have been beneficial to investigate whether the difference in means is more significant between grades 6 and 7 than

it would have been between other grades (i.e., grades 5 and 6). A suggestion to further research would have been to conduct the same test on different age groups to be able to compare the differences in means. However, the study can conclude with a clear difference in means and point to a sizeable improvement in MLA between grades 6 and 7. This finding corresponds with Tellier (2013), who found that "something may be happening in children's development between the ages of 8 and 9, and again between the ages of 11 and 12", (Tellier in Roehr, 2013, p. 29) pointing to a significant statistical increase in means between these ages (Tellier, 2013, p. 30). As stated in section 3.5.2, the participants in grade 6 did speak to each other on a couple of occasions. The researcher noticed the disturbance and corrected the participants' disorder promptly. Consequently, there is a low risk that the disturbance has altered the results. However, it is plausible that some found the test difficult and therefore sparked the disturbance.

5.2 The correlation between the participants' English competence and MLA

On the issue of the correlation between MLA and the participants' English competence, it is necessary to state that the participants' English competence was reported by the participants' English teachers. This study has not collected any other data to state the participants' English competence. If the study had assessed the participants' English competence by a language test, the outcome in assessment could have been different. Still, the teacher-reports demonstrates what level of proficiency the teachers believe the participants are at and can be used as a comparison measurement against their state of MLA. As shown in Table 15 and 16, no correlations between the participants' state of MLA and English competence was discovered. It must be acknowledged that P1-7C was assessed on level 1 in English competence and received the lowest score in the test. However, P16-7C was assessed as a participant on level 4 in English competence and received the second-lowest score on the test. Likewise, P17-7C received the highest score in grade 7 and was considered a level 4 student in English competence. P9-7C was one out of the two participants in grade 7 who was considered at the top level in English competence (level 6) but received a score of 17 points below P17-7C. The study detected the same tendencies in grade 6. The participant who received the highest MLA score was assessed as a level 1 student in English competence. Participant P1-6A was the only participant in grade 6 that was considered as a level 6 student in English competence but received 17 points in the MLA test, which is 6 points below the median score for grade 6. These results show little to no correlation between the participants'

English competence and the state of MLA. It is feasible that the participants' level of English-competence could have been higher or lower if the study had included a language test to report their competence. Still, the presented difference in MLA and reported English competence is noteworthy. There might be different reasons for the demonstrated difference. One of them can be that the tasks in the MLA test differed from the usual tasks the participants have worked with previously. Some might have found the dissimilarity confusing which resulted in weaker test-results. However, there were many participants with a low teacher reported English competence that received a high score in the MLA test. Since these participants are grade 6 and 7 pupils, it is likely that parts of the English course are conducted in Norwegian. Therefore, the study investigated if the participants who received 1, 2 and 3 in English competence were participants with a different L1 than Norwegian. If this was the case, it could indicate that their Norwegian competence could have affected the teachers' overall impression of the participants' competence in the course. However, the study did not find any correlations between the teacher's assessments and the participants' L1. The instructions for assessment were to assess both their oral and written competence. Therefore, some of the participants that are not orally active in the classroom might have received a lower level of competence than if the assessment of oral competency were excluded. Though, this study cannot make any conclusion in this regard.

5.3 The potential positive relationship between MLA and bilingualism

As stated in section 2.3.3, some studies have found that bilingual children develop a control of processing and a control of attention more advantageously than monolingual children (Bialystok, 2001, 2004) and might have an advance in developing their metalinguistic awareness. This is also supported by Cummins who claimed that bilinguals could have the advantage of comparing languages more easily and probably develop control of processing to a higher degree than monolingual children (Cummins, 1991, cited in Jessner, 1999). Therefore, some of the participants who claimed to be bilingual will be discussed in comparison to their state of MLA.

Participant P22-6A reported a fluent oral and written proficiency in Norwegian, English and Swedish. In addition to this, the participant reported good written and oral proficiency in French. This participant received the highest score of MLA, with 48 out of the 57 possible points. On the other hand, participant P18-7C reported a fluent oral proficiency in Norwegian

and Spanish as well as good oral proficiency in Turkic. In addition to this, it was reported some oral proficiency in French, Spanish and English and some written proficiency in German, Turkic, French, and Spanish. This participant received an MLA score of 22 points and was considered as a level 5 student in English competence. The two other participants that reported a fluent proficiency in other languages than in Norwegian in grades 6 and 7 were P6-6A, P16-6A, and P23-6A. P6-6A reported a fluent oral and written proficiency in Norwegian, English as well as a fluent oral proficiency in Somali. This participant received an MLA score of 15 points. P16-6A reported fluent oral and written proficiency in Norwegian and fluent oral proficiency in Arabic and scored 24 points in the test. Lastly, P23-6A reported fluent oral and written proficiency in Norwegian and Turkic and received an MLA score of 14 points. No clear indications for a bilingual advantage in MLA are found in this study, just as there were none for an overall correlation between MLA and English competence. However, there might be some tasks that bilingual children solved more successfully than monolingual children. This has not been investigated as it is not one of the study's aims. Nevertheless, it would be favorable to research this at a later time. To summarize, the present study neither found evidence for any clear correlation between participants' degree of MLA and their teacher-assessed English competence nor evidence for a higher MLA for participants with self-reported bi- or multilingualism.

5.3 Conclusion and further research

This study found a clear improvement in means in MLA between grades 6 and 7, which is consistent with Tellier's (2013) findings. Additionally, there was no evident correlation between the participants' teacher-reported English competence and state of MLA in either class. Moreover, the results correlate with previous research of children developing their MLA as they mature in age (Bialystok, 2001, p. 14, Tellier, 2013, p. 36-37). However, as both this study as well as Tellier (2013) found a clear difference in means between grades 6 and 7, it is encouraged to conduct more research on MLA in this age-group to investigate this difference further. Additionally, it would have been valuable to investigate how the bilingual children performed in comparison to the monolingual children per task, as well as to investigate whether there was a gender difference in the task scores. As presented in sections 2.2-2.5, several studies show clear benefits in developing pupils' metalinguistic awareness. Awareness of rhymes, lexical ambiguities and syntax can evolve the individuals reading comprehension as well as their L2 proficiency. Language teachers have numerous

possibilities in developing their pupils' metalinguistic awareness. Teachers can provide the class with texts that require pupils to find ambiguous words or metaphors, re-arrange sentences to create a different meaning, make word-games and rhymes and so forth. As the presented research in this study shows, a positive increase in children's level of MLA can have positive effects on their development of their L2 proficiency. However, as this study did not find indications for correlations between a participants MLA and English competence, further research is needed.

References

Andrade, A. I. (2004). *Janua Linguarum- The gateway to languages: The introduction of language awareness into the curriculum.* (M. Candelier, Ed.). Kapfenberg: Council of Europe Publishing. Retrieved from
http://archive.ecml.at/documents/pub121e2004candelier.pdf

Bialystok, E. (2001). *Bilingualism in development*. Cambridge: Cambridge University Press.

Bialystok, E. (2002). *Metalinguistic aspects of bilingual processing* (pp. 169-181). *Annual Review of Applied Linguistics, 21.*
http://doi.org.galanga.hvl.no/10.1017/S0267190501000101

Bialystok, E. (2004). *Bilingualism in development*. Cambridge: Cambridge University Press.

Christoffersen, L., & Johannessen, A. (2012). *Forskningsmetode for lærerutdanningene*. Oslo: Abstrakt forl.

Creswell, J., & Creswell, J. David. (2018). *Research design: Qualitative, quantitative & mixed methods approaches* (5th ed.). Los Angeles, California: Sage.

EVLANG (1997-2001). Sokrates/Lingua project «Éveil aux langues» 42137-CP-1-97-1-FR-Lingua-LD. EVLANG test material received directly from the principal investigator, Michel Candelier.

Jessner, U. (1999). *Metalinguistic Awareness in Multilinguals: Cognitive Aspects of Third Language Learning*. New York: Routledge, Taylor & Francis Group.
https://doi.org/10.1080/09658419908667129

Jessner, U. (2008). Language Awareness in Multilinguals: Theoretical Trends. In: Hornberger N.H. (eds) Encyclopedia of Language and Eduacation. Boston, MA: Springer US.

Krogtoft, M., & Sjøvoll, J. (2018). *Masteroppgaven i lærerutdanninga: Temavalg, forskningsplan, metoder*. Oslo: Cappelen Damm akademisk.

Mackey, A., & Gass, S. (2016). *Second Language Research: Methodology and Design*. New York: Routledge.

McCabe, J. A. (2017). *Floor and Ceiling Effects*. (A. Wenzel, Ed.). Thousand Oaks: SAGE Publication. doi: http://dx.doi.org/10.4135/9781483365817.n598

Morse, J., & Niehaus, L. (2009*). Mixed method design: Principles and procedures* (Vol. Volume 4, Developing qualitative inquiry). Walnut Creek, Calif: Left Coast Press.

Roehr-Brackin, K. (2018). *Metalinguistic Awareness and Second Language Acquisition*. New York: Routledge.

Roehr, K., & Gánem-Gutiérrez, G. A. (Eds.). (2013). *The Metalinguistic Dimension in*

Instructed Second Language Learning. London: Bloomsbury Academic.
doi:10.5040/9781472542137

Roehr, K., & Thepseenu, B. (2013). University-Level Learners' Beliefs about Metalinguistic
Knowledge. In *The Metalinguistic Dimension in Instructed Second Language
Learning* (pp. 95-118). London: Bloomsbury Academic.
doi:10.5040/9781472542137.ch-005

Sartori, M. E., Cassuto, P., & Giolfo, M. (2017). *Approaches to the history and dialectology
of Arabic in honor of Pierre Larcher.* Leiden: Brill. doi:10.1163/9789004325883_001

Simard, D., Foucambert, D., & Labelle, M. (2013). Examining the Contribution of Syntactic
and Metasyntactic Abilities to Reading Comprehension among Native and Non-Native
Speakers. In *The Metalinguistic Dimension in Instructed Second Language
Learning* (pp. 45-70). London: Bloomsbury Academic.
doi:10.5040/9781472542137.ch-003

Tellier, A. (2013). Developing a Measure of Metalinguistic Awareness for Children Aged 8–
11. In *The Metalinguistic Dimension in Instructed Second Language Learning* (pp.
15–44). London: Bloomsbury Academic . doi:10.5040/9781472542137.ch-002

Thurén, T., Gjerpe, K., & Gjestland, D. (2009). *Vitenskapsteori for nybegynnere.* Oslo:
Gyldendal akademisk.

Tunmer, W., Pratt, C., and Herriman, M. (ed.). (1984). *Metalinguistic Awareness in
Children: Theory, research and implications* (pp. 285-289). Springer, Berlin.

Simard, D., & Gutiérrez, X. (2018). The Study of Metalinguistic Constructs in Second
Language Acquisition Research. In Garrett, P., & Maria, C. i C. J. (Ed.), *The
Routledge handbook of language awareness* (pp. 205-222). New York: Routledge,
Taylor & Francis Group.

Utdanningsdirektoratet. (2019). *Læreplan i Engelsk.* Retrieved April 8, 2020, from
https://www.udir.no/lk20/eng01-04/kompetansemaal-og-vurdering/kv3

Utdanningsdirektoratet. (2019). *Kjerneelementer.* Retrieved April 8, 2020 from
https://www.udir.no/lk20/eng01-04/om-faget/kjerneelementer

Utdanningsdirektoratet. (2019). *Motivere og ha positive forventninger.* Retrieved April 8,
2020, from https://www.udir.no/laring-og-trivsel/klasseledelse/motivasjon-og-
forventninger/

Zhang, D., Chin, C., & Li, L. (2017). Metalinguistic awareness in bilingual children's word
reading: A cross-lagged panel study on cross-linguistic transfer facilitation (pp. 6-
23). In *Applied Psycholinguistics.* National Institute of Education, Nanyang

Technological University.

Zipke, M. (2008). *Teaching metalinguistic awareness and reading comprehension with riddles: Understanding that words and sentences can have more than one meaning improves comprehension by allowing readers to think flexibly about the appropriate meaning* (pp. 128-137). The Reading Teacher.

Appendices

Appendix 1: Table of tasks and its components

ITEM	MLA task (operationalization)	Degree of complexity	MLA	Language	Construct investigated: awareness of….	Linguistic domain
Tellier item 1	cognates	simple	Crosslinguistic awareness	**Natural languages** Spanish Portuguese Italian Dutch Romanian Esperanto English Norwegian	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 2	cognates	complex	Crosslinguistic awareness	**Natural languages** Italian German Esperanto	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 3	Plural marking	simple	General MLA	**Artificial languages**	Systematicity in inflectional morphology	inflectional morphology
Tellier item 7	Case marking	complex	General MLA	**Artificial languages**	Systematicity in inflectional morphology; additional element: what constitutes a sentence (word order)	inflectional morphology & syntax
EVLANG item 1.5	Chinese numbers	Simple	General MLA	**Natural language** Chinese	Systematicity in word formation	Word formation (compounding)
EVLANG item 2.5	Sentence in unknown L	complex	General MLA	**Artificial language**	Systematicity in what constitutes a sentence (word order)	Syntax

Appendix 2: Task 1

1:

Hvilke setninger betyr det samme?

1. El gato es blanco.

2. O elefante é grande.

3. La rana è verde.

4. De olifant is groot.

5. Tigrul este negru şi verde.

6. La rano estas verda.

7. The cat is white.

Par:

Setningen som er igjen betyr:

Norsk: _____

Engelsk: _____

Appendix 2: Task 1

2. Benytt alle språkene du kan for å hjelpe deg med å finne ut hva setningene nedenfor setningene betyr. Skriv hele setningen, eller så mange ord som du kan.

Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato.

_____

_____

Der Hund, die Katze und die Maus sind sehr gute Freunde.

_____

_____

Estas bela tago do mi volas vanilan glaciaĵon kaj limonadon

_____

_____

Appendix 3: Task 2

Appendix 4: Task 3.1

3. Se på dette ordparet.
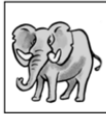
en *elefant* ➡ *elefanto*

Mer enn en elefant ➡ elefantoj

....

Prøv å finne ordpar fra listen nedenfor. Et av ordene i ordparet skal bety **en** og et av ordene i ordparet skal bety **mer enn en.** Skriv numrene på ordparene i boksene nedenfor.

1. oligare

2. garbatáns

3. saminii

4. rahak

5. garbatan

6. oligares
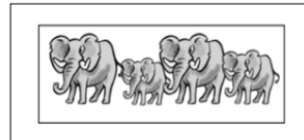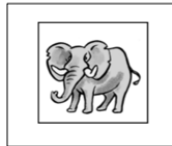
og

og

og

og

Appendix 5: Task 3.1

7. rahaki

8. samini

Nedenfor står et ord fra det samme språket som ordet "samini" er fra. Arkari er ordet for **en.** Hva sier du om det er **mer enn en?**
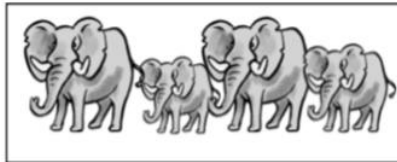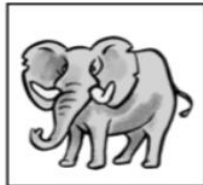
Arkari

Appendix 6: Task 3.2

4. Nå skal vi jobbe med et språk som kalles "ograma". Se på navnene på disse to dyrene.

| ogrtsi | | omsu |

....

Det finnes to ulike måter på dette språket å beskrive hva som skjer på bildet nedenfor.

| ogrtsi grarid omsu-ch | | omsu-ch grarid ogrtsi |

....

Hvordan vil du beskrive det som skjer på bildet nedenfor? Skriv det på språket "ograma" på to ulike måter.

Appendix 7: Task 4

Appendix 8: Task 5.1

5. Nedenfor ser du tall som er skrevet på kinesisk.

1	yi
2	er
3	san
4	si
5	wu
10	shi
11	shiyi
12	shier
20	ershi
23	ersisan

Bruk listen med tall som hjelp og eksempelet som modell for å skrive numrene til de kinesiske tallene nedenfor.

Eksempel: shisan: 13

shisi: _____

ershiwu: _____

sanshisan: _____

Appendix 8: Task 5.1

Appendix 9: Task 5.2

Nå er det din tur. Bruk eksempelet nedenfor som modell og skriv følgende tall på kinesisk:

Eksempel: 13: shisan

15: _____

24: _____

....

Appendix 9: Task 5.2

6. Her er tre setninger på språket "krarid". Du finner den norske

oversetningen nedenfor:

En	inu	betsiaki	*Jeg har sett jaguaren*
Jeg	*jaguaren*	*har sett*	
En	baka	betsiaki	*Jeg har sett fisken*
Jeg	*fisken*	*har sett*	
Min	baka	betsiaki	*Du har sett fisken*
Du	*fisken*	*har sett*	

Se nærmere på hvordan setningene er laget og prøv å skriv setningen

nedenfor på språket "krarid".

Du har sett jaguaren

_____

Appendix 10: Task 5.2

Noen spørsmål om deg:

Hvor gammel er du?

Jeg er _____

1. **Er du gutt, jente eller identifiserer du deg som noe annet?**

Gutt ☐   Jente ☐   Annet ☐

2. **Hvor lenge har du bodd i Norge?**

Hele livet ☐   Mindre enn to år ☐   Mindre enn ti år ☐

3. **Hvilke språk snakker du?**

Sett et kryss under "muntlig" og et kryss under "skriftlig" for hvert av språkene.

Språk	Muntlig			Skriftlig		
	Litt	Bra	Flytende	Litt	Bra	Flytende

Appendix 11: Language background, p. 1

4. Hvilke språk bruker du oftest i følgende situasjoner:

Når du forteller moren din om dagen på skolen:

_____

Når du leker med søskene dine hjemme:

_____

Hvis du vil kjøpe noe og må be faren din betale for det:

_____

Når du forteller vennene dine om et program du har sett på TV:

_____

Når du spiller dataspill:

_____

Når du ser på TV eller YouTube:

_____

Appendix 12: Language background, p. 2

Appendix 13: Consent Form in Norwegian

Samtykkeskjema

Språklig bevissthet til 6. og 7. trinn

Marianne Bakken
Tlf: 410 91 376
E-mail: Marianne_bkkn@hotmail.com/Marianne.bakken@bergen.kommune.no

Introduksjon og invitasjon til studie
Hei. Mitt navn er Marianne Bakken. Jeg har jobbet som vikar på [1][...] skole i tre år og skal i år skrive en masteravhandling i Undervisningsvitenskap. Jeg skriver om språklig bevissthet. Kort fortalt handler det om å forstå former og strukturer til språk. Jeg skal undersøke 6. og 7. klassinger sin språklige bevissthet ved å gi de en prøve som vil vare i ca. 60 minutter og som vil gjennomføres i en vanlig klassetime. Elevene vil blant annet bes å se på strukturen til et språk de ikke kjenner og å finne ut om ukjente ord er i entall eller flertall.

Denne prøven vil ikke bli brukt for å vurdere elevene eller for senere undervisning. Prøven vil kun bli brukt som datamateriale for masteroppgaven. Det vil si at det ikke vil ha noen konsekvenser for barnet ditt om du velger at han/hun ikke skal delta.

Ved å forstå elevenes grad av metalingvistisk bevissthet vil språklærere ha bedre muligheter til å tilpasse undervisningsopplegg til hver enkelt elev og å forstå enkeltelevens utfordringer i språkopplæring. Dette er noe jeg håper masteroppgaven min kan bidra til.

Anonymitet og konfidensialitet
Elevene vil ikke bli bedt om å skrive navn, fødselsdato eller noe annet som kan identifisere de. Skolen eller området i Bergen vil heller ikke bli oppgitt i masteroppgaven, deltakerne blir oppgitt som «klasse 6» og «klasse 7» i en skole på Vestlandet.

Prøvesvarene og alle andre eventuelle opplysninger som blir hentet inn vil være konfidensielle og kun bli sett av meg og masterveilederen min.

Samtykke:
Det er opp til deg og barnet ditt om du vil at barnet ditt skal bli med på prøven. Jeg hadde satt stor pris på deltakelsen for å få så mange svar som mulig, men det er som sagt helt frivillig. Du kan også velge å trekke samtykket ditt når som helst og uten å oppgi grunn og uten at det vil ha noen konsekvenser. Om barnet ditt kan delta i prøven, vær så snill å signer navnet ditt og dato nederst i skjemaet.

Kontakt og dato for prøven:
Om du/de har spørsmål, kontakt meg gjerne på telefon eller mail som er oppgitt øverst i skjemaet. Prøven vil skje i uke 47. Jeg hadde satt stor pris på om skjemaet blir levert til kontaktlærer før fredag 8. november.

Barnet mitt kan delta

Signatur_____ Dato_____

---

[1] Name of school is eliminated to ensure the participants' anonymity.

Appendix 13: Consent Form in Norwegian

Samtykkeskjema

Språklig bevissthet til 6. og 7. trinn

Marianne Bakken
Tlf: 410 91 376
E-mail: Marianne_bkkn@hotmail.com/Marianne.bakken@bergen.kommune.no

Introduksjon og invitasjon til studie

Hei. Mitt navn er Marianne Bakken. Jeg har jobbet som vikar på [1][...] skole i tre år og skal i år skrive en masteravhandling i Undervisningsvitenskap. Jeg skriver om språklig bevissthet. Kort fortalt handler det om å forstå former og strukturer til språk. Jeg skal undersøke 6. og 7. klassinger sin språklige bevissthet ved å gi de en prøve som vil vare i ca. 60 minutter og som vil gjennomføres i en vanlig klassetime. Elevene vil blant annet bes å se på strukturen til et språk de ikke kjenner og å finne ut om ukjente ord er i entall eller flertall.

Denne prøven vil ikke bli brukt for å vurdere elevene eller for senere undervisning. Prøven vil kun bli brukt som datamateriale for masteroppgaven. Det vil si at det ikke vil ha noen konsekvenser for barnet ditt om du velger at han/hun ikke skal delta.

Ved å forstå elevenes grad av metalingvistisk bevissthet vil språklærere ha bedre muligheter til å tilpasse undervisningsopplegg til hver enkelt elev og å forstå enkeltelevens utfordringer i språkopplæring. Dette er noe jeg håper masteroppgaven min kan bidra til.

Anonymitet og konfidensialitet

Elevene vil ikke bli bedt om å skrive navn, fødselsdato eller noe annet som kan identifisere de. Skolen eller området i Bergen vil heller ikke bli oppgitt i masteroppgaven, deltakerne blir oppgitt som «klasse 6» og «klasse 7» i en skole på Vestlandet.

Prøvesvarene og alle andre eventuelle opplysninger som blir hentet inn vil være konfidensielle og kun bli sett av meg og masterveilederen min.

Samtykke:

Det er opp til deg og barnet ditt om du vil at barnet ditt skal bli med på prøven. Jeg hadde satt stor pris på deltakelsen for å få så mange svar som mulig, men det er som sagt helt frivillig. Du kan også velge å trekke samtykket ditt når som helst og uten å oppgi grunn og uten at det vil ha noen konsekvenser. Om barnet ditt kan delta i prøven, vær så snill å signer navnet ditt og dato nederst i skjemaet.

Kontakt og dato for prøven:

Om du/de har spørsmål, kontakt meg gjerne på telefon eller mail som er oppgitt øverst i skjemaet. Prøven vil skje i uke 47. Jeg hadde satt stor pris på om skjemaet blir levert til kontaktlærer før fredag 8. november.

Barnet mitt kan delta

Signatur_____ Dato_____

---

[1] Name of school is eliminated to ensure the participants' anonymity.

4. Hvilke språk bruker du oftest i følgende situasjoner:

Når du forteller moren din om dagen på skolen:

_____

Når du leker med søskene dine hjemme:

_____

Hvis du vil kjøpe noe og må be faren din betale for det:

_____

Når du forteller vennene dine om et program du har sett på TV:

_____

Når du spiller dataspill:

_____

Når du ser på TV eller YouTube:

_____

Noen spørsmål om deg:

Hvor gammel er du?

Jeg er _____

1. Er du gutt, jente eller identifiserer du deg som noe annet?

Gutt ☐   Jente ☐   Annet ☐

2. Hvor lenge har du bodd i Norge?

Hele livet ☐   Mindre enn to år ☐   Mindre enn ti år ☐

3. Hvilke språk snakker du?

Sett et kryss under "muntlig" og et kryss under "skriftlig" for hvert av språkene.

Språk	Muntlig			Skriftlig		
	Litt	Bra	Flytende	Litt	Bra	Flytende

1:

Hvilke setninger betyr det samme?

1. El gato es blanco.

2. O elefante é grande.

3. La rana è verde.

4. De olifant is groot.

5. Tigrul este negru şi verde.

6. La rano estas verda.

7. The cat is white.

Par:

 and

 and

 and

Setningen som er igjen betyr:

Norsk: _____

Engelsk: _____

2. Benytt alle språkene du kan for å hjelpe deg med å finne ut hva setningene nedenfor setningene betyr. Skriv hele setningen, eller så mange ord som du kan.

Ho una tazza di caffè con latte e zucchero, e una torta al cioccolato.

_____

_____

Der Hund, die Katze und die Maus sind sehr gute Freunde.

_____

_____

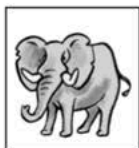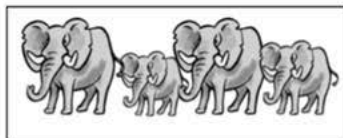Estas bela tago do mi volas vanilan glaciaĵon kaj limonadon

_____

_____

7. rahaki

8. samini

3. Se på dette ordparet.

en *elefant* ➡ *elefanto*

Mer enn en elefant ➡ elefantoj

....

Prøv å finne ordpar fra listen nedenfor. Et av ordene i ordparet skal bety **en** og et av ordene i ordparet skal bety **mer enn en.** Skriv numrene på ordparene i boksene nedenfor.
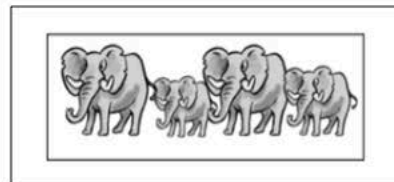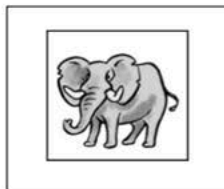
1. oligare

2. garbatáns

3. saminii

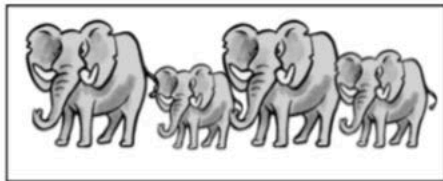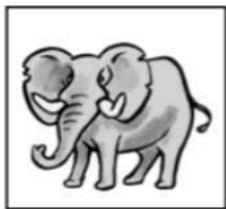4. rahak

5. garbatan

6. oligares

	og	
	og	
	og	
	og	

Nedenfor står et ord fra det samme språket som ordet "samini" er fra. Arkari er ordet for **en.** Hva sier du om det er **mer enn en?**

Arkari

4. Nå skal vi jobbe med et språk som kalles "ograma". Se på navnene på disse to dyrene.

ogrtsi

omsu

Det finnes to ulike måter på dette språket å beskrive hva som skjer på bildet nedenfor.

ogrtsi grarid omsu-ch

omsu-ch grarid ogrtsi

Hvordan vil du beskrive det som skjer på bildet nedenfor? Skriv det på språket "ograma" på to ulike måter.

5. Nedenfor ser du tall som er skrevet på kinesisk.

1   yi
2   er
3   san
4   si
5   wu

10   shi
11   shiyi
12   shier

20   ershi
23   ersisan

Bruk listen med tall som hjelp og eksempelet som modell for å skrive
numrene til de kinesiske tallene nedenfor.

> *Eksempel:* shisan: 13

shisi: _____

ershiwu: _____

sanshisan: _____

Nå er det din tur. Bruk eksempelet nedenfor som modell og skriv følgende tall på kinesisk:

Eksempel: 13: shisan

15: _____

24: _____

....

6. Her er tre setninger på språket "krarid". Du finner den norske oversetningen nedenfor:

En	inu	betsiaki	Jeg har sett jaguaren
Jeg	*jaguaren*	*har sett*	

En	baka	betsiaki	Jeg har sett fisken
Jeg	*fisken*	*har sett*	

Min	baka	betsiaki	Du har sett fisken
Du	*fisken*	*har sett*	

Se nærmere på hvordan setningene er laget og prøv å skriv setningen nedenfor på språket "krarid".

Du har sett jaguaren

_____

ITEM	MLA task (operationalization)	Degree of complexity	MLA	Language	Construct investigated: awareness of….	Linguistic domain
Tellier item 1	cognates	simple	Crosslinguistic awareness	**Natural languages** Spanish Portuguese Italian Dutch Romanian Esperanto English Norwegian	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 2	cognates	complex	Crosslinguistic awareness	**Natural languages** Italian German Esperanto	Lexical relationships between Ls; (elements of "what constitutes a sentence")	Lexical semantics & syntax
Tellier item 3	Plural marking	simple	General MLA	**Artificial languages**	Systematicity in inflectional morphology	inflectional morphology
Tellier item 7	Case marking	complex	General MLA	**Artificial languages**	Systematicity in inflectional morphology; additional element: what constitutes a sentence (word order)	inflectional morphology & syntax
EVLANG item 1.5	Chinese numbers	Simple	General MLA	**Natural language** Chinese	Systematicity in word formation	Word formation (compounding)
EVLANG item 2.5	Sentence in unknown L	complex	General MLA	**Artificial language**	Systematicity in what constitutes a sentence (word order)	Syntax