

DO WE HAVE TO SPEAK ENGLISH?

A longitudinal perspective on the impact of early English input on the development of English comprehension skills in Norwegian pupils.

Odin Lein Strand

Master's thesis in English

Department of Language and Literature

Norwegian University of Technology and Science

Spring 2014

ABSTRACT

This study is a continuation of previous research done on a group of Norwegian pupils. When this group started first grade they partook in a project by Dahl and Vulchanova (2014) which studied how input based teaching, and some use of English outside of English lessons, affected the early acquisition of English as a second language. They found that the pupils exposed to naturalistic input outperformed the control group after just eight months in both receptive vocabulary size and listening comprehension skills. However, Sivertzen (2013) found that two years after the discontinuation of the input based teaching programme the pupils no longer outperformed their peers in receptive vocabulary size.

The same participants were the focus of this study; at the time of testing they were halfway through fifth grade. Two tests were used, one listening comprehension test and one reading comprehension test. Subsequently, the results of these tests were analysed in SPSS, showing significant difference in performance between the extra input group and the control group. Results from existing research indicate that receptive vocabulary can be used as a good predictor of language comprehension skills, which the findings of this study are in conflict with. It is hypothesized that the significant difference in performance has to do with the quality of input they received during first and second grade, and that this has facilitated language comprehension skills that cannot necessarily be detected by vocabulary size testing.

ACKNOWLEDGEMENTS

For this project I would first of all like to thank Mila Vulchanova for her support in making this project come true, and her inexhaustible knowledge of research and theories within the field of psycholinguistic. I would also like to direct special thanks to Anne Dahl for letting me use and develop her listening comprehension test, and for always making time when I have questions about SLA research. Furthermore I would like to give special thanks to Bjørn Grønnesby and his colleagues for taking the time to programme the test, I am forever grateful for your competence and help. I would also like to thank Luke James Barber for contributing with his impeccable British accent in the recordings for the listening comprehension tests. Lastly I would like to direct my gratitude towards Camilla Hellum Foyn for sharing her competence in statistics, you turned the part of the thesis I feared the most into a learning experience that opened my eyes to the world of statistics.

I would also like to thank the schools for being ever so helpful and flexible in the process of the data collection. Furthermore, I am very grateful for the effort the pupils put into the tests, as well as the parents who allowed the pupils to partake and took the time to fill in the questionnaires.

At last I would like to thank my partner, Anette Meek, for always lending an ear and helping me out when I needed a second pair of eyes, and for always supporting me regardless of how I felt about the task at hand, without you I would not have been able to finish the thesis.

TABLE OF CONTENTS

ABSTRACT	III
ACKNOWLEDGEMENTS	V
LIST OF ABBREVIATIONS	IX
1.0 INTRODUCTION	1
2.0 THEORETICAL BACKGROUND	3
2.1 SECOND LANGUAGE ACQUISITION	3
2.2 THE CRITICAL PERIOD AND THE AGE OF LANGUAGE ONSET	5
2.3 ACQUIRING A VOCABULARY	7
2.4 BECOMING LITERATE IN THE L2	8
2.5 VOCABULARY AS A DETERMINANT FOR COMPREHENSION SKILLS	10
2.6 THE EFFECTS OF INPUT BASED LANGUAGE TEACHING IN SLA	11
3.0 METHOD	15
3.1 PARTICIPANTS	15
3.1.1 EXTRA ENGLISH INPUT CONDITIONS GROUP	16
3.1.2 NORMAL ENGLISH INPUT CONDITIONS GROUP	16
3.2 BACKGROUND INFORMATION ON THE CHOSEN SCHOOLS	17
3.3 PROCEDURE	18
3.4 CONSENT FORMS AND QUESTIONNAIRES	18
3.5 TESTS	18
3.5.1 LISTENING COMPREHENSION TEST	19
3.5.2 READING COMPREHENSION TEST	20
4.0 ANALYSIS & RESULTS	23
4.1 MANOVA	23
4.2 ANOVA	25
4.3 LISTENING COMPREHENSION TEST	26
4.4 READING COMPREHENSION TEST	28
4.5 DATA FROM THE QUESTIONNAIRES	28
4.6 LIMITATIONS	30
5.0 DISCUSSION	33
5.1 THE IMPLICATIONS OF THE MANOVA AND ANOVA RESULTS	33
5.2 LISTENING COMPREHENSION SKILLS	35
5.3 LITERACY AND READING COMPREHENSION	37
6.0 CONCLUSION	41
6.1 QUESTIONS FOR FURTHER RESEARCH	43
7.0 BIBLIOGRAPHY	45
APPENDICES	47
APPENDIX 1: PARENTAL CONSENT FORM	47
APPENDIX 2: QUESTIONNAIRE ABOUT LANGUAGE BACKGROUND	48
APPENDIX 3: LISTENING COMPREHENSION TEST	51
APPENDIX 4: READING COMPREHENSION TEST	58
APPENDIX 5: FREQUENCY DATA FROM THE LISTENING COMPREHENSION TEST	60

LIST OF ABBREVIATIONS

L1	First language
L2	Second Language
NI	Normal English input conditions group
EI	Extra English input conditions group
CP	Critical Period
UG	Universal Grammar
SLA	Second language acquisition
CPH	Critical period hypothesis
CDI	The MacArthur Communicative Development Inventory
ANOVA	Analysis of Variance
MANOVA	Multivariate Analysis of Variance

1.0 INTRODUCTION

It is a well-established view within language research that children acquire language(s) more effortlessly than adults. Although, several different theoretical approaches have been put forward to explain why this is the case. This study will be a contribution to this field of research, as it is a continuation of a longitudinal project that was started almost five years ago by Anne Dahl and Mila Vulchanova when they tested a group of Norwegian first graders. The same group was subject for Tonje Gauslaa Sivertzen's masters thesis last year, where Sivertzen retested the group with the PPVT-4 test, which was one of the tests conducted by Dahl and Vulchanova in the initial study. The role of this study has been to retest and expand one of the other tests that was conducted on the group in the first grade, a listening comprehension test developed by Dahl and Vulchanova. The following has been the central research question for this study:

To what extent does vocabulary and comprehension skills correlate with each other in Norwegian second language users of English, and does early exposure to authentic language leave long lasting advantages in language comprehension skills in young language learners.

The current Norwegian curriculum has fixed amounts of teaching hours for each subject. For the period between first and fourth grade Norwegian pupils receive a total amount of 138 hours of English teaching, which equals less than one hour per week. In comparison, pupils in upper secondary school receive a grand total of five 45 minutes lessons a week, in their last year(s) of English teaching (Utdanningsdirektoratet, 2013a). This distribution where the highest density of English teaching takes place when pupils have reached the age of 16 years stands in stark contrast to research on second language acquisition. Dahl and Vulchanova's study (2014) set out to document the effect input based teaching and colloquial use of English outside the English classroom setting had on first graders. The results after eight months displayed resounding improvement in both receptive vocabulary and comprehension skills when compared to a control group that had received regular English teaching with no extra focus on input- or alternative teaching methods (Dahl, 2014; Dahl & Vulchanova, 2014). Since there was such positive effects of the extra English input the school proceeded with the program throughout second grade, even though the research project was officially completed. After second grade they no longer had a specific focus on this issue (Sivertzen, 2013). Thus, when

Sivertzen set out to test the vocabulary almost 2 years after the initial project was finished one would expect that the tests would reveal an advantage compared to peers that had undergone regular English teaching. However, even though the raw scores showed differences in vocabulary size, it turned out that the difference was not statistically significant. It was hypothesized that the levelling out in development was caused by the discontinuation of the naturalistic input programme (Sivertzen, 2013).

In this study the English comprehension skills of the pupils will be assessed by two tests; a listening comprehension tests where the pupils are asked to match pictures with spoken input; a reading comprehension test where the pupils are presented with a short fictional text that they have to answer eight questions about afterwards. The results of these two tests will be used in a Multivariate Analysis of Variance to analyse how the pupils compare to peers that have not received input based teaching during the first years of schooling, thus give a better understanding of the correlation between vocabulary size and comprehension.

2.0 THEORETICAL BACKGROUND

This chapter is divided into 6 subchapters. 2.1 *Second Language Acquisition* deals with a general overview of the field of SLA, as well as placing it in the context of this study. 2.2 *The Critical Period and the Age of Language Onset* discusses how age effects SLA. 2.3 *Acquiring a Vocabulary* distinguishes the terms incidental and intentional language learning and present the CDI scale. 2.4 *Becoming Literate in the L2* explains the processes leading up to literacy. 2.5 *Vocabulary as Determinant for Comprehension Skills* presents data that emphasise the importance of vocabulary in comprehension. 2.6 *The Effects of Input Based Teaching in SLA* presents theory and recent studies that highlight the importance of exposure to the target language in SLA.

2.1 SECOND LANGUAGE ACQUISITION

The study of SLA, *second language acquisition*, is a term with a wide scope in that it describes not only the second language people acquire, but all languages acquired subsequent to their L1. The term also encompasses the acquisition of dialects, naturalistic and formal training, as well as foreign language and lingua franca settings (Doughty & Long, 2003; Saville-Troike, 2006). In an attempt to describe the diversity of SLA Macaro (2010) says the following.

“Second language learning is neither cumulative nor neat and tidy. And it is in that messy, non-linear, but semi-permanent process of learning the subject, and in the myriad situations in which a second language is engaged, that the complexity of second language learning resides” (p. 4).

Macaro’s attempt to describe the SLA field touches upon an important note, which is that SLA research is an applied form of linguistic research, that draws upon both research from cognitive psychology and linguistics. This cross-field transfer has to do with the simple fact that language learning a highly unique cognitive process, and the study of it can shed more light on the nature of the human mind. One other rather obvious field of study that SLA research contributes to, is the pedagogic aspect of language teaching (Doughty & Long, 2003), which will be a part of the scope of this study since it looks at early English input.

In the current Norwegian curriculum English has the status as a second language, not foreign language (Kunnskapsdepartementet, 2009). Norwegian pupils are being taught English from first grade up until tenth grade, as well as in upper secondary school,

which is optional. Whereas, they first start learning what is labelled foreign language in eight grade, most commonly Spanish, German or French (Utdanningsdirektoratet, 2013a). In the *purpose* section of the current English curriculum, it is argued that English has become a lingua franca, and that Norwegian pupils need to acquire high competence in English both for personal gain and to meet the needs at the job market and society. (Utdanningsdirektoratet, 2013b).

A holistic perspective on the process of acquiring a language, both L1 and L2, can be divided into three sequential stages; *initial stage*, *intermediate stage*, and *final stage* (Saville-Troike, 2006, p. 17). The *initial stage* is where the speaker first acquires the language. The difference between L1 and L2 acquisition at this point is mainly that a person acquiring an L2 already have labels and concepts from his or her L1 (Saville-Troike, 2006, p. 17). *Intermediate stage* describes the period from which the speaker is able to produce and comprehend simple utterances all the way up to the stage of ultimate attainment of a language. At this stage we find the largest discrepancy in cognitive maturity. L2 learners already possess the ability to produce complex utterances in their L1, while that is the ultimate goal for L1 learners at this stage. L2 learners will at this stage start to develop an *interlanguage* where they apply *transfer* of attained skills in their L1 in the process of attaining competence in the L2. *Positive transfer* will occur when there is correlation between two languages, thus making it possible to apply structural and grammatical rules from the L1 onto the L2. Whereas, *negative transfer* is the opposite, when the speaker applies knowledge from the L1 that causes disturbance in the language which leads to the production of grammatical anomalies and incorrectness (Saville-Troike, 2006, pp. 16-17). A more in-depth presentation of the various processes that go into both the *initial and intermediate stage* will be given throughout this chapter. The *final stage* is reached when the L1 user has reached native linguistic competence, however for the L2 user the final stage is a more undefined stage. Some L2 learners reach near native competence, however most L2 users are at times prone to make utterances that comes off as odd to native L1 speakers of the language. The *final stage* for many L2 users is when their language acquisition starts to stall, which is referred to as fossilization in linguistic literature (Saville-Troike, 2006, p. 21). One can assume that the final stage in English acquisition for the majority of Norwegian pupils takes place when they finish learning English at upper secondary school.

2.2 THE CRITICAL PERIOD AND THE AGE OF LANGUAGE ONSET

The book *Biological Foundations of Language* by Eric Lenneberg published in 1967 was a continuation of the work by Penfield and Roberts on theories about what we today know as the *critical period hypothesis* (CPH), hypothesizing that language acquisition is biologically constrained, starting early in life and ending at puberty (Newport, 2006, p. 737). Lenneberg's work paved the way for extensive research that has been conducted on the field. Today, there exists a broad consensus regarding that children acquire language in a more naturalistic and effortless way than adults, especially phonology and grammar is best acquired at young age (Newport, 2006, p. 737). Studies using fMRI, PET and ERP scanners to monitor brain activity have established a strong left hemisphere activation in L1 and bilingual language production, while L2 activation has been shown to have a more scattered activation pattern, which at times is partially or non-overlapping with the areas activated in L1 production. This scattered activation pattern of late learned L2's shows that they are less lateralized, and that greater variances are found amongst L2 activation patterns, which is then used as an argument to explain why people that learn the L2 at a young age often attain higher proficiency levels than adult learners (Newport, 2006, p. 739). However, some data indicate that the discrepancy between L1 and L2 processing becomes less apparent with increased L2 proficiency (Boxtel, 2005, p. 28).

The CPH has been subject for debate since it was first introduced by Penfield, Roberts and Lenneberg, most significantly the offset of CP, and CP in L2 has characterized the debate (Bialystok & Hakuta, 1999; DeKeyser & Larson-Hall, 2005; Muñoz & Singleton, 2011; Newport, 2006). Of the premises that have been disputed are the measures for ultimate attainment in language acquisition, with especial focus on how L2 speakers perform compared to L1 monolingual speakers of the language. This means that the measures for ultimate attainment of a language traditionally have been native-like competence. This view of ultimate attainment has been tested by several researches; Cook (2010) argues that L2 focus should be on the L2, and not on the comparison to L1 (p. 156). Cook's argument is that ultimate attainment is a measure for L1 acquisition, while there is no reason why L2 acquisition should have the same endpoint of its acquisition (Cook, 2010; Muñoz & Singleton, 2011, p. 4). Munoz and Singleton (2011) share the same view in their review *A critical review of age-related research on L2 ultimate attainment*, where they argue that a looser focus on L2 ultimate attainment in comparison to L1 ultimate attainment could possibly shed more light on the L2

acquisition process. DeKeyser and Larson-Hall (2005) strengthen this argument when they point out that there is yet to be produced examples of pupils that have acquired near native L2 competence from being part of early L2 immersion programmes. The amount of input in L2 teaching is simply not extensive enough to facilitate near native competence (DeKeyser & Larson-Hall, 2005, p. 101), thus, ultimate attainment of an L2 cannot be the same as native like competence.

Closely related to the debate of ultimate attainment is the research question related to age of language acquisition, more precisely the age of which the critical period ends. Traditionally, the CPH explanation has been to acquire a language after the age of puberty requires a greater cognitive effort, making it more difficult to attain high proficiency levels for adolescent and adult learners (Muñoz & Singleton, 2011). The *universal grammar* (UG) take on SLA has traditionally been that before puberty the learner is able to tap into the universal grammar skills, thus the UG and the L1 serve as the basis for SLA to take place. This would then explain the cognitive effort that goes into acquiring a language as an adult (Bialystok & Hakuta, 1999, p. 170). However, recent research shows that there is little to no evidence of an abrupt change around the age of puberty, rather that there is a steady decline throughout puberty and into adulthood before it flattens out and becomes fossilized (Bialystok & Hakuta, 1999; Muñoz & Singleton, 2011, p. 10). DeKeyser and Larson-Hall (2005) conclude that the evident difference when it comes to age of acquisition is that children learn implicitly, but initially at a slower rate, whereas adults rely on explicit language teaching that facilitates a more rapid initial growth in language competence. However, for the acquisition of more complex grammatical rules and structures, young age of language onset will be beneficial in that the learner has a longer period of exposure to the language, thereby causing the child to acquire it implicitly (DeKeyser & Larson-Hall, 2005, pp. 101-104).

In a response to CPH Ellen Bialystok (1997) points out that the main difference in SLA in adults and children is the way the language is conceptualized in the learners. Adults are more prone to extend familiar structures from other languages, while children are more apt to create new categories, which for adults often lead to negative transfer (p. 132; Saville-Troike, 2006). Bialystok goes on to attest that the difference makes it seem that children are better language learners, but in reality it is only evidence of differences in how language is acquired at the different stages in life (Bialystok, 1997). The major difference between young and adult language learners is that adults rely to a much

greater extent on explicit teaching, compared to young learners that are more inclined to implicitly acquire new languages (Cook, 2010). This is then used as an argument against CPH, pointing out that maturational limits and differences in the nature of the input in language acquisition are not the one and same (Bialystok, 1997, p. 132). Thus, according to the data presented by Bialystok one cannot argue that SLA is constrained by a CP. The level of attainment relies on the speaker having learned a natural language as their L1 during his or her childhood, and that they are somehow able to recreate some of the experimental advantages that children enjoy (Bialystok, 1997, pp. 133-134). At the moment there is simply not enough evidence to support the CPH claim in SLA that there is a discontinuity in language acquisition after puberty (Boxtel, 2005, p. 26).

2.3 ACQUIRING A VOCABULARY

In psycholinguistic literature there is a much-used distinction between two forms of acquiring language, with especial focus on vocabulary, namely *incidental* and *intentional language learning*. Hulstijn (2003) points out that there has been a wide variety of meanings that have been assigned to the terms over the past forty years, and that they have been used within different research fields. Despite this diverse use of the terms there is surprisingly little consensus about what exactly goes into the terms, and more prominently which language processes the term involves (p. 349). However, as a whole, one can label *incidental learning* as a term covering the processes where an individual is acquiring language as a by-product of what he is engaged with. In SLA research the term is mostly used in describing the process of learning vocabulary through reading (Hulstijn, 2003, p. 358). On the other hand, *intentional learning* has mostly been assigned to the cognitive processes of rehearsal and memorizing techniques where the intended outcome of the activity is a fixed amount of learned lexical knowledge (Hulstijn, 2003, p. 359). Hulstijn (2003) points out that even though *intentional learning* involves a deliberate attempt to accommodate new information to the lexicon, no learning can take place without attention and notion being paid to the input. Thus, both *intentional* and *incidental language learning* share involvement in the acquisition process which consequently blurs the distinctions between the two (pp. 360-361). In terms of acquiring a vocabulary *incidental learning* involves acquisition of both abstract and factual knowledge, whereas intentional knowledge is restricted to factual knowledge (Hulstijn, 2003, p. 361).

Extensive amounts of research have been done on the first words (lexical items), that are learnt when an individual first starts to acquire a language. A widely recognized model that is used to describe this initial language acquisition is the CDI, *The MacArthur Communicative Development Inventories*. The CDI is a scale that describes in which order words are learnt, and more notably, which categories of words that are acquired first. The CDI scale is comprised of data gathered from extensive parental report forms, where the parents are asked to report which words their children understands as well as which words they produce at different stages in the development, and later the children will be brought in for tests based on their reported vocabulary (Dale & Fenson, 1996; Karmiloff & Karmiloff-Smith, 2002). From the parental reports it has been derived that the first acquired words belongs to the categories of nouns, verbs and adjectives, Dale and Fenson (1996) reports that other vocabulary acquisition tests draw upon the same results, thus one can assume that the CDI is accurate in its descriptions (p. 125). The implications CDI has for this study is that one can draw upon the knowledge of which categories of words that are first learnt, meaning that one can anticipate the participants from the extra input group to perform better at nouns, verbs and adjectives, and maybe showcase a greater competence when it comes to function words like pronouns, prepositions, auxiliaries and conjunctions.

2.4 BECOMING LITERATE IN THE L2

An important aspect in the process of acquiring a language is to become literate in the target language. Cook (2010) explains that learning to read and write changes the way a person thinks, as well as the structuring of the language in the brain. Cain (2009) on the other hand has found that early verbal abilities and vocabulary knowledge are important predictors of later reading comprehension. Furthermore, the writing system of the language affects the processing of the language (Cook, 2010, p. 149). In the case of this study both languages are written in the roman letter system, thereby it is assumed that reading and writing in both languages does not differ in means of how difficult it is to acquire. Norwegian is defined as a language with a relatively shallow orthography, whereas English is defined as having a deeper orthography. The terms *deep* and *shallow* orthography refer to the degree of which the language contains inconsistent correspondences and morphological influences in spelling. As a result, decoding skills are easier to acquire in Norwegian than English, which one would expect to find traces of

in literacy acquisition and consequently vocabulary acquisition (Lervåg & Aukrust, 2010, p. 613; Seymour, Aro, & Erskine, 2003, p. 145).

The first step in the process of learning to read is to map sounds onto letters, a process referred to as *phonological recording*. After mastering this process the child will subsequently be able to access already acquired words from their spoken lexicon (Ziegler & Goswami, 2006). Albeit, phonological recording in languages with deep orthography is more complex than simple sound-letter matching, English in particular is exceptionally inconsistent in both pronunciation and spelling (Ziegler & Goswami, 2006, p. 430). However, when acquiring an L2, this process will in large be predicted by the input the pupils are exposed to, since it is not expected that the pupils have any L2 vocabulary of significance when they start formal training. Ziegler & Goswami (2006) propose that the main difference in the process of becoming literate in a deep orthography compared to a shallow one, is that the learner of the deep orthography has to develop recording strategies for grains of language of varying size, whereas the learner of the shallow orthography only needs a strategy for decoding smaller grains of language, and still be able to read and write without errors. The inconsistency of the deeper orthographies makes it impossible for a learner to just rely on small grains of the language. They need to rely on several recording strategies at the same time, which is believed to put restraints on the acquisition process, thus being the reason why children learning to read English as their L1 does this at a slower rate than most other European children learning to read their L1 (p. 431).

The processes discussed so far have explained the decoding side of becoming literate; the other aspects are the skills that support comprehension (Cain, 2009). Cain (2009) explains that reading comprehension is the process where the reading “results in a representation of the situation described by the text, rather than a description of the text itself: a situation model” (p. 11). Furthermore, just being able to decode and read words does not ensure text comprehension. Proficient text comprehenders are able to combine information from a text with general knowledge, which then leads to inferencing. In this process working memory plays a key role, in that there is a simultaneous processing and storing of information (Cain, 2009). Cain (2009) points out three skills that are believed to influence reading comprehension development; these are *integration and inference*, *comprehension monitoring* and *knowledge and use of structures*. All three of these skills

rely on working memory, thus it is expected that teaching that emphasises these three skills will facilitate positive development of reading comprehension skills (p. 13).

2.5 VOCABULARY AS A DETERMINANT FOR COMPREHENSION SKILLS

A recent Norwegian longitudinal study by Lervåg and Aukrust (2010) on how vocabulary skills can be seen as predictors for comprehension skills in young Norwegian L1 and L2 learners sheds new light on the role of vocabulary in comprehension. They found that the L1 group initially had better reading comprehension skills, and a more rapid growth in these skills. The research also showed that vocabulary and decoding skills were a predictor for initial development of reading comprehension skills as well as future development of these skills (Lervåg & Aukrust, 2010, p. 617). At the beginning of reading, decoding was revealed to be the most dominant component. It is assumed that decoding is easier in consistent orthographies like Norwegian, compared to the more inconsistent orthography of English (Lervåg & Aukrust, 2010, p. 617). The findings indicated that L2 readers are slower in their growth of reading comprehension skills compared to L1 learners, which is assumed to be caused by a stagnation of decoding development that is found in L2 learners when they reach a certain level of competence. This stagnation will then affect the growth level of reading comprehension skills, which causes the L2 learner to rely more on the vocabulary than L1 learners for reading comprehension. Additionally, having a large vocabulary facilitates the skill to acquire new vocabulary through reading (Lervåg & Aukrust, 2010). In sum, Lervåg and Aukrust's (2010) findings showed that vocabulary plays a large part in reading comprehension, and that vocabulary is a better determinant for L2 comprehension skills than it is for L1 comprehension.

A large part of language comprehension is literacy skills, according to Bialystok, Lok and Kwan (2005). Literacy skills are only transferred across languages when both languages are written within the same writing system, which is the case in this study; both Norwegian and English uses the roman letter system. The system which the language is written in is significant in that the acquisition of the various sets of symbols taps onto different sets of cognitive skills (Bialystok et al., 2005, p. 44). Bialystok et al. (2005) list three skills that are crucial for acquiring literary comprehension skills; these are oral proficiency, metalinguistic awareness and general cognitive development (p. 44). A large portion of what makes up oral proficiency skills is the vocabulary, and several

studies on bilingualism have reported superior vocabulary size in monolinguals when compared to bilingual peers. However, the lesser vocabulary size should not be taken as a sign of lesser proficiency in L2, rather just a trait of bilingualism. Metalinguistic awareness refers to how one applies phonological knowledge and awareness in alphabetic reading, as well as recognizing and applying structures from other languages, depending on how closely related the languages are. The third skill required, cognitive development, includes factors such as working memory, specific language factors and orthography. Research on this area indicates that bilinguals have a superior cognitive development, but only in applying it to the two, or more, languages they are proficient in (Bialystok et al., 2005, pp. 44-45).

2.6 THE EFFECTS OF INPUT BASED LANGUAGE TEACHING IN SLA

Studies have shown that non native speakers of a language are prone to develop a “foreigner register” in their L2 inventory, which often is comprised of shorter sentence production, less refined use of prepositions, more frequent use of yes-no questions instead of wh-questions (Gor & Long, 2009). Gor and Long (2009) point out that this “foreigner register” often is grammatically correct, even though it is easily distinguished from language produced by an L1 speaker (2009, p. 445). The development of this “foreigner register” intuitively has to do with the input the language learner is being exposed to while acquiring the language. In pedagogy, a tradition of exposing the children to as much as possible target language has developed. It is based on the idea that when you are learning an L1 you have no other language to fall back to, thus in L2 teaching the children should not rely on their L1 competence to succeed, but learn to access the L2 directly (Cook, 2010, p. 147).

Grosjean (2008) has been a leading figure in developing the term *language mode*, which is a term used to describe how one or more languages are being activated at the same time. The basic concept of *language mode* is that the bilingual operates with one base language and one comparative language (Grosjean, 2008). Meaning that if a Norwegian-English bilingual is speaking to a Norwegian monolingual he or she would be in a Norwegian monolingual mode with Norwegian as the base language, whereas if two Norwegian-English bilinguals are speaking to each other either of the two languages could be the base language, and the speakers would be in a bilingual mode where they can access and use segments from both languages (Grosjean, 2008). In regards to input

based teaching, language mode can be used as an argument for why the teaching of an L2 should be conducted in the target language, especially at lower levels. If the L1 is the main spoken language when trying to teach the L2, the pupils will most likely be in a monolingual language mode, making the shift into the L2 more abrupt and less likely to be successful.

In the wake of research on the use of target language in pedagogy, a teaching tradition referred to as input based learning has emerged (Barcroft, 2004). In this paper the term input based teaching will be used to describe teaching that emphasises naturalistic use of the target language; oral and practical use of high frequency language through songs, talking about objects and reading aloud. In short, focusing on exposing the pupils to authentic language input that encourages them to use the language actively. The general conception about SLA in the psycholinguistic field is that input and exposure to the target language is crucial for successful acquisition to take place (Dahl & Vulchanova, 2014). A study by Hatch on L1 acquisition, published in 1978, established that participation in conversations facilitate growth of linguistic competence. According to Hatch the quality and diversity of the input that the learner is exposed to in large determines the course of development in the learner (Ellis, 1985, p. 20). Matthew J. Traxler (2012) acknowledges that teaching techniques that minimize L1 activation in the process of L2 activation will have greater chances of facilitating L2 learning, as the learner does not have to access L1 label from his mental lexicon in order access the corresponding L2 label (2012, p. 437). Based on Hatch's findings, Traxler's view on teaching an L2, and the body of research on L1 acquisition one can assume that the same processes apply in SLA. As Dahl & Vulchanova put it "[a]mount and quality of input are undoubtedly crucial factors in SLA" (2014, p. 2).

A study by Vibeke Grøver Aukrust (2007) followed a group of Turkish-Norwegian children learning Norwegian as their L2 throughout the two last years of kindergarten and the subsequent first year at school. The findings of this study showed a linkage between teacher talk in kindergarten and vocabulary results in the first graders. However, it was shown that it was not the quantity of exposure that mattered, but the quality of the teacher talk addressed to the children (Aukrust, 2007, pp. 32-33). The teachers that used an extensive vocabulary also introduced more complex words and structures, as well as not demanding complex answers from the L2 speakers. This lends reason to believe that non-threatening learning environments, in combination with high

quality input, facilitates positive learning outcome in L2 learners with limited L2 capacity (Aukrust, 2007, p. 34). Other factors that were shown to play a role were the parental education levels, a high correlation between L2 vocabulary test results and parental education were shown (Aukrust, 2007, p. 33).

Anne Dahl and Mila Vulchanova (2014) tested how extra English input in Norwegian first graders affected their English acquisition. The formal amount of English lessons was not increased, but the density of English exposure was increased through speaking English during circle time and classroom management. This study found that the first graders that had been exposed to naturalistic and frequent use of English outperformed counterparts that had received conventional English instructions. However, this study tested for receptive vocabulary, which is known to be predictive of comprehension, but productive vocabulary was not tested (Dahl & Vulchanova, 2014, p. 758).

The follow-up study by Tonje Sivertzen (2013) retested the receptive vocabulary of the same group of pupils two years after the discontinuation of the extra input focus in the teaching. Her findings indicated that even though there was a difference in raw scores between the groups, there was no statistically significant difference (2013). When compared to the first grade results that showed a significant difference, it became evident that the vocabulary growth in the extra English group and in the control group were parallel, explaining the small, but not significant difference found by Sivertzen (Sivertzen, 2013, p. 37).

3.0 METHOD

The methods that are applied in the data collection for this study are deductive methods that obtain quantitative data, by that both the reading comprehension test and the listening comprehension test score the performance of the participants either right or wrong. However, the questionnaires, that the parents had to fill out, had some qualitative questions where the parents were asked to elaborate if they had any extra information that was not covered by the qualitative questions where they had to tick of boxes. Postholm and Jackobsen (2011) acknowledge that quantitative data gathering is the most effective method in means of establishing general information about the participants in the study, but they also point out that qualitative and quantitative methods should not be viewed as opposing alternatives, and that research often is conducted within both scopes (p. 42).

This chapter is divided into the following sections; 3.1-3.1.2 presents the groups of participants; 3.2 presents background information on the participant schools; 3.3 describes the procedures of the data collection; 3.4 discusses the consent form and questionnaire; 3.5-3.5.2 describes the reading comprehension test and the listening comprehension test.

3.1 PARTICIPANTS

For this study, the body of test subjects were made up of Norwegian fifth graders that all lived in one of Norway's largest cities. All of the participants came from neighbourhoods of equal sociocultural status, which is reflected in the mean scores of the schools on the national test in which all Norwegian pupils have to partake. The cumulative mean score from 2007-2013 showed that two of the three schools scored slightly above average for the municipality mean score, whereas one was slightly below (Trondheimsskolen, 2013). Based on the mean scores of the participating schools one can estimate that there were no formal differences in how the teaching was done at the test schools.

The reason for the selection of this specific age group is that two of the classes in this study have been subject for two previous studies (Dahl, 2014; Dahl & Vulchanova, 2014; Sivertzen, 2013), making this the latest addition of data in a longitudinal study. A total of 88 pupils were asked to participate in the study, but since participation was not mandatory 55 brought back filled in forms. The total amount of participants that were

included in the project was 39, after omitting those who reported Norwegian as their L2, bilinguals, and pupils that had not been part of the classes during Dahl and Vulchanova's study (2014).

3.1.1 EXTRA ENGLISH INPUT CONDITIONS GROUP

The *Extra English Input Conditions Group* (EI) was made up of pupils that participated in the study of Dahl & Vulchanova four years ago, and in Sivertzen's MA thesis that was conducted about eight months before the data collection for this study.

In the process of selecting test subjects a total of 53 consent forms and questionnaires were handed out. It was ascertained that all the participants had attended the same school since first grade, based on the fact that Dahl and Vulchanova's (2014) study was conducted during first grade. Of the 53 distributed forms 36 pupils brought approved consent forms and filled in questionnaires. Of these, 4 had to be excluded because they were bilingual, in addition to 8 that had to be excluded because they had not attended first grade at this school. The group consisted of 24 selected candidates, 10 male and 14 female.

3.1.2 NORMAL ENGLISH INPUT CONDITIONS GROUP

The body of the *normal English input conditions group* (NI) was made up of two fifth grade classes from two primary schools that reside in areas that are of similar socio-cultural status. The reason why the control group was made up of pupils from two different schools was because of a very low turnout of approved consent forms from the school that was first approached, thus an additional school was added to the project in order to compensate for an uneven number of participants in the two groups. For all future references these two schools will be referred to as a single group, NI, since no sign of deviation from standardized English teaching were reported in neither of the schools.

In total 45 consent forms and questionnaires were handed out to the two classes, of which 19 were responded positively. Of these 4 had to be excluded because they reported that they were bilingual. The final group consisted of 15 selected candidates, 7 female and 8 male.

3.2 BACKGROUND INFORMATION ON THE CHOSEN SCHOOLS

The background for this study was the work that was initiated by Anne Dahl and Mila Vulchanova four years ago on the EI group, as well as Tonje Sivertzen's retesting of EI for her MA thesis last year.

Dahl and Vulchanova's study (2014) set out to investigate how extra English input during the first year of teaching would affect the growth in both English vocabulary as well as comprehension in first graders. In this study the teachers teaching the EI group used English more extensively both during English lectures and outside the regular English classroom contexts. When used outside of the English lectures the language was used for simple instructions, morning meetings, songs and reading aloud. The teachers in the NI group were told to teach English just as they were used to, with no instructions in regards of which language that had to be spoken during the lectures (Dahl & Vulchanova, 2014, p. 3). In means of approaches to teaching, the two schools differed in that the EI group had a more *input based teaching*. This is a more practical approach to language teaching typical of Norwegian English teaching. The input based approach applies more use of pictures, stories that are read by the teacher and working with physical objects (Dahl & Vulchanova, 2014, p. 4). The results from Dahl and Vulchanova's study showed that during the short period of eight months the pupils of the EI group had developed a larger receptive vocabulary than the NI group (2014, p. 11), as well as showing a significant superiority in sentence comprehension (Dahl, 2014).

For her MA thesis, Tonje Gauslaa Sivertzen (2013) retested the EI group, three years after they had partaken in Dahl and Vulchanova's project. The teachers at the EI school had stopped focusing on extra English input after the second grade, which is why Sivertzen retested the vocabulary size, to look for long lasting linguistic benefits. For Sivertzen's study the NI group was not the same as in Dahl and Vulchanova's, but the teacher reported that they had been exposed to instruction typical of Norwegian schools, thus it proved to be equal in means of functioning as a control group (Sivertzen, 2013). What Sivertzen found in her results was that the significant difference in vocabulary size that was documented three years earlier, no longer was significant. The EE group showed higher mean raw scores, but factorial ANOVA analysis showed that the difference in growth was not statistically significant (2013, p. 42).

For this study the body of the EI group is made up by pupils from the same classes that made up the EI group for both Dahl and Vulchanova's and Sivertzen's studies.

3.3 PROCEDURE

Before the data collection, all of the participants were given oral information about the project as well as getting written information that was passed on to their parents. To be able to conduct the study I relied on parental consent, because the participants were under the age of 18. Also, since participation in this project by no means was mandatory for the pupils, it meant that the parents and children could reserve themselves from partaking. This in turn led to the need of two control groups because the first NI class that was approached only produced eight approved consent forms. The experience from Sivertzen's study was that it was very time consuming when she first had to pass on consent forms, then questionnaires to the ones that responded positively to the consent form. Therefore, in an attempt to make the process of the information gathering as simple as possible, both the parental consent form and the questionnaire were handed out together. All of the filled in forms were brought back to the school in sealed envelopes, in order to keep all the information in the forms as anonymous as possible, as well as keeping the information in the forms hidden from third party eyes.

3.4 CONSENT FORMS AND QUESTIONNAIRES

In addition to asking for parental consent for participation in the project, the form included a questionnaire that was designed to map out each participants language environment outside of the school (see appendix 1 and 2).

The questionnaire was split into three parts; personal background, language background, and other factors in language learning. Of these, *language background* made up the main body of the questionnaire. The main function of the *language background* part was to rule out those who were bilingual, or spoke Norwegian as their L2. Secondly, the questionnaires could be used to cross reference the results of the tests if there was an individual result that significantly deviated from the other participants.

3.5 TESTS

The decision to have two tests was based on the fact that Sivertzen's (2013) study found little to no long lasting significant effects on growth of vocabulary, thus by running two

tests when testing for comprehension one could expect to get a more nuanced result that might reveal otherwise overlooked aspects. Unfortunately, all of the test subjects included in this study did not participate in both the tests, since the testing had to be done over a period of several days, making it impossible to compensate for some subjects being absent.

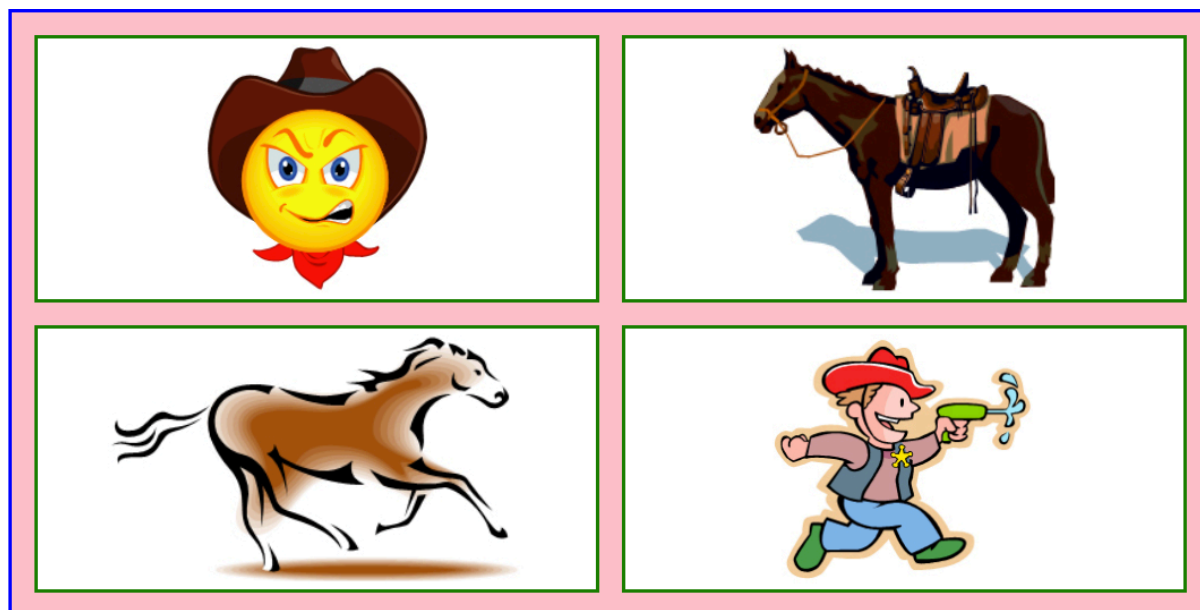
3.5.1 LISTENING COMPREHENSION TEST

The *listening comprehension test* used for this project, was first developed by Anne Dahl, Mila Vulchanova and programmed by Bjørn Grønnesby and colleagues, and used as one of the tests that was conducted in their project on extra English input (Dahl, 2014; Dahl & Vulchanova, 2014).

During the test the pupils were situated in front of a computer where they were presented with frames consisting of four pictures. The participants would then press a play button on the screen to initiate a voice recording of a sentence that corresponded with one of the pictures on the screen. In every frame there was one picture that corresponded to the recorded sentence, two that were semi-related and one that was non-related¹. After hearing the sentence the participants were instructed to click on the picture that he or she thought was the most related to what he or she just heard. In Anne Dahl's study the test had 30 frames, and the voice recordings was a mix of recordings from a female American English native speaker and a female British English native speaker. For this project the test was extended with 23 extra frames with corresponding sentences that were aimed at testing a more complex English. These sentences were based on glossary lists and texts from the fourth grade text book *Steps: Textbook 4* (Munden, Musk, & Wessman, 2007). New recordings of the original and the new sentences were made; a male native speaker of British English with a neutral dialect read the sentences for the recordings. The choice of replacing the original recordings were based on the belief that it could cause disturbance in the test results if there were both male and female voices. The original 30 frames were kept for several reasons. If a participant performed poorly on the tasks intended for first graders it might be indicative of specific learning difficulties. Also, presenting the participants with tasks that were likely to be perceived as simple would most likely invoke a sense of achievement, which was anticipated to

¹ The following picture frames had one correct picture and three semi-related pictures: T37, T39, T46, T51, T52 (appendix 3).

make them less intimidated by the frames that were more challenging. All 53 picture frames and corresponding sentences are included in appendix 3.



Example of a picture frame from the original test by Dahl and Vulchanova.

At all the schools the test was conducted in a separate room from the rest of the class. The participants were tested one at a time because of two reasons. Firstly, the pupils were much more likely to not be distracted by noise from fellow pupils or adopt other pupils' attitude towards the test, if tested all at once. Secondly, the test and the results were stored at one of the NTNU's servers; the test itself could be run in any Internet browser, however, in order for it to log the results, all 53 answers had to be answered. Because of this, the material constraint was much smaller when testing one pupil at a time. The risk of malfunctioning computers would be very high if an entire class were to do the test at the same time, as well as putting larger constraint on the network, risking that some data might not be stored correctly.

3.5.2 READING COMPREHENSION TEST

The format of the reading test is one that most Norwegian pupils are familiar with. First they were presented with a short story they had to read, and then they answered eight questions that referred to details from the short story. For this test the short story was borrowed from a webpage called *Learn English Kids*, which is a resource page for

English teaching set up by The British Council (Learn English Kids, n.d.)². Both the short story and the questions can be found in Appendix 4.

In contrast to the listening comprehension test, the reading test was done collectively. All the participants were given a copy of the short story, and instructed to read the story while paying attention to details. After they were finished reading they were given the eight questions to answer. When conducting the reading test in the various classes it became clear that there was a lot of tension associated with the test situation. Therefore all the participants were instructed that they could write in Norwegian if they could not remember the correct English word or were unsure about the spelling. When the test results were encoded into SPSS correct English answers were given the value label 2, correct answers in Norwegian 1, and incorrect answers 0. Correctness in spelling was not scored.

It is expected that this test will in large tap onto the decoding part of reading skills, as short answer tests and cloze tests rely heavily on the reader being able to decode what is asked for. In tests where the reader is presented with longer passages of text, the reader has more words to base his or her inference on, thus it was easier to come up with the correct answer (Lervåg & Aukrust, 2010, p. 618). The questions in the reading comprehension test (appendix 4) can be divided into two categories; one tested the decoding skills, where the participants were asked to list factual objects from the story; the other aimed to test comprehension by making the participants draw on inferences and facts from the text in order to produce correct answers.

² <https://learnenglishkids.britishcouncil.org/en/short-stories/the-voyage-the-animal-orchestra>

4.0 ANALYSIS & RESULTS

This chapter is divided into six subchapters. *4.1 MANOVA* presents the main analysis for this study where the scores from both of the tests were fact analysed. *4.2 ANOVA* presents two analyses, one of each test. *4.3 Listening Comprehension Test* highlights the tasks where the largest discrepancy between the two groups was found. *4.4 Reading Comprehension Test* lays out the frequencies of distribution for each question of the reading comprehension test. *4.5 Data From the Questionnaires* lays out the most relevant data from the questionnaire. *4.6 Limitations* points out some of the limitations of the data set for this study.

4.1 MANOVA

The data collected from both tests were run through a *Multivariate Analysis of Variance* (MANOVA) with the statistical analysis program SPSS. MANOVA is a complex statistic tool that makes it possible to analyse multiple dependent variables together (Leech, Barrett, & Morgan, 2005, p. 162). Leech et al. (2005) explains that data that goes into a MANOVA has to be related conceptually, but one should pay attention to the level of correlation, too low correlation will mean that there is no reason to compare the data in a MANOVA, whereas too high correlation means you are risking multicollinearity. The *General Linear Model* provides the researcher with the multivariate F that is a linear combination of variables that maximally distinguishes the groups, which is the multivariate result in MANOVA (p. 162).

In this study the MANOVA analysis was used to generate a statistical representation of how the two groups, EI and NI, compared to one another. The dependent variables in the analysis were the individual total scores of the two tests, while the fixed factors were the two groups, EI and NI. The amount of participants that were included in the MANOVA was reduced to 19 EI participants and 12 NI participants, because the MANOVA requires two dependent variables per participant. Thereby excluding the participants that had only partaken in one of the tests, in this case excluding 5 EI participant and 3 NI participants.

	Group	Mean	Std. Deviation	N
Reading test total score	NI group	4.0833	5.08935	12
	EI group	8.8421	4.58576	19
	Total	7.0000	5.25991	31
Listening test total score	NI group	96.7500	4.69284	12
	EI group	101.1053	3.21273	19
	Total	99.4194	4.34951	31

Table 1: Descriptive statistics. Here the descriptive mean scores are shown. The maximal score was 16 for the reading test and 106 for the listening comprehension test. Standard Deviation indicates how much deviation one can expect to find in the results of a single individual's score.

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.999	16039.867 ^b	2.000	28.000	.000
	Wilks' Lambda	.001	16039.867 ^b	2.000	28.000	.000
	Hotelling's Trace	1145.705	16039.867 ^b	2.000	28.000	.000
	Roy's Largest Root	1145.705	16039.867 ^b	2.000	28.000	.000
Group	Pillai's Trace	.260	4.913 ^b	2.000	28.000	.015
	Wilks' Lambda	.740	4.913 ^b	2.000	28.000	.015
	Hotelling's Trace	.351	4.913 ^b	2.000	28.000	.015
	Roy's Largest Root	.351	4.913 ^b	2.000	28.000	.015

Table 2: Multivariate tests output.

Source	Dependent variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected model	Reading test total	166.557 ^a	1	166.557	7.280	.011
	Listening test total	139.509 ^b	1	139.509	9.452	.005
Intercept	Reading test total	1228.751	1	1228.751	53.710	.000
	Listening test total	287917.702	1	287917.702	19506.643	.000
Group	Reading test total	166.557	1	166.557	7.280	.011
	Listening test total	139.509	1	139.509	9.452	.005
Error	Reading test total	663.443	29	22.877		
	Listening test total	428.039	29	14.760		
Total	Reading test total	2349.000	31			
	Listening test total	306978.000	31			
Corrected total	Reading test total	830.000	30			
	Listening test total	567.548	30			

Table 3: Tests of Between-Subjects Effects

Table 2 displays that there is significant output of the MANOVA, regardless of which of the four tests one chooses to apply. Leech et al. point out that Wilks' Lambda produces a good representation when assumptions are met, in this case $F(2,28) = 4.91$, $p = .015$ (2005, p. 167). With a p value of .015 one can reject the *null hypothesis of no difference* ($p < .05$), and assume that there is a significant difference between the two groups (Leech et al., 2005, p. 54). The significant F is an indication of a significant difference between the linear combination of the two dependent variables, as can be seen in table 3 in the *group* column (Leech et al., 2005, p. 167). Table 3 displays the between-subjects effects, which shows the two ANOVAs (analysis of variance) that go into the MANOVA. Table 3 shows that the between-subjects effects on the reading test was $F(1,29) = 7.28$, $p = .011$, whereas the listening comprehension test was $F(1,29) = 9.45$, $p = .005$. However, the between-subject effects in table 3 does not paint an exact picture of the gathered data, since a total of 8 participants were excluded from the analysis because they had not partaken in both tests.

4.2 ANOVA

The output of an *analysis of variance* is a linear analysis of more than one independent variable that calculates the correlation of the two or more variables. Factorial ANOVAs are typically used when the data material consist of a low number of variables, in analyses where two groups are compared with one independent variable, which is the case in this study; they are called single factor designs (Leech et al., 2005, p. 129)

In the ANOVA of the listening comprehension test 12 participants from the NI group and 21 from the EI group were included. The number of NI participants is the same as in the MANOVA, but there are two more from the EI group in this analysis than in the MANOVA.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected model	147.360 ^a	1	147.360	10.088	.003
Intercept	299052.088	1	299052.088	20473.004	.000
Group	147.360	1	147.360	10.088	.003
Error	452.821	31	14.607		
Total	327607.000	33			
Corrected total	600.182	32			

Table 4: ANOVA output of the listening comprehension test.

The result the ANOVA of the listening comprehension test, as can be seen in table 4, shows that the result, $F(1.31)=10.1$, $P=.003$, is significant. This was the expected outcome, considering that the NI group remains the same while the EI group got expanded, thus making the group sizes more uneven.

In the ANOVA of the reading test, 15 NI and 22 EI participants were included.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected model	148.380 ^a	1	148.380	6.371	.016
Intercept	1510.109	1	1510.109	64.836	.000
Group	148.380	1	148.380	6.371	.016
Error	815.188	35	23.291		
Total	2721.000	37			
Corrected total	963.568	36			

Table 5: ANOVA output of the reading comprehension test.

The output of the ANOVA of the reading comprehension test, seen in table 5, does show slightly higher correlation, $F(1.35)=6.38$, $P=.016$. The outcome is significant, but the elevation of the p value, compared to the MANOVA between-subjects effect (table 3), is indicative of the difference in performance between the two groups being smaller when the body of NI participants is elevated. However, the result is significant $p < .05$, thereby one can refute the null hypothesis, and make the claim that there is a significant difference in performance between the two groups.

4.3 LISTENING COMPREHENSION TEST

Thus far the total scores have been presented and discussed; this section will have a more pragmatic approach and present some data on the tasks where the two groups had semi correct and incorrect answers. By highlighting the tasks with the most deviance one might shed more light on specific areas where the two groups deviate from each other. It was expected prior to the testing that both groups would score equally well on the first 30 tasks since these were based on first grade material and the level of inference required to comprehend these should be well acquired for fifth graders. This expectation was more or less met; there were quite a few questions where one or two participants had chosen the semi correct alternative. The only exception was task 16 (appendix 3), where 33.3 percent of the NI group and 14.3 percent of the EI group chose the semi correct answer (appendix 5).

Of the 23 tasks that were added to the original test, there was a couple that were expected to be more difficult for the test participants, common for all of these was that they either had more complex sentence structure, required the participant to do inferencing, or tested prepositions where small nuances in the pictures separated the right and wrong option.

The first of these was task 34 (appendix 3) where it was expected that the riddle would confuse the participant, thus creating impoverished comprehension. This was also the case, of the NI group only 8.3 percent answered the correct option, 66.7 semi correct and 25 percent wrong; the EI group had 52.4 percent correct, 42.9 percent semi correct and 4.8 percent incorrect (appendix 5).

Task 41 (appendix 3) had an *expletive there* construction, which was expected to be challenging. On this task 50 percent of the NI group chose correctly, and 50 percent a semi-correct answer, whereas in the EI group 38.1 percent chose correctly, 57.1 percent semi correct and 4.8 percent chose the incorrect option (appendix 5).

Task 44 and 45 (appendix 3) both tested prepositions. On task 44 the NI group had 16.7 percent correct, 66.7 percent semi correct, and 16.7 percent incorrect; the EI group had 47.6 percent correct, 47.6 percent semi correct, and 4.8 percent incorrect. On task 45 the NI group scored 41.7 percent correct and 58.3 percent semi correct; the EI group scored 57.1 percent correct and 42.9 percent semi correct (appendix 5).

However, the most surprising result was the amount of semi correct answers on tasks 51 and 52 that had sentences with negation (appendix 3). On task 51 the NI group scored 33.3 percent correct and 66.7 percent semi correct, whereas the EI group had 61.9 percent correct and 38.1 percent semi correct. For task 52 the NI group had a distribution of 50 percent correct and semi correct, while the EI group had 61.9 percent correct and 38.1 percent semi correct (appendix 5). But it was not the distribution of semi correct answers that was the most surprising. When conducting the tests a large portion of the participants made remarks about the pictures, claiming that all of the pictures showed pets when asked to single out the one that did not. Therefore the results might be coloured by the lack of concept about livestock versus pets instead of being an indication of negation causing disturbance in the inference.

4.4 READING COMPREHENSION TEST

In sections 4.1 and 4.2 the results of the reading comprehension tests were discussed in terms of total performance of the two respective groups. However, in addition to the total score one should also take the descriptive statistics of each question into consideration. The following tables will present the frequencies of distribution of answers of each question.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Wrong	31.8	40.9	63.6	40.9	18.2	72.7	54.5	50
Correct Norwegian	-	-	-	-	-	-	-	-
Correct English	68.2	59.1	36.4	59.1	81.8	27.3	45.5	50

Table 6: This table shows the frequencies of distribution of answers, by the 22 EI participants, at the reading comprehension test.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Wrong	66.7	53.3	80	66.7	53.3	86.7	80	86.7
Correct Norwegian	-	-	-	-	-	-	6.7	-
Correct English	33.3	46.7	20	33.3	46.7	13.3	13.3	13.3

Table 7: This table shows the frequencies of distribution of answers, by the 15 NI participants, at the reading comprehension test.

The three horizontal columns in tables 6 and 7 represent the three grades of which the answers were categorized into when transcribed into SPSS. The vertical columns represent each of the eight questions. It was anticipated that the alternative of writing Norwegian answers would give more pupils the option to further display their comprehension skills, but as seen in the tables only one participant of the NI group answered one question in Norwegian. When comparing the two tables, the distribution of answers clearly shows that the EI outperformed the NI group on all eight questions. Although, one should keep in mind that the variance within each group is high, as seen by the high value of the standard deviation (table 1), thereby masking high performances by NI participants.

4.5 DATA FROM THE QUESTIONNAIRES

The data collected from the questionnaires were filled out by the pupils and the parents was transcribed into SPSS where the results were produced into frequency tables. Because of the size difference of the EI and NI group all the data from the questionnaires

will be presented in percentage to give a better picture of how the two groups compare on the different questions. A selection of the most relevant questions will be discussed, as parts of the questionnaire was intended to unveil discrepancies rather than gather data.

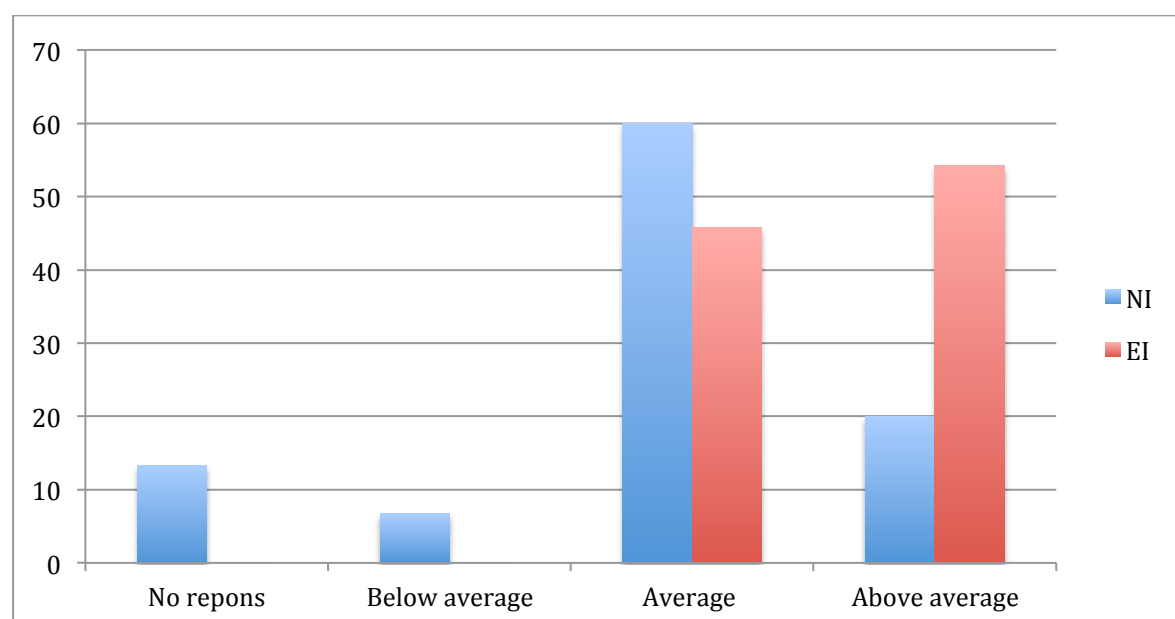


Table 8: Self-assessment of overall English skills displayed in percentage.

Table 8 shows how the two groups assessed themselves on overall English skills. As can be seen from the graph, the two groups assessed their skills quite differently. Of the EI group, 45.8 percent assessed themselves as average while 54.2 percent assessed themselves as above average. Whereas in the NI group 60 percent assessed themselves as average, 20 percent as above average, 6.7 percent as below average and 13.3 did not respond on that particular question. The ones that had not responded had scored themselves as average and below average on the other self-assessment questions, so it is likely that the lack of response on this particular question has to with misinterpretations of the form rather than low self-assessment.

On the question about listening skills in English, the two groups had close to equal scores, which can be indicative of that the amount of spoken English that the participants are exposed to is similar.

However, on the question about how often they read texts in English the results are more divided. In the EI group, 75 percent reports that they read English on a weekly basis, whereas 16.7 percent read English every day, and 4.2 percent read several times a month, and 4,2 percent occasionally read texts in English. Meanwhile in the NI group

only 26.7 percent reports that they read on a weekly basis, 13.3 percent read at a daily basis, 26.7 percent read several times a month and 33.3 percent report that they only read occasionally.

On the question regarding how often they read English texts that are not part of homework assignments the EI group reported that 12.5 percent do so on a daily basis, 25 percent on a weekly basis, 12.5 percent on a monthly basis, 37.5 percent occasionally, and 12.5 percent reported never reading any extra curricular texts. Amongst the NI group 13.3 percent reported reading on a daily basis, 6.7 percent on a weekly basis, 53.3 percent occasionally, and 26.7 percent reported that they never read English texts that are not part of a homework assignment.

When asked which subtitle option they usually choose when watching films, 93.3 percent of the NI group reported that they used Norwegian subtitles, whereas 6.7 percent (1 participant) used English subtitles. In the EI group 83.3 percent reported that they used Norwegian subtitles, while the remaining 16.7 percent reported that they did not use subtitles.

4.6 LIMITATIONS

Of the limitations in this data, it is the amount of participants that is the biggest weakness, especially the discrepancy in size amongst the two groups. If the groups had been more equal in size and consisted of a larger number of participants it would have produced more robust data that could have shed more detailed information. This is especially reflected in the high value of the standard deviation in the descriptive statistics (table 1), which is indication of the data being coarse, thus one should be careful in drawing conclusions that are too decisive.

Another limitation is the questionnaire, where the parents and pupils are asked to fill out the form together. It is impossible to check whether this guideline is met, or if the parent(s) have filled out the form singlehandedly. This is a risk one has to take when using questionnaires, however, the information reported in the questionnaires does for the most part correlate with the results of the two tests. This then makes the data from the questionnaires more reliable, as one may assume that the parents have good knowledge about their children's language proficiency. Still, the data from the questionnaire will be

used with caution, and for the most part serve as perspectives rather than weighing arguments.

5.0 DISCUSSION

In this chapter, the results of the tests will be discussed in light of the theory presented in chapter 2, in an attempt to explain the findings, and to discuss how they correspond with current theories of SLA research. This chapter is divided into three parts; 5.1 *The implications of the MANOVA and ANOVA results* discusses how we can interpret the overall performance of the EI group; 5.2 *Listening comprehension skills* presents a in-depth discussion of the listening comprehension skills; 5.3 *Literacy and reading comprehension* discusses in detail the relationship between vocabulary and comprehension.

5.1 THE IMPLICATIONS OF THE MANOVA AND ANOVA RESULTS

The results of the MANOVA analysis reveal interesting findings. The output of the analysis reveals that the EI group outperformed the NI group significantly when the results of both reading test and the listening comprehension test were analysed, shown in table 1,2 and 3 (chapter 4.1). In the coming section will argue and hypothesize why this difference exists, and more importantly what we can learn from it.

To put the findings of this study to perspective we are first going to take a brief overview of what has already been found in the previous studies on the EI group. Dahl and Vulchanova (2014) documented that during the period of eight months the EI pupils drastically expanded their L2 receptive vocabulary, and they concluded that it had to be a product of the quality of the input and the focus on the use of the target language during English lessons (2014). Dahl (2014) also ran the listening comprehension test at the end of the eight month period, which revealed that there was a significant difference ($p = .001$) between the two groups in sentence comprehension (p. 14). Tonje Sivertzen (2013) retested the receptive vocabulary of the EI group two years after the discontinuation of the extra input program, and found that the significant difference in receptive vocabulary size had levelled out. The EI group had higher raw scores, but no significant difference was found when analysed by ANOVA.

In light of the previous studies on the EI group, and the results of this project, it becomes evident that the EI group has gained language competence that does not necessarily show on receptive vocabulary tests, but becomes apparent when testing for listening comprehension and reading comprehension. Therefore, one can argue from the

results of the MANOVA and the ANOVA that the general English proficiency of the EI group is higher than that of the NI group, even though Sivertzen found no statistical difference in receptive vocabulary size. One possible explanation for the difference that is found between the two groups is that the EI group had a higher quality of input during their initial stages of acquiring English, which according to CPH would mean that they acquired the language more effortlessly than their counterparts, the NI, who had their spurt in acquisition of receptive vocabulary at a later stage (Sivertzen, 2013). Sivertzen's hypotheses suggested that the EI group hit their language spurt in English at an earlier stage due to the extra input, and that the corresponding age equivalent chart of the PPVT-4 test could explain the lack of significant difference in receptive vocabulary size (Sivertzen, 2013). However, what Sivertzen's hypothesis does not fully take into account is the cognitive effort that goes into the language acquisition, and it would seem from the results of the MANOVA and ANOVA analysis that the age at which the language spurt took place does have an impact on comprehension skills and general language proficiency, even though the extra input program was discontinued. Thus, according to Bialystok's argument about how there is not enough evidence to support the claim that SLA is constrained by a CP; one can argue that it seems like the naturalistic input the EI group received facilitated learning that they still are reaping positive benefits from. Therefore, it will be hypothesized that the naturalistic input based teaching made the language acquisition for the EI group more implicit, thereby making a solid foundation for future acquisition of the L2, and despite the similarities between the receptive vocabulary size of the two groups, it is hypothesized that the differences in comprehension skills are caused by the more explicit early English teaching the NI group received (Bialystok, 1997; Boxtel, 2005; Cook, 2010; DeKeyser & Larson-Hall, 2005).

A striking find in the data gathered from the questionnaire is that the self-assessment of the participants' English skills correlates well with the results of the MANOVA and ANOVA analysis. Two interpretations of this correlation will be addressed. First, the high percentage of EI participants that have scored themselves as above average might indicate that the EI group has a more developed meta-cognitive awareness that makes them able to assess their own performance and skills compared to their peers. Meaning that they have passed the threshold of the *intermediate stage*, and possess a more developed *interlanguage* than their peers in the NI group (Saville-Troike, 2006). Secondly, the questionnaires do indicate that Norwegian parents generally have a

good knowledge and insight of their children's language skills, and that there is a pattern in terms of how both the parents and the children assess their language skills that points towards the EI being generally more proficient in English.

The test results, seen in the light of *language mode theory* as presented by Grosjean (2008), could be interpreted in two ways due to the fact that the instructions for both tests were given in Norwegian. One interpretation would indicate that the EI group has a higher proficiency in the L2 because they are able to switch modes more rapidly, thus achieving better scores in the test. The other interpretation would be that the NI group performed worse on the test because the Norwegian instructions have set them in a Norwegian monolingual mode, making the switch over to testing even more difficult. However, neither of these interpretations mutually exclude the other, and since it is reported that most English teaching in Norwegian lower grades takes place in Norwegian (Dahl & Vulchanova, 2014), it would be safe to assume that the NI group is used to accessing the L2 with Norwegian as the base language, thus the test setting probably seemed more familiar when done in Norwegian.

5.2 LISTENING COMPREHENSION SKILLS

The EI group outperformed the NI group significantly on the listening comprehension test, as previously shown in the results of both the MANOVA and the ANOVA, and in detail in chapter 4.3. This section will try to shed some light on how these findings correspond with current theories in this field.

When Dahl (2014) tested the EI group, with the listening comprehension test after they had been exposed to the extra English input for eight months, she found that they scored significantly better than the NI input group. Thus, we know that the EI group had attained better comprehension skills from being exposed to naturalistic input of high quality (2014). The findings of this study also imply that the EI group was better at comprehending spoken English than their NI peers, although at a slightly lower level of correlation. The initial conclusion that can be drawn from the listening comprehension test of this study is that the EI group have retained much of the advantage that they acquired during their first year at school, even though their growth in receptive vocabulary has levelled out compared to the NI peers (Sivertzen, 2013).

The major difference between the receptive vocabulary test and the listening comprehension test is that the participants have to infer and comprehend complete sentences and match them to pictures compared to matching single words with pictures. Enhanced skills in sentence comprehension, and the ability to single out the meaning bearing components of a sentence, are skills believed to be acquired by exposure to spoken language. In other words, exposure to oral use of the target language may have led to incidental language learning. Moreover, the pupils have grown familiar with being addressed in English on subjects that does not necessarily involve topics where explicit learning is the intended outcome (Hulstijn, 2003). Therefore, what might have happened during the first year at school is that the EI pupils acquired both abstract and factual competence about English, compared to the NI group that is likely to have received most of their early English teaching in Norwegian, consequently making the teaching explicit.

Furthermore, the English teacher that taught the EI group was a native speaker of English, giving reason to believe that this group was not exposed to “foreigner language”, as described by Gor and Long (2009). It is not the intention to undermine the work of Norwegian teachers, but studies have reported that very little of the time spent on teaching English is actually done in the target language (Dahl & Vulchanova, 2014, p. 2). But it has to be emphasised that any exposure, be it from a native speaker or not, will expose the pupils how the language is structured, which gives the learner far more insight into how the language functions compared to cramming single words taken out of context.

Another contributing factor that could possibly shed some light on why the EI group outperforms the NI group is the fact that they most likely acquired the high frequency words of English at a much earlier stage than the NI group. We know from the work that has been done on the CDI that the first words that are acquired by L1 speakers of English are meaning bearing words, and that function words are learned later (Dale & Fenson, 1996; Karmiloff & Karmiloff-Smith, 2002). It is likely that the teaching of meaning bearing words like verbs, nouns and adjectives are the first words that the NI group were taught during their first year in school, but the difference would be that if the spoken language in the classroom was Norwegian, they would not have been exposed at all to function words that are used in combination with meaning bearing words. Thereby prolonging the period leading up to the stage where the pupils start acquiring function words. When looking more closely at the results of some of the tasks it becomes apparent

that the EI group did indeed outperform the NI group on tasks testing for function words and complex sentence structure (see chapter 4.3).

5.3 LITERACY AND READING COMPREHENSION

The results of the reading comprehension test did reveal that there was a large difference in reading comprehension skills between the two groups. This can be seen in table 1 (chapter 4.1), where the descriptive statistics of the participants that were included in the MANOVA show that the mean score of the EI group was more than twice the NI score. Although this reveals a discrepancy, one should also note that the standard deviation was higher for the NI group, which indicates that there was a larger span between the lowest and the highest scores in this group. In the between-subjects effect output of the MANOVA, in table 3 (chapter 4.1), the significance is $P = .011$, but in table 5 (chapter 4.2) of the ANOVA the significance equals $P = .016$, meaning that with an increased body of NI participants the difference between the two groups is smaller. Compared to the listening comprehension test, and the PPVT-4 picture that Sivertzen ran, this is actually the first test that involves reading as part of the testing that has been done on the EI group. As presented in the theory chapter, receptive vocabulary can be used as a determinant for literacy skills and reading comprehension, thus the findings in Dahl and Vulchanova's (2014) study as well as Sivertzen's (2013) will be used in the discussion of the test results.

One aspect that has to be kept in mind, which makes the participants in this study quite unique, is that they started acquiring reading and writing skills in both L1 and L2 at the same time, since both Norwegian and English are subjects from the first grade (Utdanningsdirektoratet, 2013a), in comparison to most of the SLA literature that describes a sequential acquisition where the learner already has literacy skill in his or her L1. However, since there are so few English lessons per week during the first years of school, one can assume that literacy in Norwegian is acquired at a much more rapid rate than in English. There are several reasons for this. Firstly, Norwegian is the L1 of the participants and they are able to apply the vocabulary and the linguistic knowledge they already possess into comprehension (Lervåg & Aukrust, 2010). Secondly, it is shown that orthographic differences do affect the phase that literacy is acquired in, the lesser sound-letter correspondence in English is likely to cause Norwegian pupils to have a prolonged stage in which phonological recording and decoding dominates their English literacy

compared to their literacy skills in Norwegian (Lervåg & Aukrust, 2010; Seymour et al., 2003; Ziegler & Goswami, 2006). Thirdly, the process leading up to literacy in a language with a deep orthography, like English, has been shown to take longer time compared to languages with a more transparent orthography (Ziegler & Goswami, 2006).

However, a more intricate question that has to be addressed is why the EI group outperformed the NI group, when Sivertzen (2013) found that there was no significant difference in the receptive vocabulary size of the EI and the NI group in her study, and we know that L2 vocabulary is an important predictor of pupils' L2 comprehension skills. Sivertzen (2013) did find a difference in the raw scores, indicating that the EI group did perform better, although not significantly so when analysed, which Sivertzen contests to be a product of the discontinuation of the extra input program, causing the language development of the EI group to normalize (2013, pp. 33-34). Lervåg and Aukrust (2010) found that L2 vocabulary could predict the L2 reading comprehension in a much larger extent than L1 vocabulary predicts L1 reading comprehension (p. 617). With that in mind we could have expected to see little to no difference in the results on the reading comprehension test, whereas quite the opposite was the actual result. One possible explanation that is not in conflict with Sivertzen's findings is the explanation proposed by Traxler (2012) saying that input based teaching might facilitate an access to the L2 labels in the mental lexicon that is not routed via the L1 lexicon (p. 437). This could then explain why the EI group performed that much better, as they have roughly the same sized L2 lexicon as the NI group, but are able to access it faster when reading, consequently making them better at comprehending what they read. Another explanation could be what Cain (2009) points out, that early vocabulary knowledge and verbal skills are important predictors of later reading comprehension, thereby explaining why the EI group's performance exceeds the NI group's in reading comprehension skills. Additionally, the two groups have learnt to read under different conditions of instruction, which will be mirrored in the reading patterns of the pupils, thereby contributing to the explanation of why the EI group outperformed the NI group at the reading comprehension test.

It should also be kept in mind that the format of the reading comprehension test could possibly have invoked stress in the participants to such a degree that it affected their performance negatively. However, this is not very likely to be the reason why the NI group performed poorly compared to the EI group. All the participants were given

thorough explanations and the chance to ask as many questions as they wanted beforehand, and the conditions were the same; no disturbance took place. Another argument that helps validate the results of the reading comprehension test is the information that was gathered via the questionnaires. The results of the reading comprehension test have a high correlation with the data from the questionnaires. 91.7 percent of the participants of the EI group reported that they read English on a weekly basis or more frequently, whereas in the NI group only 40 percent read on a weekly basis or more frequently. These numbers are good indicators of how much written English the groups are exposed to. Although pedagogical, one might ask whether the NI group receives less homework or if they simply do not do their homework, or if they get time to do it at school. Furthermore, what is perhaps more telling is when they were asked to assess how much non-homework English they read. Here, nearly twice as many, 37.5 percent EI participants report reading English on a weekly basis or more frequently, whereas from the NI group only 20 percent read on a weekly basis or more frequently. Moreover 26.7 percent of the NI group report that they never read any non-homework related English texts, while no EI participants reported to never reading. This is much in line with Cain's (2009) statement "[c]hildren who fail to understand adequately what they read may lack the motivation to read in their leisure time" (p. 13). From these numbers a distinct pattern appears, indicating that the EI group as a whole does read more English, be it homework or not, compared to the NI group. There are many variables that can explain this difference, but it is striking that the correlation between the self-assessment data and the test results are this high, giving all the more reason to conclude that the extra English input program might have facilitated a prolonged positive effect that is reflected in the EI groups' literacy skills.

6.0 CONCLUSION

This study's aim has been to assess whether the pupils that partook in both Dahl and Vulchanova's (2014) and Sivertzen's (2013) study, still reap benefits from the extra English input that they were exposed to during their first year at school. In this study pupils from three schools partook, one being the EI group, and the other two accumulating the NI group. Of these, the total number of participants was 39, with the distribution of 24 in the EI group, and 15 in the NI group. It was hypothesized that the EI group would still show signs of positive benefits from having been part of the extra input program, but that it was expected to be significantly less dominant at the time of the testing, as three years had passed since the extra English program was discontinued.

The data gathered in this study have throughout this paper been discussed in light of the previous studies that have been conducted on the EI group. The extra English program was initiated by Dahl and Vulchanova (2014) when the EI group started school in first grade. They found that there was a significant difference in acquisition between the EI and the NI group during the first year. They tested both the receptive vocabulary and the listening comprehension skills, and hypothesized that the input based teaching and the quality of the target language exposure was the reason for the significant difference that was found in both receptive vocabulary and listening comprehension (Dahl, 2014; Dahl & Vulchanova, 2014). Sivertzen (2013) retested the receptive vocabulary of the EI group three years after the discontinuation of the initial study. She found that the growth rate of the EI group's receptive vocabulary had levelled out with the NI group's, and that there was no longer a statistically significant difference between the two groups.

At the time of testing the subjects of this study were halfway through fifth grade, less than a year after Sivertzen's data collection took place. Therefore the choice of doing the listening comprehension test was based on the idea that a short period of time between the data collection would give a more nuanced image of the EI group's skills. The listening comprehension test was first developed by Dahl and Vulchanova, and was extended for the purposes of this study. The reading comprehension test was added with the intent of mapping comprehension skills, in this case, the reading comprehension skills. The test results were analysed in SPSS using MANOVA and ANOVA analyses. The MANOVA result, $F(2,28) = 4.91$, $p = .015$, makes it possible to attest that the EI

group surpass the NI group in comprehension skills. It is hypothesized that this discrepancy is caused by the age of which the EI group was exposed to the extra English input, and that the acquisition of the L2 happened more implicitly during the initial stages. L2 teaching where the L1 is being used actively does intuitively become more explicit, in that the language is acquired when being in a near monolingual language mode. This means the findings do confirm the aspect of the CPH that an L2 is best acquired at young age.

When analysed in detail the listening comprehension test reveals that the EI group performed better than the NI group at the picture frames that were anticipated to come of as more difficult. From these results it is hypothesized that the EI group is better at singling out the meaning bearing components of spoken input, as well as being used to authentic language. During the testing several pupils from the NI group made remarks regarding the poor pronunciation of the native speaker who read for the recordings, which can be an indication of them not being used to hearing different dialects of English, or native speakers for that part. It was also hypothesized that the lack of authentic input in the NI group has made them acquire function words at a later stage, and thus in a less incidental way.

The reading comprehension test did reveal that there were large differences in comprehension, and more notably there was a very low percentage of the NI participants that were able to produce an answer. Previous research have indicated that L2 vocabulary size is a good indicator of L2 comprehension, but for the results of this study it was hypothesized that the input based teaching has made the EI group more apt to access the L2 lexicon directly, not via the L1, thus making them better at comprehension and inferencing even though they are approximately equal in vocabulary skills. Furthermore, the data from the questionnaire does indicate that the EI group is more literate and reads more.

In sum, with the findings of this study and the body of research that has been done on the EI group, this indicates that they are still under positive influence of early language input, almost three years after the discontinuation of the initial project. The findings of this study could be used as an argument of a reallocation of the teaching hour distribution in Norwegian schools. As the current curriculum has the most condensed

English teaching at age 16 to 18, as opposed to what this study, and most SLA research have proven; language is best learnt at a young age.

6.1 QUESTIONS FOR FURTHER RESEARCH

The results of this study do bring up several questions that could be pursued into future research projects. Some of these questions will be addressed in the following section.

One of the most prominent questions is what the development of the EI group would have looked like if the extra input program had not been discontinued. Moreover, would they have shown the same steep progress, as documented by Dahl and Vulchanova, throughout the teaching, or would a regression in the learning curve appear, as they come of age?

Another question that could be addressed is what role did it play that the teacher of the EI group was a native speaker of English? Could the same acquisition have taken place if the teacher was an L2 speaker of English? These are important questions that need to be addressed before one can use the data of this study as an argument for change in the teaching hour distribution, as most Norwegian English teachers are L2 speakers of the language.

None of the tests of this study, or the ones of the previous studies, have tested the participants' grammar knowledge. It could therefore have shed new light on the effects of the input based teaching if the pupils were tested by an acceptability judgment test or similar tests that tap into their grammar skills.

7.0 BIBLIOGRAPHY

- Aukrust, V. G. (2007). Young Children Acquiring Second Language Vocabulary in Preschool Group-Time: Does Amount, Diversity, and Discourse Complexity of Teacher Talk Matter? *Journal of Research in Childhood Education*, 22(1), 17-37.
- Barcroft, J. (2004). Second Language Vocabulary Acquisition: A Lexical Input Processing Approach *Foreign Language Annals*, 37(2), 200-208.
- Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition *Second Language Research*, 13(2), 116-137.
- Bialystok, E., & Hakuta, K. (1999). Confounded Age: Linguistic and Cognitive Factors in Age Differences for Second Language Acquisition In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* Mahwah: Lawrence Erlbaum Associates, Publishers
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems *Scientific Studies of Reading*, 9(1), 43-61.
- Boxtel, S. v. (2005). *Can the late bird catch the worm? Ultimate attainment in L2 syntax.* (P.H.D), Radboud University Nijmegen, Utrecht.
- Cain, K. (2009). Making Sense of Text: Skills That Support Text Comprehension and Its Development. *Perspectievs on Language and Literacy*, 35(2), 11-14.
- Cook, V. (2010). Chapter 4: The Relationship between First and Second Language Acquisition Revisited In E. Macaro (Ed.), *The Bloomsbury Companion to Second Language Acquisition* London: Bloomsbury
- Dahl, A. (2014). *Målspråkseksposering, språktilegnelse og prosessering i engelsk hos norske skolebarn.* Paper presented at the Språkdagen 2014, NTNU Dragvoll.
- Dahl, A., & Vulchanova, M. D. (2014). Naturalistic acquisition in an early language classroom. *Frontiers in Psychology. Special issue: Learning a non-native language in a naturalistic environment: Insights from behavioural and neuroimaging research*, 5.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behaviour Research Methods, Instruments and Computers*, 28(1), 125-127.
- DeKeyser, R., & Larson-Hall, J. (2005). What Does the Critical Period Really Mean? In J. F. Kroll & A. M. B. DeGroot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* Oxford: Oxford University Press Inc.
- Doughty, C. J., & Long, M. H. (2003). Chapter 1: The Scope of Inquiry and Goals of SLA. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* Oxford Blackwell Publishing Ltd.
- Ellis, R. (1985). The L1 = L2 Hypothesis: A Reconsideration *System*, 13(1), 9-24.
- Gor, K., & Long, M. H. (2009). Chapter 19: Input and Second Language Processing In W. C. Ritchie & T. K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition* Bingley: Emerald Group Publishing Limited.
- Grosjean, F. (2008). *Studying Bilinguals* Oxford: Oxford University Press
- Hulstijn, J. H. (2003). Chapter 12: Incidental and Intentional Learning. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* Oxford Blackwell Publishing Ltd.
- Karmiloff, K., & Karmiloff-Smith, A. (2002). *Pathways to Language: From fetus to adolescent* Harvard: Harvard University Press
- Kids, L. E. (n.d.). The voyage of the animal orchestra. 2013, from <http://learnenglishkids.britishcouncil.org/en/short-stories/the-voyage-the-animal-orchestra>

- Kunnskapsdepartementet. (2009). Fremmedspråk og språklig fordypning. 2014, from <http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stmeld/2010-2011/meld-st-22-2010--2011/4/3.html?id=641277>
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: use and interpretation* (2 ed.). Mahwah: Lawrence Erlbaum Associates, Inc.
- Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *The Journal of Child Psychology and Psychiatry*, 51(5), 612-620.
- Macaro, E. (2010). Chapter 1: Second Language Acquisition: The landscape, the Scholarship and the Reader In E. Macaro (Ed.), *Bloomsbury Companion to Second Language Acquisition* London: Bloomsbury Academic.
- Munden, J., Musk, N., & Wessman, K. (2007). *Steps: Textbook 4*. Oslo: Gyldendal Norsk Forlag.
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44(1), 1-35.
- Newport, E. (2006). Critical Periods in Language Development. *Encyclopedia of Cognitive Science*, 737-740.
- Postholm, M. B., & Jackobsen, D. I. (2011). *Læreren med forskerblikk: Innføring i vitenskapelig metode for lærerstudenter* Kristiansand: Høyskoleforlaget.
- Saville-Troike, M. (2006). *Introducing Second Language Acquisition* Cambridge: Cambridge University Press.
- Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, 94(2), 143-174.
- Sivertzen, T. G. (2013). *The role of input in early second language acquisition* (MA), NTNU Norwegian University of Science and Technology Trondheim
- Traxler, M. J. (2012). *Introduction to Psycholinguistics*. Chichester: Wiley-Blackwell.
- Trondheimsskolen. (2013). Resultater for barneskolene. 2014, from <http://www.trondheimsskolen.no/sample-page>
- Utdanningsdirektoratet. (2013a). Kunnskapsløftet fag- og timefordeling og tilbudsstruktur. 2014
- Utdanningsdirektoratet. (2013b). Læreplan i engelsk: Formål 2014
- Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science*, 9(5), 429-436.

APPENDIX 1: PARENTAL CONSENT FORM

Bakgrunnsinformasjon for forskningsprosjekt om sammenheng mellom ordforråd og språkforståelse hos norske 5. klassinger

Først og fremst, tusen takk for at du har sagt ja til å delta i dette forskningsprosjektet. I dette skjemaet ber vi om informasjon om ditt **barns språkbakgrunn**. Bakgrunnsinformasjon er nødvendig for at resultatene fra undersøkelsen skal kunne brukes, det er derfor viktig å besvare spørsmålene så korrekt som mulig. Fyll gjerne ut skjemaet sammen eller i samråd med ditt barn.

Alt av opplysninger som blir oppgitt i dette skjemaet vil bli behandlet uten direkte gjenkjenner opplysninger. En kode vil knytte dine opplysninger gjennom en deltakerliste, og det vil kun være autorisert personell som har adgang til deltakerlisten og som kan finne tilbake til informasjonen. Ved prosjektets slutt vill all informasjon bli anonymisert, det vil ikke være mulig å identifiserer barnet ditt i resultatene av studien når den publiseres.

Legg merke til at skjemaet har 3 sider, for at det skal kunne brukes i mitt prosjekt så må alle punkter være utfylt. Ved spørsmål vedrørende skjema, kontakt meg på odinlein@stud.ntnu.no.

Hvis du ønsker at ditt barn skal delta i dette prosjektet sendes utfylte spørreskjema med til skolen i en forseget konvolutt som samles inn av kontaktlærer.

Med takknemlig hilsen

Masterstudent Odin Lein Strand,

Professor Mila Vulchanova, NTNU

Del A: Personlig informasjon

Fødselsår: _____

Kjønn ☐ Gutt ☐ Jente

Bostedskommune: _____

Godkjennelse av foresatt til deltagelse i prosjekt

Underskrift foresatt: _____

APPENDIX 2: QUESTIONNAIRE ABOUT LANGUAGE BACKGROUND

Del B: Språklig bakgrunn

Morsmål

Er norsk morsmålet ditt?

☐Ja ☐Nei

Hvis ja, har du andre morsmål i tillegg?

☐Ja ☐Nei

Hvis ja, hvilke(t) språk?

Hvilket språk bruker dere hjemme?

Engelsk og andre fremmedspråk

I engelsk, hvordan vurderer du ferdighetene dine på hvert av disse områdene

	Under middels	Middels	Over middels
Lesing			
Skrijving			
Snakke			
Lytte			
Totalt			

Har du bodd i, eller hatt lengre opphold i et land hvor engelsk er hovedspråk?

☐Ja ☐Nei

Hvis ja, hvor lenge varte dette oppholdet?

Har du vært på kortere (under 14 dager) reise i et land hvor engelsk er hovedspråk?

☐Ja ☐Nei

Hvilke språk kan du utover morsmålet ditt og engelsk? (Hvis du ikke snakker andre språk, hopp over denne)

Språk	Nivå		
	Under middels	Middels	Over middels
Tysk			
Fransk			
Spansk			
- angi språk			
- angi språk			
- angi språk			

Hvor ofte leser du tekster på engelsk?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Hvor ofte leser du tekster på engelsk som ikke er lekse?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Hvor ofte leser du på engelske nettsider?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Hvor ofte lytter du til/hører du engelsk?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Hvor ofte ser du på engelskspråklige filmer/serier/tv-programmer?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Hvis du kan velge, hvilket alternativ foretrekker du når du skal se film/serier/tv-programmer?

☐ Undertekst på norsk ☐ Undertekst på engelsk ☐ Ingen undertekst

Hvor ofte spiller du engelskspråklige dataspill?

☐ Hver dag ☐ Flere ganger per uke ☐ Flere ganger i måneden ☐ Av og til ☐ Aldri

Andre faktorer i språklæring

Har du, eller har du hatt, problemer med synet utover normal brillebruk?

☐ Ja ☐ Nei

Har du, eller har du hatt, problemer med hørselen?

☐ Ja ☐ Nei

Har du, eller har du hatt, språkvansker av noe slag (spesifikke språkvansker, lese-/lærevansker eller lignende)?

☐Ja ☐Nei

Hvis ja, spesifiser?

Har du, eller har du hatt, andre diagnoser som kan tenkes å påvirke språklæring (ADHD, Aspergers syndrom eller lignende)?

☐Ja ☐Nei

Hvis ja, spesifiser?

APPENDIX 3: LISTENING COMPREHENSION TEST

The listening comprehension test can be accessed at www.hf.ntnu.no/odintest, simply type in any name into the slot to start the test. The original test designed by Anne Dahl my also be accessed at www.hf.ntnu.no/annedahl. Raw data from the test may be accessed afterwards from www.hf.ntnu.no/odintest/resultat.php. Below are the picture frames from the test used for this study, presented with the corresponding sentences.

<p>T1 The girl is eating an apple.</p> 	<p>T2 The boy is out in the rain.</p> 
<p>T3 There are trees outside the house</p> 	<p>T4 The man is wearing a white shirt.</p> 
<p>T5 This boy has a nice bird.</p> 	<p>T6 The little girl is jumping.</p> 
<p>T7 The car in this picture is red.</p> 	<p>T8 The people in this picture are happy.</p> 

T9 The boy is very sad.



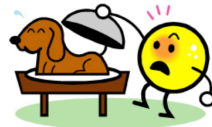
T10 Where is the black shoe?



T11 There are six hands in this picture.



T12 The girl really likes to sing.



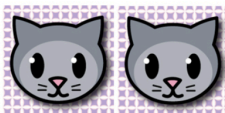
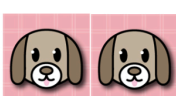
T13 She has washed the dog.



T14 Daddy is reading to the boy.



T15 There are four cats in this picture.



T16 This funny rabbit has pink ears.



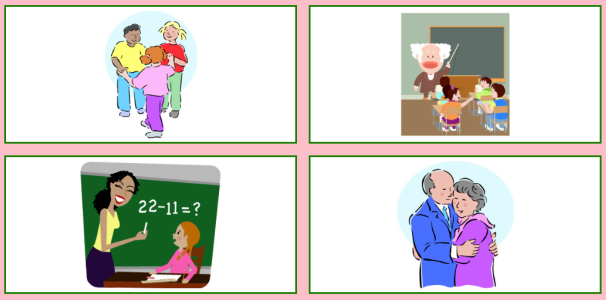
T17 Can you see the yellow flower?



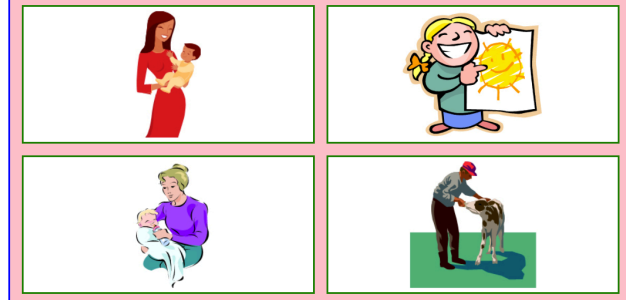
T18 This elephant is playing with a ball.



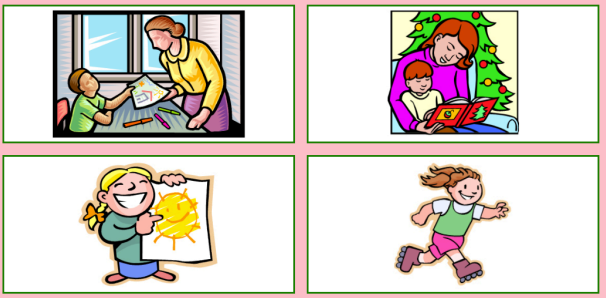
T19 The teacher is very old.



T20 The mother is feeding the baby.



T21 The boy has drawn his mother a picture.



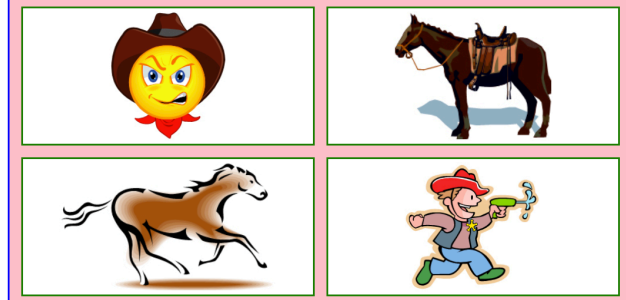
T22 The cow jumped over the moon.



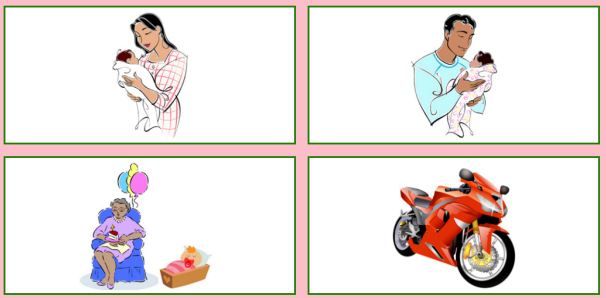
T23 Three children are playing.



T24 The horse is running.



T25 She is holding the baby.



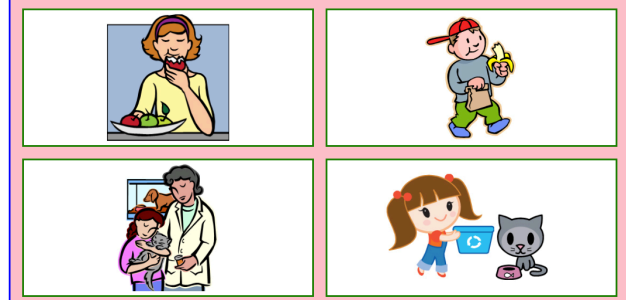
T26 He is sleeping under a tree.



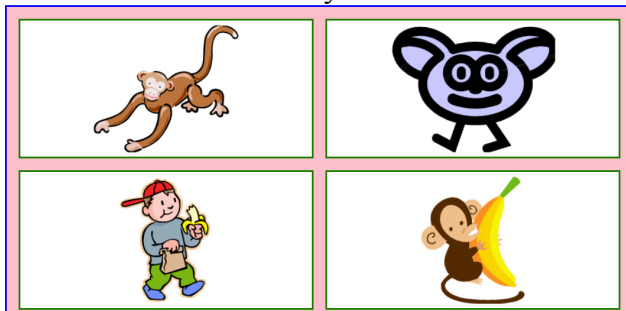
T27 They are playing in the water.



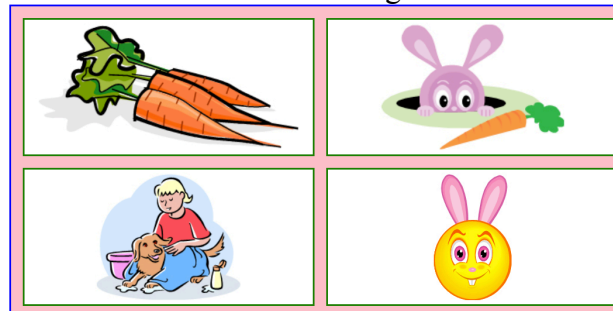
T28 The girl is holding the cat.



T29 The monkey has a banana.



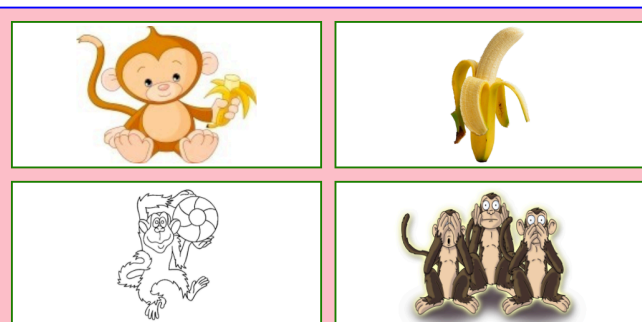
T30 The rabbit is looking at the carrot.



T31 The funny cat in the picture is brown and beige.



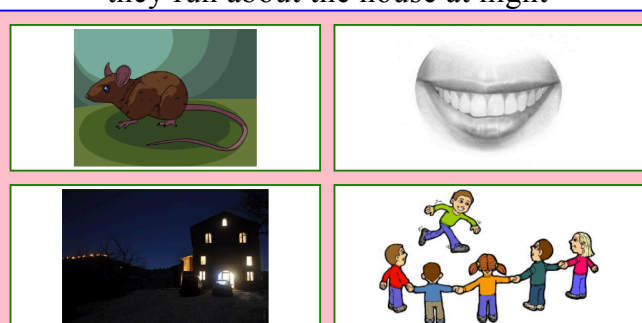
T32 The monkey is eating a banana.



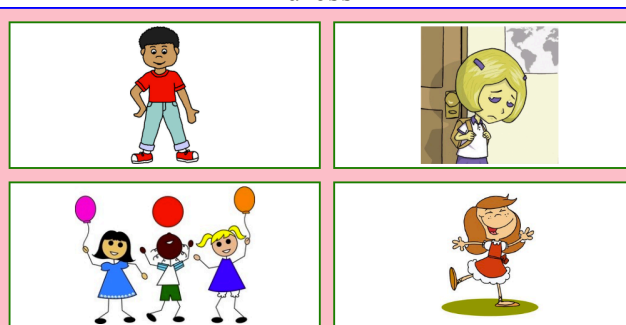
T33 the police officer is patrolling the street.



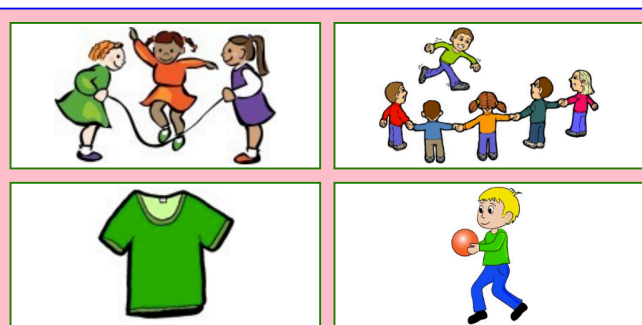
T34 Their ears are pink, their teeth are white, they run about the house at night



T35 The girl was very pleased with her new red dress



T36 One of the boys is wearing a green shirt



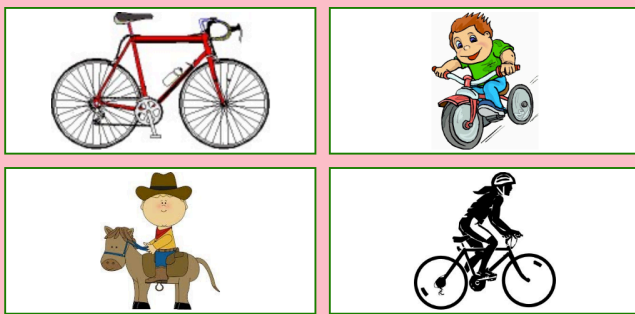
T37 The children in the picture are eating ice cream



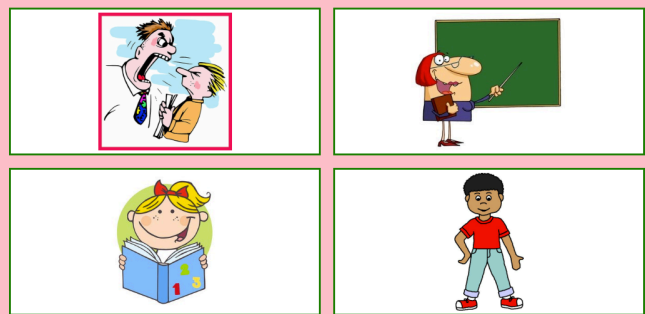
T38 John likes to read books at night



T39 Riding his bike is Jonathan's favourite activity.



T40 The teacher got angry with the boy because he had forgotten to do his homework



T41 There are two persons leaving the room



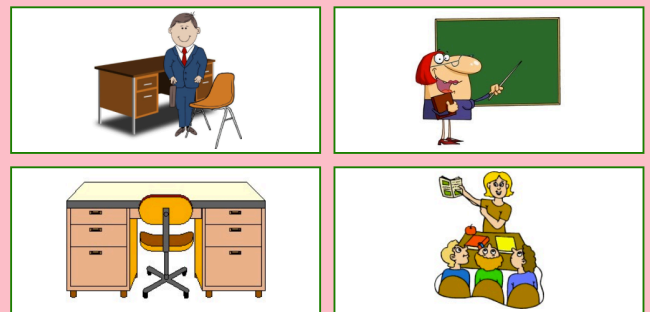
T42 In this picture there are eight different animals



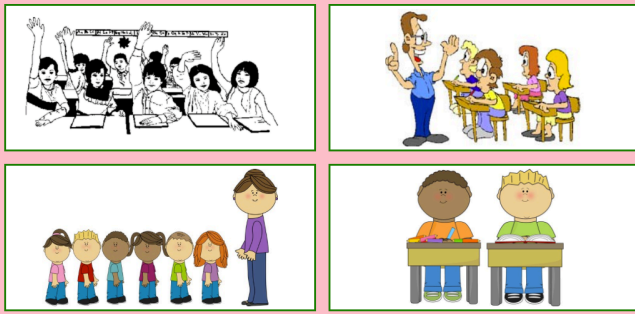
T43 There used to be 9 cats in the flock, now there are only 3.



T44 The teacher is standing behind the desk.



T45 The boys are sitting in front of the class



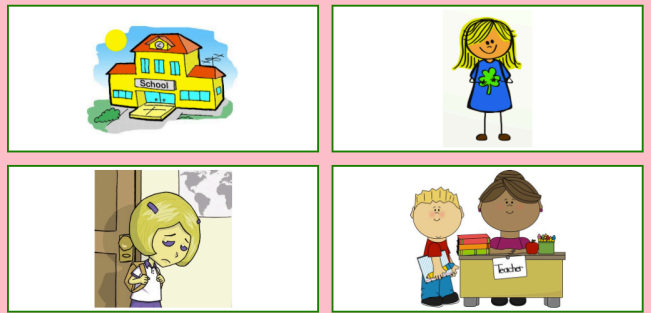
T46 The boys and girls outside play together



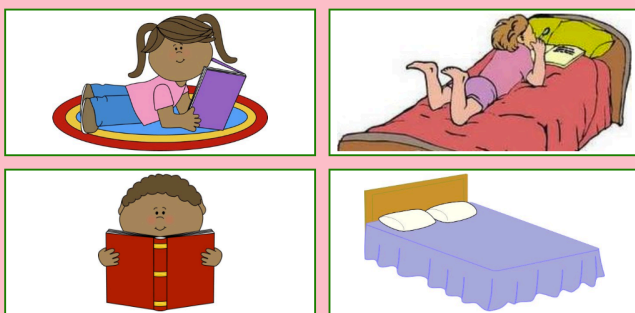
T47 In a dark, dark town, there is a little green house



T48 Matilda doesn't like being at school



T49 Amy is reading on her bed



T50 Where is the girl in the red and white coloured dress?



T51 Which of the girls does not have a pet?



T52 One of these animals is not a pet, which one?



T53 The girl is watching cartoons



APPENDIX 4: READING COMPREHENSION TEST

Below is the short story the pupils were presented with, found at the teacher resource webpage, British Council – Learn English kids (Learn English Kids, n.d.). The pupils were presented with the following questions after having read the story.

The voyage of the animal orchestra

It's a sad day. Our Ship, Symphony, hit a rock this morning and we are sinking. We must abandon the ship and swim for our lives.

Day 1. We're alive! We swam all day and all night until we reached land. Who know what's on this island? First we must sleep and rest.

Day 2. Today we walked around the island. We climbed a tree and all we saw was the deep blue sea and the hot sand! Now we must find food.

Day 3. Today we went swimming and fishing. There were sea-urchins in the sea! Now we must find fresh water. Who knows how we can carry it?

Day4. Today we climbed a volcano. At the top there was a pool of fresh rainwater. It was delicious! Now we must explore more of the island.

Day 5. Today we walked across the island. There were banana trees and coconut trees! Now we must make a shelter. Who knows how we can make one?

Day6. Today we made a shelter out of bamboo and palm leaves. We have fish, fruit, milk, water and shelter. Now we must have some music!

Day 7. Today the band practised on the beach. There was a ship on the horizon but it didn't see us. Who know how we can stop the ship?

Day 364. This morning the band was playing on the beach (the music was a bit loud) when a ship sailed by! I blew my seashell and the ship stopped!

It's a miracle! The ship heard the band and came to rescue us. We're finally leaving the deserted island. We're going home. Hip hip hurray! Hip hip hurray!

Questions:

1. Why did the ship sink?

.....

2. For how long did they have to swim?

.....

3. What could they see from the top of the tree?

.....

4. What was on the top of the volcano?

.....

5. What kind of food did they find in the trees?

.....

6. For how long did they stay on the island?

.....

7. What kind of materials did they use when building the shelter?

.....

8. What did the animals do to stop the ship?

.....

APPENDIX 5: FREQUENCY DATA FROM THE LISTENING COMPREHENSION TEST

Frequency table of the EI group's results on the listening comprehension test.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
Correct	95.2	100	100	100	95.2	100	100	95.2	100	100	100	100	100	100	100
Semi-correct	4.8	-	-	-	4.8	-	-	4.8	-	-	-	-	-	-	-
Wrong	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	T16	T17	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30
Correct	85.7	90.5	100	95.2	100	100	100	100	100	90.5	100	100	100	100	100
Semi-correct	14.3	9.5	-	4.8	-	-	-	-	-	9.5	-	-	-	-	-
Wrong	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	T31	T32	T33	T34	T35	T36	T37	T38	T39	T40	T41	T42	T43	T44	T45
Correct	100	100	95.2	52.4	100	52.4	100	95.2	95.2	100	38.1	100	71.4	47.6	57.1
Semi-correct	-	-	4.8	42.9	-	47.4	-	-	4.8	-	57.1	-	28.6	47.6	42.9
Wrong	-	-	-	4.8	-	-	-	4.8	-	-	4.8	-	-	4.8	-
	T46	T47	T48	T49	T50	T51	T52	T53							
Correct	66.7	100	95.2	100	95.2	61.9	61.9	100							
Semi-correct	33.3	-	4.8	-	4.8	38.1	38.1	-							
Wrong	-	-	-	-	-	-	-	-							

The data in this table is the valid percentage of the distribution of answers for each individual task.

Frequency table of the NI group's results on the listening comprehension test.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
Correct	91.7	100	100	91.7	91.7	100	100	91.7	100	100	100	100	100	100	100
Semi-correct	8.3	-	-	8.3	8.3	-	-	8.3	-	-	-	-	-	-	-
Wrong	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	T16	T17	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30
Correct	66.7	83.3	91.7	83.3	91.7	91.7	91.7	100	100	91.7	91.7	100	100	100	91.7
Semi-correct	33.3	16.7	8.3	16.7	8.3	8.3	8.3	-	-	8.3	8.3	-	-	-	8.3
Wrong	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	T31	T32	T33	T34	T35	T36	T37	T38	T39	T40	T41	T42	T43	T44	T45
Correct	75.0	100	91.7	8.3	100	50.0	50.0	91.7	91.7	100	50.0	100	58.2	16.7	41.7
Semi-correct	25.0	-	8.3	66.7	-	50.0	50.0	8.3	8.3	-	50.0	-	41.7	66.7	58.3
Wrong	-	-	-	25.0	-	-	-	-	-	-	-	-	-	16.7	-
	T46	T47	T48	T49	T50	T51	T52	T53	-						
Correct	75.0	100	66.7	91.7	66.7	33.3	50.0	83.3							
Semi-correct	25.0	-	33.3	8.3	16.7	66.7	50.0	16.7							
Wrong	-	-	-	-	16.7		-	-							

The data in this table is the valid percentage of the distribution of answers for each individual task.