The development of vocabulary and reading comprehension among Icelandic second language learners

Sigríður Ólafsdóttir

Ph.D. Dissertation

UNIVERSITY OF ICELAND

School of Education
The development of vocabulary and reading comprehension among Icelandic second language learners

Sigríður Ólafsdóttir

**Supervisor**
Dr. Freyja Birgisdóttir

**Co Supervisors**
Dr. Hrafnhildur Ragnarsdóttir  
Dr. Hetty Roessingh

**Doctoral Committee**
Dr. Sigurgrímur Skúlason

**Opponents**
Dr. Anne Vermeer  
Dr. Jóhanna Thelma Einarsdóttir

Ph.D. dissertation

Faculty of Education Studies  
School of Education, University of Iceland  
2015
The development of vocabulary and reading comprehension among Icelandic second language learners

A dissertation for a Ph.D.-degree in Education

© 2015, Sigríður Ólafsdóttir
All rights reserved
ISBN
Printed at: Háskólaprent ehf.
Reykjavík, 2015
This work is dedicated

to my Icelandic second language participants
I sincerely hope that they will benefit from their valuable contribution
Abstract

The development of vocabulary and reading comprehension among Icelandic second language learners

The strong relation between vocabulary and reading comprehension has long been recognized. As bilingual children tend to develop less vocabulary in each language compared to monolinguals in their only language, they frequently demonstrate a continuous deficiency in reading comprehension and academic achievement. Two prior studies with Icelandic second language compulsory school learners (Ice2) suggest that they make limited progress in the acquisition of Icelandic vocabulary across grades. However, only one of these studies was longitudinal and neither explored how vocabulary contributed to their reading comprehension and its development. The majority of vocabulary studies with second language learners in other countries have included very young children or have spanned a relatively short period of time. The few studies that have extended over a longer period of time are mainly based on comparisons of achievement levels across grades, and therefore do not take an advantage of the opportunity offered by longitudinal data.

The main purpose of this study was to explore the shape and the rate of vocabulary and reading comprehension growth among Ice2 learners from grade four to grade eight, and to detect the relation between these two skills. We also aimed to examine the extent to which linguistic and contextual factors might influence their emergence, and drive change in these skills. Finally, we sought to study the students’ proficiency in expressing their ideas in print in support of their arguments, and if this was related with their word skills.

Two age groups of Icelandic second language learners were tested on vocabulary (PPVT, Órskilningur, and Órðalykill) and reading comprehension measurements (NARA and Lesskilningspróf Námsmatstofnunar) at three time points over a period of three years: the younger group (49 children) in grades four, five, and six, and the older group (41 children) in grades six, seven, and eight. Additionally, a measurement of reading fluency was administered at Testing 1, to fourth and sixth graders, and a writing test at Testing 3, to sixth and eighth graders. A comparison group of age peers with Icelandic as their first language
(Ice1; 41 children in the younger group and 55 in the older group) was tested concomitantly. Growth model analysis was used to analyse the data, based on an accelerated design, which made it possible to obtain a continuous developmental view on the children’s vocabulary and reading comprehension growth from grade four through grade eight (Duncan, Duncan, Strycker, Li, & Alpert, 1999).

Our findings demonstrated that the Ice2 children had lower scores than the Ice1 children on all vocabulary measurements in grade four and the gap between the two groups widened during the study period. The Ice1 learners’ superior vocabulary skills were also reflected in their greater use of higher order, academic tier 2 words. For both groups, the number of these words went up in line with an increased ability to express their ideas in print in support of their arguments. While both the Ice2 and the Ice1 children made an unstable improvement in reading comprehension, the gap between the two groups remained the same. The findings revealed that Icelandic vocabulary skills in fourth grade positively predicted the rate of growth in reading comprehension, for both the Ice2 and the Ice1 learners. This means that those children who started with better word skills increased their reading comprehension scores at a faster rate than those who started with poorer vocabulary. The difference in reading comprehension that emerged on the Lesskilningspróf Námsmatstofnunar between fast and slow Ice2 readers in fourth grade disappeared during the time of the study as the slow readers increased these reading comprehension scores faster and managed to catch up with the others. On same test, the reading comprehension gap between slow and fast Ice1 readers remained the same, as the slow readers did not manage to accelerate their growth to reach the fast readers. The same applied to the Ice2 children’s performance on the NARA reading comprehension measurement, which is based on orally administered open-ended questions. The slow initial Ice2 readers did not reach the reading comprehension scores of the fast readers over these years. According to the results of the study, Ice2 learners’ age of arrival, as well as origin of their first language, are significant influencing factors in their acquisition of vocabulary and reading comprehension. The older the Ice2 learners were when they arrived in Iceland, the faster they increased their vocabulary and reading comprehension scores. There was also a significant difference between the vocabulary and reading comprehension skills of Ice2 learners with European first languages and their Ice2 peers with non-European first languages, always in favour of the former. However, first language
proficiency and literacy, as estimated by parents, first language instruction, and maternal education of the Ice2 learners did not influence vocabulary and reading comprehension skills and their rate of growth.

The implications of the findings are clear for Ice2 learners: solid Icelandic word skills are, first and foremost, fundamental for their academic progress in Icelandic speaking schools, even those who have the longest residence in the country need support, and in particular Ice2 children whose first language is not European.
Ágrip

Þróun orðaforða og lesskilnings íslenska grunnskólanema sem hafa annað móúrmál en íslensku

Sterk tengsl á milli orðaforða og lesskilnings hafa lengi og margoft verið sanreynd. Þar sem tvítyngd börn þróa gjarnan minni orðaforða í hvoru tungumál er eintyngd börn í sinu eina máli hættir þeim til að dragast aftur úr í lesskilningi og almennu námsgengi. Niðurstöður teggja íslenskra rannsókna benda til að grunnskólanemendur sem hafa íslensku sem annað tungumál (ísl2) auði hægt við íslenskan orðaforða sinn á milli ára. Æðins önnur þessara rannsókna var langsmiðsrannsókn og hvorug þeirra kannóð áhrif orðaforða þátttakenda á þróun lesskilninga þeirra. Flestar orðaforðarannsóknir á tvítyngdum börnum í öðrum lóndum hafa náð til mjög ungra barna eða spannað að fyrir þeim fá ár. Í þeim fáu rannsóknum sem hafa staðið yfir í lengri tíma þeirra þátttakenda á milli ára verið þó gorin saman og því hafa ákveðinn tækifæri sem langsmiðsgögn bjóða upp á ekki verið nær.

Meginmarkmið þessarar rannsóknar var að kanna hve hratt og hvernig orðaforði og lesskilningur þróast hjá ísl2 nemendum í fjórða bekk til áttunda bekkjar grunnskóla, einnig tengsl á milli þessara færniþätta. Þú var leitast við að skoða hvaða áhrif mímallegir, félagslegir og æðir umhverfisþættir hefðu á þróun orðaforða og lesskilnings. Að sústustu var könnuð færni barnanna í að þjá hugmyndir sínar í ritun og færa rök fyrir þeim, og hvort mætt tengja hana við orðaforða þeirra.

Tvær aldursþópar ísl2 grunnskólanema voru prófæðir þrisvar á þriggja ára tímaliblendi, yngri höpurinn (49 börn í fjórða, fimmta og sjötta bekk og eldri höpurinn (41 barn í sjötta, sjóunda og áttunda bekk). Orðaforðaprófr (PPVT, Orðskilningur og Orðalykill) og lesskilningsþrófr (NARA og Lesskilningsþrófr Námsmatssofnunar) voru lögð fyrir í hverri prófþróllögn. Lesfimiprófr var lagt fyrir í fyrstu þróllögn, í fjórða og sjötta bekk, og ritunaprófr í þróju þróllögn, í sjötta og áttunda bekk. Samanburðarhópur jafnaldra sem hafa íslensku sem móðurmál (ísl1; 41 barn í yngri höpur og 55 í þeim eldri) var prófæður samhlíða. Gögn voru greind með þróunarlíkandi sem gerir kleift að rekja samfellda þróun orðaforða og lesskilnings frá fjórða bekk upp í áttunda bekk (Duncan, Duncan, Strycker, Li, & Alpert, 1999).

Niðurstöður okkar sýndu að ísl2 börnin höfuðu minni orðaforða í fjórða bekk en ísl1 samanburðarhópurinn og bilið á milli hópanna breikkaði á rannsóknartímanum. Yfirburðir ísl1 nemendanna kom einnig fram í notkun

Niðurstöður benda eindregið til að íslenskur orðaföði isl2 nemenda sé lykill að framgöngu þeirra í námi í íslenskum skólum, að jafnhvel þau sem hafa dvalið hér lengst þurfi stuðning, og sérstaklega þau sem eiga móðurmál sem ekki eru evrópsk.
Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Freyja Birgisdóttir, who was a highly professional mentor throughout the whole journey. I had already experienced her expertise in my MA research, when she advised on the statistical work and read through and commented on the results chapter. In the PhD study she gave me priceless advice when narrowing down the topic, selecting measurements, and in organizing and writing out the thesis. Moreover and importantly, she gave me the scaffolding I needed to improve my academic writing proficiency in English, my second language. I thank her for reading my papers in such detail that it resulted in a vigorous and robust thesis. I am also thankful for her valuable support and encouragement at, sometimes, tough times.

I would like to express my appreciation and thanks to Hetty Roessingh for her inspiration, generosity, and encouragement. These date back to the years when I was teaching Ice2 children in a compulsory school in Reykjavik. Through e-mail communications with her, from reading her articles, and after mutual visits, I got this incurable interest and keen devotion to second language vocabulary acquisition. I thank her once again for presenting at a conference in Reykjavik in 2006, sharing with the audience her experience of highly effective instructional programs. I am also grateful for her endless hospitality when I, in company of Anh Dao Tran, visited her in Calgary same year. Together we all went to a TESL conference in Winnipeg at which Roessingh was a key speaker. My stay in Calgary last year was also an important and valuable milestone in the process. During my time there I got to know experts in the measuring of vocabulary and writing proficiency, Brock J. Wojtalewicz and Geoffrey Pinchbeck. I thank them sincerely for enjoyable and inspiring conversations and important advice. Also in Canada, Roessingh introduced me to Susan Elgie who read through my thesis and gave important feedback, for which I am very grateful.

I would like to thank Sigurgrímur Skúlason for the numerous conversations I had with him, over the phone, via e-mail, and at meetings, discussing the statistics. I am grateful to him for being there all the time, willing to participate, cooperate, and give advice. I also thank him for providing validated measurements for the research, based on item response theory.
I thank Hrafnhildur Ragnarsdóttir for sharing my vocabulary interest when I came to her in 2008, eager to do a MA study on vocabulary acquisition of Ice2 children, which was meant to be a starting point to a further and a continuous effective work in this field. It was encouraging to see books in her room written by Catherine Snow, by whom and her colleagues I had already read some articles. I thank her for translating the NARA test particularly for this PhD research and for taking part in selecting other measurements. I especially appreciate the meeting she arranged with Catherine Snow at the start of the project, at which I presented the research and obtained valuable advice.

I thank Anna Guðrún Júlíusdóttir, since all of this started during our cooperation when we were searching for ways to improve our teaching methods with Ice2 children in two preparation departments for immigrant children in Reykjavík. Our mission was that our pupils would feel as winners every school day, boosting up their self-esteem, enabling them to build up ambitious educational goals. I would like to thank Anna Guðrún Júlíusdóttir for being an inspiring and enthusiastic partner. Together we launched the website Námsgagnavefurinn Katla where we presented teaching material we created, with Icelandic vocabulary in central role, and the theoretical framework that supported our work.

I thank my former students, Shaira Cagatin and Tijana Dröbnjak, who translated permission letters to parents in their first languages, Tagalog and Serbian. It is a real pleasure to witness how they have succeeded as Icelandic second language learners in the Icelandic school system.

Special thanks to my fellow doctoral students, of whom some have already graduated, Anh Dao Tran, Anna Guðrún Edvardsdóttir, Auður Pálisdóttir, Edda Óskarsdóttir, Hrund Þórarins Ingudóttir, Skúlína Kjartansdóttir, Susan Elizabeth Gollifer, and Þóra Sæunn Úlfsdóttir. I thank them for rich discussions, sharing good and bad moments during the journey, and supporting each other in keeping the right spirit: endurance, devotedness, and enthusiasm.

I also thank my brother, Skúli S. Ólafsson, who defended his own PhD thesis a year ago, for his support and cooperation. It was a gratifying experience to have him around all the time, struggling with similar issues.

I would like to express my appreciation to my husband, my children, my daughter-in-law and son-in-law, who always supported me and encouraged me to take on and keep working on this large project. I would also like to
thank my parents who brought me up in the way that I was able to complete this work, equipped with the strength, skills, and attitude needed. Finally, I thank my whole family for showing endless patience, in particular when the study time tended to stretch out in the end.

This PhD study would, however, never have been possible without my cooperative participants. I send them my deepest gratitude for putting their efforts into completing all the demanding tasks. I sincerely hope that their work will give results for their own personal benefits. I thank their parents for giving their approval for the participation of their children.
# Table of contents

Abstract .......................................................................................................................... v  
Ágrip ................................................................................................................................. viii  
Acknowledgements ......................................................................................................... x  
Table of contents............................................................................................................ xiii  
List of Figures and Tables .............................................................................................. 17  
1 Introduction .................................................................................................................. 19  
2 Review of the literature ............................................................................................... 25  
2.1 Vocabulary definitions – what is involved in knowing a word ............................. 26  
2.1.1 Breadth of vocabulary knowledge .................................................................... 26  
2.1.1.1 High- and low-frequency words ............................................................... 26  
2.1.1.2 Formulaic language ................................................................................. 29  
2.1.2 Depth of vocabulary knowledge ...................................................................... 30  
2.1.3 Receptive and productive vocabulary .............................................................. 31  
2.2 Measuring vocabulary knowledge ....................................................................... 32  
2.2.1 Reliability and validity of vocabulary measurements ..................................... 33  
2.2.2 Types of measures .......................................................................................... 34  
2.2.2.1 Context independent and context depended measurements .................... 34  
2.2.2.2 Discrete and embedded measurements .................................................. 36  
2.2.2.3 Selective and comprehensive measurements ......................................... 36  
2.2.3 Measuring the breadth of vocabulary ................................................................ 37  
2.2.3.1 Selection of words .................................................................................. 37  
2.2.4 Measures of receptive and productive vocabulary ....................................... 38  
2.2.5 Measuring the depth of vocabulary .................................................................. 40  
2.2.6 Measuring bilinguals’ vocabularies ................................................................. 41  
2.3 The development of vocabulary ............................................................................. 42  
2.3.1 Vocabulary growth among L1 children ......................................................... 42  
2.3.1.1 SES, quality of input and the acquisition of vocabulary ......................... 44  
2.3.1.2 Implicit and explicit vocabulary acquisition .......................................... 45  
2.3.2 Vocabulary growth among L2 children ............................................................ 48
2.3.2.1 Ethnicity, SES and the acquisition of vocabulary ..........49
2.3.2.2 Exposure to each language .....................................50
2.3.3 The vocabulary gap between L1 and L2 children ..........52
  2.3.3.1 Vocabulary growth among L2 learners compared to L1 peers 56
2.3.4 Influence of cognates in first and second language ..........59
2.3.5 Reciprocal relation between first and second language vocabulary acquisition .............................................62
  2.3.5.1 Age of arrival and second language vocabulary acquisition 62
2.4 The educational context of Ice2 learners ..........................64
2.5 Summary ........................................................................65
2.6 What is reading comprehension? ......................................68
  2.6.1 The simple view of reading ........................................68
    2.6.1.1 The changing relationship between decoding skills, linguistic comprehension and the development of reading comprehension ........................................69
  2.6.2 The effect of orthography on the acquisition of basic decoding skills 70
  2.6.3 The acquisition of phonological and decoding skills among L2 and L1 learners ....................................................72
2.7 The influence of vocabulary on reading comprehension ..........76
  2.7.1 Second language vocabulary strongly contributes to second language reading comprehension .....................................78
  2.7.2 First language skills and second language reading comprehension - influence of age of arrival .................................81
2.8 The development of reading comprehension among L2 and L1 learners - The reading comprehension gap between L1 and L2 learners .86
2.9 Measuring reading comprehension .....................................89
2.10 Summary and conclusions ..............................................91
2.11 Aims and research questions ..........................................93
  2.11.1 Hypotheses ..........................................................94
3 Method .............................................................................101
  3.1 Research design and data collection .................................101
  3.2 Preliminary inquiries ....................................................102
  3.3 Ethical issues ..................................................................103
3.4 Participants........................................................................................................103
3.5 Measurements ...................................................................................................107
  3.5.1 Receptive word skills......................................................................................107
  3.5.1.1 PPVT ........................................................................................................107
  3.5.1.2 Orðskilningur (OSK) – a vocabulary subtest from the
        Icelandic testing battery for grade four-seven ...........................................108
  3.5.2 Receptive and productive word skills............................................................109
  3.5.2.1 Orðalykill (OLY) - an Icelandic word definition test.....109
  3.5.3 Free productive vocabulary: the writing test - What to do with
        the undeveloped space in the schoolyard?..................................................110
  3.5.4 Reading skills...............................................................................................111
  3.5.4.1 Reading comprehension ..............................................................111
  3.5.4.2 Decoding skills - Reading fluency...................................................113
  3.5.5 Screening for specific language impairment - A non-word
        repetition test.................................................................................................113
3.6 Analytic approach...............................................................................................114
4 Findings..................................................................................................................117
  4.1 Descriptive statistics ......................................................................................117
  4.1.1 Means and standard deviation on vocabulary measurements for
        each age cohort...............................................................................................117
  4.1.1.1 PPVT: Ice2 ..............................................................................................117
  4.1.1.2 Orðskilningur: Ice2 and Ice1 .................................................................118
  4.1.1.3 Orðalykill: Ice2 and Ice1 .......................................................................119
  4.1.2 Means, and standard deviation on reading comprehension
        measurements for each age cohort...............................................................120
  4.1.2.1 The NARA reading comprehension test...............................................120
  4.1.2.2 Lesskilningspróf Námsmatstofnunar, L.N.: Ice2 and Ice1 121
  4.1.2.3 The reading fluency test .......................................................................121
  4.1.3 Screening for language impairment: a non-word repetition test..............122
  4.1.4 Correlation for language impairment: a non-word repetition test ..........122
4.2 Research question 1 .........................................................................................127
  4.2.1 Vocabulary growth patterns.......................................................................128
  4.2.1.1 PPVT ......................................................................................................128
  4.2.1.2 Orðskilningur.........................................................................................129
Appendix 5 – The writing prompt: The empty Space behind the
Schoolyard........................................................................................................189

Appendix 6 - Rubric for holistic scoring of the writing prompt
(Edmonton Public Schools, 2008) .................................................................190

Appendix 7 - Tier 2 words used by participants ..............................191

Appendix 8 – An excellent writing sample ......................................193

References .............................................................................................................195

List of Figures and Tables

Figure 1. First languages of the Ice2 participants .................................105
Figure 2. The development of Ice2 participants’ mean raw scores on
the vocabulary measurement PPVT .............................................................128
Figure 3. The development of Ice2 and Ice1 participants’ mean raw
scores on the vocabulary measurement Orðskilningur ......................130
Figure 4. The development of Ice2 and Ice1 participants’ mean raw
scores on the vocabulary measurement Orðalykill ............................132
Figure 5. The development of Ice2 mean raw scores on the reading
comprehension test NARA .................................................................133
Figure 6. The development of mean raw scores on the reading
comprehension test, Lesskilningspróf
Námsmatssþofnunar L.N. ...............................................................134

Table 1. Vocabulary raw scores of Ice2 and Ice1 children ..............54
Table 2. Measurements administered at Testing 1, 2, and 3: to the
younger group (grades four, five, and six) and to the
older group (grades six, seven, and eight) ..............................................102
Table 3. Ice2 and Ice1 participants’ background variables (means
and standard deviations) ...............................................................106

xvii
1 Introduction

The number of Icelandic compulsory school students whose first language is other than Icelandic (henceforth Ice2 learners) has increased considerably the last few years. The share was only 0.02% in 1998, but by the year 2010 this proportion had risen to 5.5%, and to 6.5% in 2013 (Statistic Iceland, n.d.). Providing Ice2 children with Icelandic language skills that are fundamental for academic proficiency is thus a new challenge to the Icelandic school system.

Research indicates that the educational status among immigrants is lower in Iceland than in the European Union (EU) and the European Economic Area (EEA) countries (Garðarsdóttir & Hauksson, 2011), as first generation male immigrants in Iceland are almost two times as likely to be early school leavers than their Icelandic born peers. Moreover, results from the 2009 PISA assessment in Iceland revealed that the percentage of Ice2 learners among weak readers (11,6%) was almost three times higher than the percentage of Ice2 learners of the Icelandic population as a whole (4,5%) (Egelund, 2012, p. 29). Furthermore, the only two studies that have explored Icelandic language skills of Ice2 compulsory school learners (Thordardottir & Juliusdottir, 2012; S. Ólafsdóttir & Ragnarsdóttir, 2010) indicate that, unlike their monolingual peers (Ice1 learners), Ice2 children make limited progress in the acquisition of Icelandic vocabulary throughout their compulsory education.

The results of these two studies merit further investigation as vocabulary has emerged as one of the most important predictors of reading comprehension, general language proficiency, and academic achievement (Carlo et al., 2004; Hart & Risley, 2003; Laufer, 1992; Laufer & Goldstein, 2004; Laufer & Ravenhorst-Kalovski, 2010; National Center for Educational Statistics, 2011; Snow, Burn, & Griffin, 1998). Being able to comprehend and acquire knowledge from a diverse range of subject materials becomes crucial by grade four when students have started the longstanding process of reading to learn. Moreover, research indicates that children who show reading difficulties in the middle school years are more likely to drop out of school later in their academic career than are learners who do not experience such difficulties (Barrington & Hendricks, 1989; Snow, Tabors, & Dickinson, 2001, p 4).
The strong relation between vocabulary and reading comprehension has long been recognized (Blachowicz, Fisher, Ogle, & Watts-Taffe, 2006; Davis, 1944, 1972; Snow, Tabors, Nicholson, & Kurland, 1994). It was, however, in the late 1980s that vocabulary deficit was first highlighted as the main reason for academic failure among disadvantaged students (cf. Becker, 1977). A decade later, Stanovich (1986) proposed a model where he emphasised the interrelated development of phonological awareness, vocabulary growth and reading ability. The aim of his research was to find a theoretical foundation for developing tools that could break the cycle of some learners showing improvement and others showing declining performance in the acquisition of literacy. At that time vocabulary was receiving a burgeoning interest among education researchers (cf. Blachowicz, et al., 2006; Laufer, 1986; Maera, 1980) and in 2005 word meaning/vocabulary was ranked as a hot literacy research topic (Cassidy & Cassidy, 2004/2005; Cassidy & Ortlieb, 2013).

Interest in second language vocabulary learning and teaching has increased due to evidence showing that bilingual children (L2) frequently develop less vocabulary in each language than monolinguals (L1) (e.g. Droop & Verhoeven, 2003; Haraldsdóttir, 2013; S. Ólafsdóttir & Ragnarsson, 2010). This is thought to result mainly from the fact that bilingual children devote their time to two or more languages and tend to acquire, therefore, less vocabulary in each. Thus, they may know some words in one language but not the other (see Oller, Pearson, & Cobo-Lewis, 2007). Furthermore, this distributed nature of bilinguals’ lexical knowledge seems to be one of the main reasons for their frequent low performance on assessments of reading comprehension and other aspects of academic achievement (Lervåg & Aukrust, 2010; Lesaux, Crosson, Kieffer, & Pierce, 2010; Verhoeven, 1990).

Over the past decades, a great number of longitudinal studies have been conducted on the development of vocabulary skills among second language learners and its role in their educational achievement, particularly reading comprehension. Their results have generally demonstrated an increasing gap between L1 learners and L2 peers across their compulsory education. However, the majority of these studies have included young children either in kindergarten or at their first years of schooling only (e.g. Lervåg and Aukrust, 2010) or have spanned a relatively short period of time (often only two or three years) (e.g. Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, & Snow, 2011; Lesaux, et al, 2010). The few studies that
have extended over a longer period of time are mainly based on comparisons of achievement levels across grades and do not, therefore, take an advantage of the opportunity offered by longitudinal data to investigate the shape and rate of growth directly, or individual trajectories. Thus, we still need a more detailed picture on how vocabulary and reading comprehension develops over an extended period of time among second language learners, particularly over the middle school years.

Only one of the two studies that have explored Icelandic word skills of Ice2 compulsory school children (Thordardottir & Juliusdottir, 2012; S. Ólafsdóttir & Ragnarsdóttir, 2010) was based on longitudinal data and neither investigated whether the children’s vocabulary affected their reading comprehension development or other aspects of academic achievement. Moreover, the Icelandic language has a complex grammatical structure and may, therefore, be particularly difficult for second language learners to attain (Thordardottir, et al., 2002). Direct generalizations of research findings regarding second language learning across cultures may, therefore, be misleading.

In fact, educational success not only depends on solid reading comprehension but also on the ability to use ‘the language of books’ (Cummins 2008). Transitioning over to an academic language use becomes pivotal from grade four, for the acquisition and the development of academic literacy. Roessingh, Elgie, and Cover (2013) concluded from their study that higher order vocabulary use was a key variable in determining learners’ productive academic language proficiency. In no study has Icelandic academic word use of Ice2 middle school learners been investigated and how this knowledge is related with their ability to express their thoughts and ideas in print. We also have no information about the predictive power of vocabulary in productive language proficiency of this population.

The purpose of this research is to respond to this gap in knowledge. The central aim is to explore the development of vocabulary and reading comprehension among Ice2 middle school learners, and the predictive power of vocabulary for their reading comprehension growth. The current research will seek to address the limitations of previous research by employing a longitudinal design in which the development of vocabulary and reading comprehension will be tracked from grade four and to grade eight. Growth model analysis will be employed to compare the shape and the rate of vocabulary and reading comprehension growth among Ice2
children and Ice1 peers. Individual trajectories will also be explored to detect the relations between initial word and reading comprehension skills and their rate of growth. We also seek to examine the extent to which theoretically important linguistic and contextual factors may influence the emergence of these skills. Finally, we aim to study academic productive language proficiency of this population and how this is predicted by earlier word skills. This is the first study of its kind in Iceland and will generate valuable information that can inform educators and policy makers in Iceland, and beyond.

Guided by these aims the research seeks answers to the following research questions:

1. Is there an initial and developmental difference in vocabulary and reading comprehension skills between Ice2 and Ice1 middle school learners?
2. To what extent do reading fluency and Icelandic word skills of Ice2 and Ice1 learners predict their reading comprehension skills and its rate of growth?
3. To what extent are Ice2 learners’ vocabulary and reading comprehension level and rate of growth influenced by their age of arrival, first language proficiency, literacy, first language relatedness with the Icelandic language, and maternal education?
4. Do Ice1 middle school learners use a higher number of higher order (tier 2) words than Ice2 peers, and how is this related to the children’s productive language proficiency? To what extent are these predicted by earlier vocabulary skills?

To answer these questions, two age groups of Ice2 children, and a comparison group of Ice1 age peers, were tested at three time points, presenting development from grade four to grade eight. A set of measurements was used to test vocabulary and reading comprehension skills at each time point. Furthermore, reading fluency was measured at the start of the study, and writing skills at the last time point. Parents provided background information about age of arrival, first language proficiency, literacy, and instruction, and maternal educational level. The data was analysed using growth model analysis in which vocabulary and reading comprehension scores were dependent variables. This approach provided
information about the shape and the rate of growth of these skills, and the extent to which influencing factors contributed to initial status and the rate of growth. Correlation and regression analysis informed about the extent to which productive language proficiency was related to and predicted by earlier vocabulary skills.

This thesis consists of six chapters and eight appendices. In this first introduction chapter the statement of the problem and the significance of the study have been outlined. The second chapter provides literature review concerning vocabulary definitions and important issues regarding vocabulary measurements, validity, reliability and types of measures. The development of vocabulary among L1 and L2 children is discussed, as well as the difference between the two groups, and influencing factors. Reading comprehension development of L1 and L2 children is another topic in the second chapter, and the two fundamental influencing factors, decoding and vocabulary skills. The second chapter closes by a presentation of aims, research questions and hypotheses. The third chapter gives information about the research methods, design and ethical issues. The participants are introduced, measurements, and study approach. Descriptive statistics and correlation analysis are also included in the third chapter. In the fourth chapter the results are presented, under each of the four research questions. The results are discussed in the fifth chapter with reference to prior studies, and limitations of the study are also identified. In the sixth chapter, conclusions are drawn from the findings with implications and recommendations for further research.

The first four appendices include letters of approval for school principals and parents of the Ice2 children (those who participated in the prior MA study conducted by same researcher and those who only took part in the current research) and to parents of the Ice1 children. Instructions given to participants prior to the writing test are presented in the fifth appendix. The sixth appendix comprises rubrics for holistic scoring of the writings. Appendix seven includes a list of words used by the participants that were classified as higher order vocabulary. The last and the eighth appendix is an example of an excellent writing.
2 Review of the literature

Throughout the years, research findings have demonstrated a high correlation between vocabulary and reading comprehension (e.g. Baumann, 2005; Davis, 1944; Terman, 1916). Laufer investigated the percentage of known words in a text needed for reading comprehension achievement (Laufer, 1989, 1992; Laufer & Ravenhorst-Kalovski, 2010). Based on her findings, she suggests that in order to grasp the central meaning of a text at least 95% of the words have to be understood, and that 98% is the optimal level for solid understanding. However, there is also evidence that the relationship between vocabulary and reading comprehension is mutual (Verhoeven, van Leeuwe, & Vermeer, 2011); the better the understanding of surrounding words in a text, the easier it is to detect the meaning of unknown words, thereby increasing vocabulary development (Laufer, 1989). Also, good reading comprehension encourages further reading, which in turn leads to more encounters with new words. Thus, vocabulary skills contribute to reading comprehension and reading comprehension enhances vocabulary learning (e.g. Nagy, 2005; Stanovich, 1986).

A strong connection has also been detected between explicit, intensive vocabulary instruction and reading comprehension, in particular when the words taught are taken directly from the texts used for reading comprehension measures (National Institute of Child Health and Human Development, 2000). However, the effect of vocabulary instruction on reading comprehension in a more general sense (i.e. when there is little correspondence between words chosen for instruction and the reading comprehension measures) has been more difficult to demonstrate and is, therefore, more disputed (see Baumann, 2005; P. D. Pearson, Hiebert, & Kamil, 2007). Thus, although the relation between vocabulary and reading comprehension is irrefutable, both sets of skills are complex and studies indicate that the relationship between the two is highly dependent on how they are measured (e.g. Lervåg & Aukrust, 2010; Rydland, Aukrust, & Fulland, 2012).

In every language there are thousands of words. A small number of these words appear frequently, but a much larger number is rarely used. For that reason it is important that vocabulary instruction and measurements represent the specific vocabulary needs experienced by children at each stage in their education (see P. D. Pearson, et al., 2007). Thus, students not
only need to learn words that are necessary for their everyday communication, they also need to learn the specific words that belong to their academic studies, both for understanding and for productive use.

2.1 Vocabulary definitions – what is involved in knowing a word

2.1.1 Breadth of vocabulary knowledge

Vocabulary breadth (also often referred to as vocabulary size) is a term used to describe how many words a person knows. The number of words, however, can be based on various counting methods, depending on what is defined as being a word. The most basic way is to count each individual word form separately, regardless of meaning, word class or grammatical form. In that way, for example, the words work and works would be counted as two separate word forms. Lemma, on the other hand, represents a word with all its possible grammatical inflections. For example, the words teach, teaches, taught, and teaching are all counted as one lemma. A word family, however, refers to all words that are semantically related. Thus, in addition to teach, teaches, taught and teaching, the noun teacher and the adjective teachable all belong to the same word family (see Schmitt, 2000).

Children learn a great number of word forms, lemmas and word families in their early years and their vocabulary size expands rapidly during their school years. For example, it is estimated that English speaking five-year-olds generally know about 4,000 to 5,000 word families when they start school and that they learn about 1,000 new word families each school year, generating a vocabulary size of 20,000 word families by the age of 20 (Schmitt, 2000, p. 3).

School children most certainly need to learn the specific vocabularies that belong to the various academic subjects; but just as importantly they have to acquire good knowledge of high-frequency words that they encounter in all kinds of spoken and written texts.

2.1.1.1 High- and low-frequency words

In every language, there are a certain number of high-frequency words that appear very often in all sorts of speech and writing. However, although these words are used frequently they make up just a small portion of the total number of words in the word corpus of each language. For example, Carrol, Davies, and Richman (1971) found that the hundred most frequent
words in English covered 49% of the texts used in their research and the 2,000 most frequent words accounted for 81% of the same texts. Their data was based on 385-million-word corpus, evenly balanced between spoken language (unscripted conversation from radio and TV shows), fiction (books, short stories and movie scripts), more than 100 popular magazines, ten newspapers, and 100 academic journals, giving a total of almost 150,000 texts. Studies on Icelandic vocabulary have demonstrated similar findings (Magnússon & Briem, 1991; Sigurðsson, 1940). Friðrik Magnússon and Stefán Briem (1991) collected data from 100 texts each containing about 5,000 word forms, giving a total of 31,876 word forms. These texts were taken both from Icelandic and translated novels, bibliographies, theoretical texts and books for children and adolescents. They found that 58 Icelandic word forms (0.2% of total number of word forms), covered 50% of these texts (Magnússon & Briem, 1991, p. 1204).

High-frequency words not only span the majority of all written texts, they are also dominant in the spoken language used in daily communication. Children learn these words in the context of ‘here and now’ and their lived experience. Thus, these are mostly concrete words, foundational for the development of the basic interpersonal communication skills (BICS) (Cummins, 1981) necessary for everyday social interactions. Moreover, in the first years of schooling when reading texts are simple and the main objective is to practice decoding skills, high-frequency words comprise a large part of the reading material. A part of these words are however frequent morphologically complex verbs, many of which are auxiliary verbs, such as the English verb be (he is, they are), the French verb avoir (il a, ils ont), and the Icelandic verb vera (hann er, þau séu). Other high-frequency words consist of prepositions, conjunctions, adverbs, and pronouns, which form a link between words and sentences to make a coherent text. Although these words appear very often in all kinds of spoken and written texts, children only gradually develop skills in utilizing lexical items of this kind themselves in speech and writing (see Ragnarsdóttir, 2004, 2011).

In contrast to the small number of high-frequency words, the majority of words in each language are of low frequency and appear rarely both in spoken and written language. Magnússon and Briem (1991, p. 1204) found, for example, that 15,623 word forms, or 49% of total word forms in their Icelandic word collection, appeared only once or twice. These low-frequency words are seldom encountered in books dedicated for the
youngest learners. However, by grade three or four when children progress from primary to middle school, reading becomes a tool for learning and subject materials become richer and more complex. The number of this kind of words, therefore, tends to augment dramatically.

Beck, McKeown, and Kucan (2002) have divided words that are important for the acquisition of mature literacy in three tiers. The first tier comprises the high-frequency words, the most basic words that are encountered in all kinds of language settings. The third tier is made up of low-frequency words that emerge within specific fields of study and often have a precise meaning that is central to a particular area, such as mathematics, biology, geography, and so forth. Words in this category would thus be words like multiply, cardiovascular, and industry.

The second tier contains words that appear across a variety of domains, and are mainly used by mature language users, and in written texts of all sorts. Words of this kind are academic, for example the abstract words evidence, coincidence, absurd, and policy. One characteristic of words in this category is that their meaning can be expressed with high-frequency basic vocabulary. Knowing words of this kind is not only fundamental for comprehension but also enables students to express their ideas in more precise and mature ways (Beck, et al., 2002, p. 16). These words play thus a key role for the ability to use language for learning. Transitioning over to an academic language use of this kind, both for understanding and productive use, becomes pivotal from grade four, for the acquisition and development of academic literacy (cf. Roessingh, Elgie, & Kover, 2013).

Roessingh, Elgie, and Cover (2013) concluded from their study on writing samples of children in grade three (n = 79; 20 – 25% L2 learners) that skilled writers used fewer of high-frequency words and more of low-frequency words (see section 2.1.1.1). They concluded that vocabulary use was a key variable in determining the quality of the writings. Better writers

... are beginning to use words they would likely have learned independently: ‘the language of books’ (Cummins 2008). Better writers were attempting metaphor and infusing humor. These are further signs of the cognitive shift into concrete operations and emergent academic literacy. The measures of lexical diversity reveal that while many students have the lexical resources to address the demands of expository prose, there are others whose resources are limited and who may accordingly be linguistically
vulnerable as they move into upper elementary grades (Roessingh, Elgie, & Kover, 2013, p. 18).

Advanced productive vocabulary facilitates the ability to express ideas and thoughts in expository writing (see section 2.9), which is fundamental for academic proficiency. Some tier 2 words are metaphors that often only exist within a particular language and their meaning is generally tightly related to the culture in which the language is spoken. The Icelandic language is rich in such metaphors, many originating from fishing and agriculture. An English example of a metaphor is the expression *educational scaffolding*, which is based on a reference to platforms used during a construction process, an expression used in English but not in Icelandic. Obviously, metaphors can be difficult for second language learners to understand as they often have different cultural backgrounds, particularly when they have only had a short residence in the host country (see Roessingh, 2008).

2.1.1.2 Formulaic language

The two words, *educational scaffolding*, give the metaphor a meaning. Frequently the intended meaning cannot be derived from single words, but rather from word clusters or expressions, termed formulaic language (Schmitt, 2010, p. 117). Formulaic language is used in all settings of language use, ranging from daily speech to specific, academic expository texts. Examples of formulaic language are expressions like *for example* and *a lot of*, frequently used in everyday language settings, and the phrase *it should be noted that*, which is typically found in academic texts. Examples of formulaic language are idioms which meaning “cannot be derived from the meaning of its component words” (Schmitt, 2010, p. 117). Idioms are expressions such as; *this is all Greek to me*, meaning ‘I do not understand this at all’. Thus, both metaphors and idioms have to be understood as whole units. There are, however, important differences between them. For example, metaphors create an image in the mind, which idioms typically do not, and metaphors tend to be more academic or poetic than idioms.

Knowledge of individual words and formulaic language that are used in daily language settings are important for active natural communications and the development of basic interpersonal communication skills (BICS) (Cummins, 1981). Furthermore, being able to understand and use low-frequency words, idioms, and metaphors that appear across various academic fields are crucial tools for making sense of new knowledge and
for the development of higher order thinking (cf. Vygotsky, 1978), both of which are prerequisites for the development of cognitive academic language proficiency \((CALP)\) (Cummins, 1981). Moreover, students not only need to learn a vast number of high- and low-frequency words they also need to know much about the words.

### 2.1.2 Depth of vocabulary knowledge

The depth of vocabulary knowledge represents the quality of word skills. This aspect of word knowledge develops through the process of recognizing the meaning of a word as a whole, its individual parts and multiple forms, and after frequent encounters, using the word in different situations and contexts. With more knowledge about a word, such as its precise meaning, multiple meanings, and flexible use, our word knowledge becomes deeper (Kieffer & Lesaux, 2012b; Nagy & Scott, 2000; Nation, 2001, p. 27).

Vocabulary depth has been divided into three dimensions: morphological awareness, awareness of semantic relations, and syntactic awareness (e.g. Proctor, Silverman, Harring, & Montecillo, 2012). Morphemes are defined as the smallest unit of meaning in a language, which cannot be further divided. Root words, pre- and suffixes, as well as inflectional endings are all examples of morphemes. The understanding that words are composed of such parts and the ability to focus on the relations between them is generally referred to as morphological awareness (see e.g. Kieffer & Lesaux, 2012a) and is widely thought to lie at the heart of children’s ability to acquire morphologically complex words (see Birgisdóttir, 2012; Cain, 2010, p. 101; Nagy & Herman, 1984). Semantic awareness includes knowledge of semantic relations between words, for example *dogs bark*, subordinates, like *golden retriever*, and super ordinates, such as *animal*. Semantic awareness also includes knowledge of the multiple meanings of words, depending on the context in which they appear. The third aspect of vocabulary depth, syntactic awareness, refers to the knowledge of the structure of the language, i.e. the grammatical and syntactic constructions in which words typically appear. This dimension of vocabulary depth is knowing whether and how specific words are used in specific sentence contexts, and how the word forms depend on the contexts in which they appear. This involves knowing if words and sentences are grammatically correct (Cain, 2007, 2010; Chiappe & Siegel, 2006; Davidson, Rasche, & Pervez, 2010).
Learners develop morphological, semantic, and syntactic awareness through repeated encounters with words. Thus, the more frequent the word the more likely it is to be known both broadly and deeply. For example, Vermeer (2001) found high correlations between the breadth and the depth of vocabulary among Dutch L2 and L1 kindergartners and children aged four and seven, and both word dimensions were highly correlated with frequency of input. These findings suggest that L2 and L1 children who know more words also know more about the words (see also Kieffer & Lesaux, 2012b), and these skills are acquired through rich experience with words. Consistently, with age and further academic experience, learners add both to the breadth and the depth of their vocabularies, in both receptive and for productive use (see Schmitt, 2010).

2.1.3 Receptive and productive vocabulary

The term receptive vocabulary is used to refer to the words an individual understands when he reads or hears them. Productive vocabulary, on the other hand, represents the words that he can utilize in spoken and written texts (Nation, 2001). As input foregoes output (e.g. P. D. Pearson, et al., 2007), receptive understanding of a word precedes its productive use. Moreover, several encounters with a word in meaningful contexts are necessary before the word is likely to come to mind during speech or writing. When students start to use the word themselves they have reached the ultimate level of knowing the word. Laufer and Goldstein found that a small increase in the ability to recall words productively represented a larger improvement in receptive vocabulary (2004). Individuals’ receptive vocabularies are thus typically larger than their productive vocabularies (Laufer, 1998; Waring, 1997). Consistently, the less frequent the word, the more rarely it is seen or heard, the less likely it is to be remembered for active use. The difference between receptive and productive vocabulary skills thus generally increases as the frequency of the word decreases (Laufer, 2005; Waring, 1997; Webb, 2008). Corson has termed the relationship between receptive and productive vocabulary the lexical bar (1984), and that it is indeed the low-frequency academic words that are the most difficult to retrieve for active use. Moreover, Laufer (1998) found that controlled productive word skills (when learners are asked to name a picture presented to them, see section 2.2.4) progressed at a slower rate over one school year than receptive word skills, whereas free productive vocabulary (the word use in free speech or writing, see section 2.2.4) did not progress at
all. Nonetheless, using a word in a free language output, correctly and accurately, represents the optimal level of knowing a word, it is to own the word.

It is important to keep in mind that although vocabulary acquisition generally progresses from receptive to productive vocabulary and from broad to deep vocabulary skills (cf. Schmitt, 2010), the various word dimensions are inter-related and intertwined in natural language settings. Words appear in various forms and meanings and are used respectively. When vocabulary skills are assessed it is therefore not always possible, nor is it authentic, to measure only one specific dimension in isolation. It is, however, important to state which specific aspects of vocabulary skills are tapped in each measurement. Not only does that give more precise information, such clarification is highly important when vocabulary measurements and research findings are compared and interpreted.

2.2 Measuring vocabulary knowledge

Measuring vocabulary knowledge is not a simple undertaking, both because of its multifaceted nature and the large number of words that exist in each language. For example, when compared with the 26 alphabetical letters and 44 phonemes in English (see Kieffer & Lesaux, 2012b, p. 368), English learners can be expected to know about 20,000 word families by the age of 20. Lexical items can thus be collected from a universe of possible options. Furthermore, words are parts of whole units, sentences and paragraphs and frequently appear in multiple forms with different meanings. Various measures have been developed to assess vocabulary knowledge. These measures differ according to how test items are chosen, what aspects of word-knowledge they are intended to assess and the testing method they employ. Also, while some vocabulary assessments are designed to measure a specific dimension of vocabulary, such as the breadth of receptive word skills, others are designed to tap multiple aspects of vocabulary knowledge.

Three specific measures of vocabulary will be used in the present research. These include an Icelandic adaption of the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007; V. Ólafsdóttir, 2011) and two Icelandic vocabulary tests called Orðalykill (Gunnarsdóttir, Ólason, & Pind, 2004) and Orðskilningar (Gunnarsson & Skúlason, 2009). Each of these tests makes different cognitive demands and employs different methods. They will be discussed briefly in the chapters below; however, full
details of these tests will be given in the Method section (3.5.1.1, 3.5.1.2, and 3.5.2.1).

2.2.1 Reliability and validity of vocabulary measurements

As with any other measures of human functioning, the first thing to note about assessments of vocabulary is whether they constitute a valid and reliable measure of the skill in question. The term validity represents whether a particular test actually measures what it is supposed to measure. Validity of measurements can be established in a number of ways. One method that is frequently used is to explore how the test relates to a specific criterion. In the case of vocabulary, this criterion could be another established measure of vocabulary, or measures of other kinds of skill that have been shown to be related to vocabulary in a certain way (Schmitt, 2010, p. 181). Criterion validity is likely to be met, for example, if vocabulary measurements relate with reading comprehension tests. However, because of the complexity of word knowledge and the large number of words in each language, most of which appear very rarely, criterion validity can be difficult to obtain. This is particularly the case when there is little overlap between the lexical items included in each test (cf. Schmitt, 2010, p. 181).

Moreover, as all words, that are representative for vocabulary skills of certain groups of learners, cannot be included in a vocabulary measurement, test items are only samples of these skills. Construct validity is a problem of generalizing the observations made on such tests to the word skills that these observations represent (Schatzschneider & Petscher, 2011, p. 58). To demonstrate construct validity it is important to precisely define the specific word skills that the test is intended to measure, and the way in which test items were chosen.

Another way to establish the validity of vocabulary tests is to explore the content of the test and explore more closely how it fits with its pre-defined purpose. The main question concerns whether each item on the test is representative of the word knowledge that it is proposed to tap. If that is the case, the test has content validity (Schmitt, 2010, pp. 181-182).

Another estimate of the quality of a measurement is the extent to which a test gives reliable results. There are different kinds of reliability estimates. One has to do with whether a test gives comparable results across repeated
administrations (often termed test-retest reliability). Thus, if a participant completes a test today and again a day, a week or a month later, the results should be similar across the two occasions. However, it is important to keep in mind that children and adults vary in their day-to-day performance and 100% reliability may, therefore, be difficult to obtain. Inter-item reliability is another kind of reliability estimate and is used to assess the internal consistency of a test. This kind of reliability is achieved if the items in a given test correlate well with each other. High correlation between test items indicates that they are measuring the same underlying ability (Cain, 2010, p. 195). Thus, if an item correlates poorly with other items on the test, it should be discarded. Another way to look at internal consistency is to split tests into two or more parts and calculate the correlation between them. Internal consistency can only be achieved if the parts are equivalent to each other (Schmitt, 2010, p. 185).

Estimations of validity and reliability are important because they give the best indication of the quality of measurements. Vocabulary tests with confirmed validity and reliability should therefore be employed whenever possible. A number of such tests have been published and are widely used. Most of them are standardized norm-referenced tests that have gone through a strict validation process based on a large group of learners. However, no two groups of participants are the same and there is no guarantee that existing measurements will work for all participants. Therefore it is important to assess the reliability and validity of each test with the particular population that is included in each research (Schmitt, 2010, p. 184). Using measurements in research that have not been through strict validation process, but have only been translated from the original version (for example the NARA test used in the current study and the PPVT vocabulary test used in the study of Thordardottir and Juliusdottir (2012)) is not a preferred option and should only be used when validated tests are not available (see Skúlason, 2005).

2.2.2 Types of measures
Vocabulary tests can range from being totally context-independent with single words presented in isolation, to context-dependent where words are embedded in oral or written text and the participant must take an account of contextual information (Schmitt, 2010, p. 174). In most measurements, vocabulary is assessed as an independent construct. In some tests, however, vocabulary is treated as one part of a composite construct (general language
ability). Vocabulary assessments also differ according to whether they assess knowledge of words that have been selected specifically for each measurement, or through word use in oral and written texts.

2.2.2.1 **Context independent and context depended measurements**

In context-independent vocabulary tests word items are tested in isolation and participants must find or form the answers without being able to refer to any contextual clues (Read, 2000; Schmitt, 2010). An example of a widely used test of this kind is the PPVT Vocabulary Test (Dunn & Dunn, 2007) in which participants are asked to choose one picture out of four possible options that best represents the meaning of a word spoken by the administrator. This test is generally considered as a measure of vocabulary breadth (e.g. Lervåg & Aukrust, 2010). It has been adapted for use in many different languages, including Icelandic (V. Ólafsdóttir, 2011), and will be used in the current research (see section 3.5.1.1 for further details). Orðalykill (Gunnarsdóttir, et al., 2004) is another Icelandic vocabulary test that is context-independent and will be used in the present study (see section 3.5.2.1). In this test participants are asked to give a definition of individual words spoken by the administrator. Testing methods that involve recalling a word’s meaning without the help of any contextual clues requires a firm knowledge of that word, and even more than in a test like the PPVT where the test-taker has pictures he can revert to.

In context-dependent vocabulary tests, words are presented within a coherent passage. These sorts of tests assess the ability to use contextual clues from surrounding words in order to detect the meaning of a target word (Schmitt, 2010, p. 174). The Icelandic vocabulary test Orðskilningur can, to a certain extent, be considered context-dependent as words are presented within isolated sentences and their meaning is defined, or more easily identified, by surrounding words. This test is, therefore, not only a measure of vocabulary breadth, but also demands semantic awareness, a dimension of vocabulary depth (see section 3.5.1.2 for further details of this test).

Content dependent tests have the advantage over content independent tests that they are more authentic as words are rarely encountered or used in isolation in natural language settings.
2.2.2.2 Discrete and embedded measurements

When words are scored as an independent unit in a context-independent or a context-dependent measurement, the test is said to be discrete. In such cases, vocabulary is seen as a distinct construct and scored separately from other components of language skills (Read, 2000, pp. 8-10). On the other hand, when words are parts of larger units that together give a total score, the test is considered as embedded. An example of a test of this kind is when test-takers are asked to produce an oral or a written text and their vocabulary is one of several aspects of overall language ability that are being rated. Some reading comprehension measurements that are followed by comprehension questions can also be considered as embedded vocabulary tests, as the questions are often difficult or even impossible to answer without an understanding of particular key words in the text (Read, 2000). Specific words can thus be selected beforehand in both discrete and embedded measurements. However, when word skills are tested from test-takers’ language output, specific lexical items are not always chosen beforehand.

2.2.2.3 Selective and comprehensive measurements

Another distinction that can be made between different kinds of vocabulary tests is whether the tests are selective or comprehensive. When particular words are selected for a given measurement, the test is selective. Many widely used vocabulary tests (such as the PPVT, Dunn & Dunn, 2007) are of this kind. The same applies to the Icelandic vocabulary test Orðskilningar (Gunnarsson & Skúlason, 2009) and Orðalykill (Gunnarsdóttir, et al., 2004) that were mentioned above. However, in free productive tests learners are not rated according to their knowledge of particular words, but rather according to the quality of their overall vocabulary use, e.g. word range and syntactic skills (Read, 2000, pp. 10-11). Assessments of narrative and writing skills are an example of this kind of tests.

Thus, vocabulary measures differ according to whether they treat vocabulary as an independent construct or one part of a composite construct, and also whether they assess knowledge of words that have been selected specifically for each measurement, or by assessing overall quality of word use in oral or written texts. From what has been discussed above it is apparent that different methods available for vocabulary measurement tap different aspects of word knowledge. The important point is, however, that irrespective of the particular method chosen, the more words the learner knows, and the more he knows about these words, the more likely he is to
perform well on such tests. All of these different approaches challenge the breadth of participants’ lexicon be it directly or indirectly.

2.2.3 Measuring the breadth of vocabulary

Vocabulary size measurements do not produce a total number of words known by the participants; rather they give information about the size of their lexicon compared to other groups of learners (see Read & Chapelle, 2001). Because of the great number of words in each language, lexical items have to be carefully chosen for each measurement.

2.2.3.1 Selection of words

It goes without saying that it is impossible to include all words in a language in one measurement. The question therefore arises how the words on each vocabulary test are selected. A number of methods have been used in this respect, all depending on the purpose of the measurement. One frequent method, for example, is to take words by random out of a dictionary or frequency word lists. The latter approach typically involves a selection of a group of words from a pre-determined band of frequency (from the 1000 most frequent words, the 2000, and so forth). One of the most widely used vocabulary tests was developed in this way, The Vocabulary Levels Test (Nation, 1983, 1990; Schmitt, Schmitt, & Clapham, 2001). Another method, commonly used with L2 learners (Nation, 2001), is to choose words from a collection of lexical items that are important for academic purposes, if such lists are available.

Words have also been collected from both dictionaries and word frequency lists and then pre-tested with specific groups of individuals to give a norm referenced and standardized measurement tool for different age groups. The before mentioned PPVT (Dunn & Dunn, 2007) is an example of this kind of test. The development of the original American PPVT test started in 1950 when 3.885 word items were selected from The Webster’s New Collegiate Dictionary. These were pre-tested several times with English speaking individuals of a large age range, giving a total of 300 test items for the first edition. The test has since then been revised, pre-tested and standardized several times. The fourth edition is a product of a thorough revision and again with reference to dictionaries and word frequency lists. The PPVT-4 was standardized with 1.793 males and 1.747 females from age two and six months to 60, representing different levels of race, educational and social status (Dunn & Dunn, 2007). As mentioned above, an Icelandic
adaptation of this test (V. Ólafsdóttir, 2011) will be used in the current research (see section 3.5.1.1 for further details). The items in the Icelandic vocabulary test Orðalykill mentioned above were also selected from published frequency counts and pre-tested with specific age groups. A total of 352 words were chosen from The Icelandic frequency word list (1991) and The Icelandic dictionary (1993) and administered to University students. They were then asked to give the age at which they thought they had learned the meaning of each word. The words were then pre-tested three times and standardized with 300 Ice1 learners in grade one through ten (V. Ólafsdóttir, 2011).

Thus from the abundant number of words in each language, test items can be selected in multiple ways: by random, from dictionaries and word frequency lists. These items are then pre-tested and standardized with specific groups of individuals to give a measuring scale of word knowledge.

2.2.4 Measures of receptive and productive vocabulary

The main feature of vocabulary tests designed to assess receptive vocabulary is that they do not demand any oral or written expressions by the test-taker. These measures typically use multiple-choice formats, as they require the participant only to look for the answer but not to express it with his own words. Multiple-choice tests have the main limitation that test-takers can base their answers simply on random guessing. Therefore, it is often necessary to control for guessing in this type of measurements, depending on the number of options. There are some ways, however, in which the influence of guessing can be minimized. For example, options can be presented in alphabetical order, or in order of length. Items can also have as little similarities as possible, and distractors can be chosen so that they all relate to the item and the key in some way (Schmitt, 2010). Two of the vocabulary tests used in this research, Orðskilningur and PPVT, are measures of receptive vocabulary skills and use this kind of format. In both of them, distractors of the same word class, or thematic category, as the target word are presented in every item. For example in PPVT, with the target word shoulder all distracting pictures are of body parts.

Productive word skill measurements can be either controlled or free productive. In controlled productive vocabulary tests words are chosen beforehand. For example, when the main purpose is to estimate the breadth of productive word skills, participants are often asked to name a picture that is presented to them.
In free productive vocabulary tests, participants are asked to compose an oral or a written text. The writing test, ‘What to do with the empty space in the schoolyard?’ used in the present research is a test of this sort (see section 3.5.3). Information about the breadth of productive word skills can be obtained by calculating the total number of words in the text (TNT) or tokens, the number of different words (NDW) or types, and the type-token ratio reflects the use of different words in proportion to the total number of words. The higher number of word types indicates that more varied vocabulary has been used, representing larger and richer productive word skills (Schmitt, 2010, pp. 212-213). However, it is highly important to control for the length of the text, as longer texts tend to have lower type-token ratios, and importantly this method does not differentiate between the uses of high- and low-frequency words. For that purpose free productive vocabulary tests can be scored with reference to frequency band clusters by comparing the words used with frequency lists (e.g. whether the participant uses words from the 2000 most frequent words, 3000, 4000, and so forth) (see Schmitt, 2010, pp. 203-212). Still another way of measuring lexical diversity is to count words used by the learners that are beyond the most frequent words and that can be classified as tier 2 words, the words that are fundamental for the ability to use language for learning (see section 2.1.1.1). This last method of measuring productive vocabulary skills of learners was used in the current study, with a writing test as stated above (see section 3.5.3).

The main advantage of a free productive vocabulary assessment is that it is more authentic than the controlled productive assessments, as it better represents participants’ language abilities in natural language settings. Natural language used in daily surroundings, and particularly in oral story composition, comprises mainly high-frequency words concerned with ‘here and now’ experience. In order to gain insight into the full range of productive word skills, written expository texts are more effective than oral ones, as writings of this kind provide the opportunity to reach beyond conversational language. The writing test, ‘What to do with the empty space in the schoolyard?’ used in the present study (see section 3.5.3 and appendix 6) demands the writer to give descriptions, explanations, and persuasions, challenging the mastery of vocabulary for academic purposes, tier 2 words (see section 2.1.1.1). This challenging writing test thus aims at squeezing out the use of tier 2 words, as it demands “ … the learner to produce her best by taxing the full range of lexical resources available” (Roessingh,
2012b, p. 41). In other words, the writing task was designed in the way that it would raise the lexical bar (Corson, 1984), (see section 2.1.3).

It is however important to point out that the distinction between measures of receptive vs. productive vocabulary is not always very clear. For example, in many vocabulary tests, e.g. the test Orðalykill used in the current research, participants are asked to define the meaning of a given word with their own words (V. Ólafsdóttir, 2011). This type of test requires an understanding of a particular word presented orally or in writing to the participant and is, therefore, based on receptive vocabulary. However, tests based on word definitions also demand an oral or written output that frequently involves supplying synonyms or the use of other words that are related to the target word in some way. Consequently, such tests also tap productive word skills and an awareness of semantic relations (Proctor, et al., 2012), which is usually linked to vocabulary depth rather than breadth (Kieffer & Lesaux, 2012b).

2.2.5 Measuring the depth of vocabulary

As has been discussed above vocabulary depth is generally divided into three dimensions: morphological awareness, awareness of semantic relations, and syntactic awareness (see section 2.1.2). When vocabulary depth is explored specifically, vocabulary tests that tap more than one dimension of word knowledge can be analysed in a way that each dimension is scored specifically. For example, in free productive vocabulary tests, the use of various word forms can be scored for syntactic awareness, and semantic awareness can be rated by participants’ use of semantically related words, and whether their choice of words is semantically correct.

Vocabulary depth can also be measured by multiple measurements that each taps one specific dimension of word knowledge. Measures of morphological awareness, awareness of semantic relations, and syntactic awareness have been used for this purpose (cf. Kieffer & Lesaux, 2012b; Proctor, et al., 2012).

Employing multiple measurements to capture the various aspects of word knowledge contributes to the construct validity of research findings and also provides the opportunity to explore more precisely the extent to which various dimensions of word knowledge are related and how these differ between individuals and between groups of learners.
2.2.6 Measuring bilinguals’ vocabularies

Children who grow up with two languages frequently know some words in one language and other words in the other language, representing how their times are distributed between the languages. For that reason they are likely to know words related to domestic life in their first language and those related to schoolwork in the other language. Bialystok, Luk, Peets, and Yang (2010) compared the performance of monolingual and bilingual children aged from three to ten years on the English PPVT-III test. They found that the vocabulary deficiency detected among the bilinguals was largely confined to words that belong to their home language use, whereas school word knowledge for the two groups was more comparable (Bialystok, et al., 2010). Other studies have, however, demonstrated that there is indeed a shortage in bilinguals’ knowledge of academic words, often first emerging by the age of nine or ten and that negatively affects their reading comprehension and academic procedure (e.g. Chall & Jacobs, 2003; Roessingh, 2008).

The Icelandic vocabulary measurement PPVT used in the present study includes lexical items related to both home and school experience. The other two vocabulary tests, Orðskilningur and Orðalykill comprise, on the other hand, mostly words that are related with academic literacy. It is however important to note in this respect that PPVT and Orðalykill were pre-tested with Ice1 learners only. Orðskilningur is based on results of Ice1 and Ice2 participants from the Icelandic Testing Battery and then pre-tested with whole classes, and Ice2 learners therefore included. Icelandic vocabulary measurements that have been pre-tested and standardized with Ice2 learners separately, representing typical word knowledge and development of this group of learners are still missing.

Many researchers claim that bilinguals´ lexical knowledge must be measured in both (or in some cases all) their languages, as skill level in only one language is not the same indicator of language ability for bilinguals as for monolinguals (Hoff, Cynthia, Place, & Rumiche, 2012). Testing the Ice2 participants in their first languages is, however, beyond the scope of this research, indicating the need for further investigation in the field. It is nonetheless important to keep the purpose of vocabulary assessments in mind. The optimal objective of educational measurement is to give an indication of the extent to which learners are equipped to meet academic requirements and identify where support is needed. What the school system must strive for is to build up good reading comprehension in the school
language (or languages), in which word skills play a central role. It does not matter whether that language is Icelandic or English, or the first or the second language, good word skills have to be established. The Icelandic Compulsory School Act (2008, article 16, p. 7) insists that instruction in Icelandic compulsory schools shall be conducted in Icelandic. “Other languages than Icelandic may be used for instruction whenever entailed by the nature of the matter or by the National Curriculum Guide” (The Compulsory School Act, 2008, article 16, p. 7). However, Icelandic word skills are fundamental for Icelandic reading comprehension and academic procedure in Icelandic schools. Acquiring reliable information about the acquisition of Icelandic vocabulary of Ice2 children is therefore of great importance and in particular by the middle school years, when this group of children frequently starts to suffer from word skill deficiency.

2.3 The development of vocabulary

2.3.1 Vocabulary growth among L1 children

Children’s experience with words starts even before they are born when they hear the voices of their parents. Babies are already able to identify a few individual words and link with their experience by the age of six months. At this age when they start to babble, children demonstrate their first productive sign of linguistic abilities. Eleven-month old babies may understand around fifty common words and frequently say their first single words around their first birthday (see Gleason, 2005). Linking sounds and words to their meanings is termed as semantic development. Words that refer to objects and daily experience are more easily learned than abstract words and words that refer to something out of sight. With age, children’s phonological skills become more and more refined, and the morphology, syntax, and semantics of their language become more complex. They also start to learn social rules in language use – termed pragmatics. Language development appears to be similar in all language systems as children all over the world take the main steps in the acquisition process roughly at the same age and in much the same way (see Biemiller 2003; Biemiller, 2001; Biemiller and Slonim, 2001). On the other hand, children learn the grammatical structure of their language in a predictable order that is specific for each language and common for all children learning that language (Chomsky, 1965). The Icelandic language is more complex grammatically than, for example, English, Norwegian, and many other European
languages. Thordardottir and colleagues (2002) found that Icelandic two-year-olds needed more word skills than their English-speaking peers, before they were able to use grammatical regularity in their speech. However, as Icelandic children are used to frequent changes of words due to the complicated inflectional system of the language, they soon realize that something has to be done with many words. For example, research indicates that Icelandic pre-school children are more active in changing words into multiple forms in formal grammatical tests (although often erroneously), than Norwegian children at the same age (Ragnarsdóttir, Simonsen, & Plunkett, 1999). Ragnarsdóttir found (1998) that Icelandic children increase their ability to use past tense from age four to eight, and that this develops in concordance with the verbs’ frequency and regularity. This development followed a single mechanism model, as regular and irregular past tense forms were equally used for generalizations, representing influences from language input. There is some indication that the complexity of the grammatical system of the Icelandic language may influence children’s ability to reflect on its morphological structure. Birgisdóttir (2012), for example, found that the performance of Icelandic primary school children in a widely used morphological awareness task (Nunes, Bryant, & Bindman, 1997) was better than has been demonstrated in research involving English-speaking children at the same age, solving the same kind of test (Kirby et al., 2012).

At each stage in language development, there are significant individual differences in vocabulary size and quality of word skills, both in regards to understanding and productive language use. This variability may partly lie in differences in natural linguistic talent, which is probably to a considerable extent genetic (Dale et al., 1998; H. Gardner, 1983). However, there is little doubt that language development is to a large extent a product of the quantity and the quality of input from the children’s linguistic environment (Hart & Risley, 1995). Thus, children’s word learning starts in the home setting with parents. Parents talk to their infants, show them toys, play and sing for them, read books, show them pictures, and talk about the pictures. Children and parents also use gestures to communicate; they point at things, clap hands, shake their heads, and so forth (Hoff, et al., 2012). All these different forms of communication lead to word learning. However, the number and variety of words children hear, right from the very first days of life, differ greatly. According to a much-cited study by Hart and Risley (1995), children who hear the greatest number of words in their first three
years of life hear three times more words than those who hear the fewest, and they also learn new words three times faster (Hart & Risley, 1995). Also, two-year-olds who hear longer utterances build productive vocabularies significantly faster than age peers who hear shorter utterances (Hoff, 2003). Differences in language input may rise for a variety of reasons. One of the most frequently cited, however, is discrepancy in children’s social economic status or SES.

2.3.1.1 SES, quality of input and the acquisition of vocabulary

Social economic status (SES) is a term used for a set of background variables, usually comprised of educational level, occupational prestige, and income. Together, these variables are thought to create basic conditions of life, in social order, in which individuals are situated at different levels (Hoff, 2006, p. 60). In research, however, it varies whether SES is indexed by a compound variable, or a single indicator, most frequently maternal education. In either case, SES has emerged as a strong and robust influencing factor on children language development.

One of the best known studies on the link between SES and language acquisition was carried out by Hart and Risley (1995). They observed 42 children and their families from different social and educational backgrounds, during their first two and a half years of life. They discovered that children from professional families heard almost three times more words per hour than children from welfare families, and more variety of words as well. At the age of three, these differences in input were reflected in the children’s language abilities. Those with higher SES used a greater number and range of words and their styles of interaction were more productive than those of the children with lower SES. Similar findings have been obtained when mother’s education has been used as an indication of SES. For example, Hoff (2003) found that two-year-old children whose mothers were college educated were exposed to richer language than children who had mothers who had only completed high school, and this difference was reflected in richer vocabulary use of the former group. This pattern appears to emerge worldwide; Hoff and Tian (2005), for example, obtained similar results in their research with two-year-old Chinese children. It was again the input provided by the mothers that accounted for the differences in vocabulary outcomes that emerged between different SES groups. Similarly, in a cross-sectional study, Ragnarsdóttir, Birgisdóttir, and
Gestsdóttir (2009) found large individual differences in the vocabulary size among four-, five- and six-year-old Icelandic-speaking children and these differences were partly explained by the educational level completed by their mother.

These results give cause for concern as research indicates that children with small vocabularies learn new words at a slower rate than those with richer word knowledge (see Stanovich, 1986). Also, this interaction between vocabulary size and SES can have a spiralling effect. - Young children from families of lower SES who start school with poor vocabulary skills are less likely to catch up with their age peers than children from families of higher SES, who also demonstrate poor vocabulary skills in their early age. Thus, the difference between the two groups is likely to grow through the course of school as children with higher SES may have more resources and better familial support (see Emmelot, Schooten, Timman, Verhallen, & Verhallen, 2001; Snow, et al., 2001). These results indicate a stronger need for scaffolding among children with low SES within the school setting.

2.3.1.2 Implicit and explicit vocabulary acquisition

Most words are learned implicitly from the natural settings in which they appear in daily life, both in social communication of various kinds and from reading (e.g. Grabe & Stroller, 2002; Nagy, Anderson, & Herman, 1985). Implicit vocabulary acquisition can be defined as “learning based on unconscious processes of generalization and abstraction” (Schmidt, 1990, p. 1). The individual is focused on the message behind the spoken or written language rather than on the vocabulary itself, and automatically picks up new words. Compared to spoken language, written texts include much more variety and a larger number of words and many low-frequency words only appear in print (see section 2.1.1.1). Reading experience is, therefore, one of the most important influencing factors in the acquisition of vocabulary.

Research indicates that those who have good reading skills tend to read more than those with poor reading proficiency (see Nagy & Herman, 1984, p. 327). Consequently, good readers encounter a greater number of novel words and are, therefore, likely to expand their vocabularies through reading to a larger extent than poor readers who read less. For example, Verhoeven, van Leeuwe, and Vermeer (2011) found that decoding skills in third grade \((n = 2.790)\) promoted lexical growth from grade three to grade six. Moreover, readers need to understand the majority of words in a text to be able to use context to derive the meaning of unknown words (Laufer &
Ravenhorst-Kalovski, 2010). This gives children (and adults) with large vocabularies a clear advantage, resulting in a widening vocabulary gap between children with different initial levels of vocabulary skills. Additionally, studies indicate that in order to be able to guess correctly the meaning of unfamiliar words from context, surrounding words have to be very accurately understood and learners need to have mastered certain reading strategies (cf. Huckin & Coady, 1999, pp. 181-193), which are more likely to be acquired by those with rich vocabulary skills.

Furthermore, several encounters with words are necessary for learners to engrave them in memory and to be able to recall them for active use in speech and writing. Thus, solid word skills not only result in faster rate in receptive vocabulary acquisition but can also result in a more effective productive word increase (see section 2.1.3) which then leads to more advanced productive language proficiency. Roessingh, Elgie, and Cover (2013) found from their study on writing samples of children in grade three that more skilled writers used more of low-frequency words (see section 2.1.1.1). Based on the fact that a large variability emerged in writing samples of children attending same schools the researchers suggested that those with better writing proficiency had learned the low-frequency words from their parents and from reading, which then gave them a large advantage over those with poor lexical skills.

Even though most words are learned implicitly through quality oral and written language input, learning new words only indirectly may lead to slower word acquisition among all learners than an explicit approach (Collier, 1995; Fraser, 1999; Huckin & Coady, 1999; Lipka, Lesaux, & Siegel, 2006; Paribakht & Wesche, 1997; Wesche & Paribakht, 2000). In every language there are thousands of words - a small number of them are used very frequently, but the vast majority appears very rarely (see section 2.1.1.1). Thus, a great majority of words appear too sporadically in order to be mastered implicitly by learners. This means that reading alone is not sufficient to close the vocabulary gap between learners.

To meet the vocabulary needs of learners with smaller vocabularies, and to enhance word learning of all learners, an approach that emphasizes both implicit and explicit vocabulary learning is, therefore, fundamental (e.g. Dickinson & Tabor, 2001; D. Gardner, 2004; Laufer, 2005; Schmitt, 2008; Vadasy & Nelson, 2012). The explicit approach should include frequent rehearsals of target words, rich explanations and productive work with target words, and opportunities to use the words in all kinds of activities.
(e.g. Roessingh & Douglas, 2013; Snow & Lawrence, 2011). Furthermore, Lesaux and colleagues (2015) found that those L2 and L1 adolescents with the smallest vocabularies benefitted the most from an intensive vocabulary instruction, as they made the largest improvement in both reading comprehension and in writing proficiency, while however all participants profited from the intervention.

Given the strong reciprocal association between comprehension, linguistic proficiency and vocabulary knowledge (see section 2.7); a common vision is needed among all teachers, practitioners, and policy makers, emphasizing implicit and explicit approaches to vocabulary instruction:

Although individual teachers may be successful in using a variety of strategies for vocabulary instruction, what is needed is a comprehensive, integrated, school wide approach to vocabulary in reading and learning. By integrated, we mean that vocabulary is a core consideration in all grades across the school and in all subject areas across the school day (Blachowicz, et al., 2006, p. 526).

Throughout the years little importance has been assigned to explicit word learning in Icelandic schools (e.g. Aðalnámskrá Grunnskóla, íslenska, 2007). In the section on Icelandic in the Icelandic curriculum (Aðalnámskrá grunnskóla: almennur hluti 2011 og greinasvið 2013) vocabulary is still not specified as an influencing factor on reading comprehension (Aðalnámskrá grunnskóla: almennur hluti 2011 og greinasvið 2013, pp. 98-99), and no mention is made of explicit vocabulary instruction as a valuable way to bridge the vocabulary and reading comprehension gap that exists between learners (cf. Birgisdóttir, 2010). On the other hand, rich emphasis is placed on the development of grammar, the act of creating new words, the ability to use a rich variety of words, and to look up words in dictionaries (Aðalnámskrá grunnskóla: almennur hluti 2011 og greinasvið 2013, pp. 101-105). In spite of the fact that some schools have implemented explicit vocabulary instruction, like Orð af orði (Engilbertsson, 2010), little focus on word learning has dominated in traditional Icelandic schoolbooks used in many schools. In these books work with Icelandic grammar and other formal aspects of the language is emphasised, while tasks intended to bolster children’s vocabulary tend to be rather limited and in little context
with other learning materials and topics of study (e.g. Gestsdóttir, 2005; Jónsdóttir & Adalsteinsdóttir, 1999; Kristmundsson, 2007, 2008).

As young children begin their school careers with very different word skill levels, there is little doubt that word learning needs to play a vital role in educational policy and practice. This may be particularly important for children who grow up using more than one language as many studies indicate that bilingual children tend to develop significantly smaller vocabularies in each of their languages than their monolingual peers. Additionally, children using another language at home than the school language, do not have the same resources in that language as those who use the school language at both settings. For that reason L2 children run the risk of school failure (see August, Carlo, Dressler, & Snow, 2005; Snow & Lawrence, 2011).

2.3.2 Vocabulary growth among L2 children

Children who grow up with more than one language are called bilinguals. The language they learn at home, generally from their parents, is their first language. Some children acquire two (or in some cases more) first languages because two languages are spoken at home. Other children learn their first language at home and another language outside the home, for example in kindergarten or at school. That language becomes their second language. Those who learn two languages simultaneously from birth are called simultaneous bilinguals. Children who learn the second language before the age of four are generally also called simultaneous bilinguals, as they are still in the early language acquisition period (see Grosjean, 2010). Children who acquire their first language from birth and then start to learn a second language after the age of four are called successive or sequential bilinguals.

Simultaneous bilinguals show similar linguistic development as children who only grow up with one language, termed monolinguals. The babbling starts around the same age, the capacity to perceive different sounds emerges at a similar age, and the first word is spoken at 11 months of age, on average. Simultaneous bilinguals generally develop each language separately, demonstrating similar acquisition patterns as monolinguals learning the same language. Bilinguals also learn words in the same manner as monolinguals, they first learn concrete words and words that refer to here and now experiences (see section 2.3.1). Their word skills then develop with age, become richer and deeper (Grosjean, 2010).
Bilinguals learn words in each language from the people around them who speak those languages. Some bilingual children occasionally blend languages. An example is when the French word chaud and the English word hot becomes shot (see Grosjean, 2010). The more words children hear, the more words they will learn in both languages (Cummins, 1979; B. Z. Pearson & Fernandez, 1994). Children who grow up with two or more languages are no different to other children who grow up using only one language – they naturally differ in linguistic talent and they too belong to families of various SES, providing different levels of quality language input. Bilinguals also belong to various ethnic groups, representing various cultural upbringing habits. All these factors appear to be related to word skills development in monolingual as well as bilingual children (see section 2.3.1). However, there are also a number of influencing factors that specifically have to be considered in relation to the language acquisition of bilinguals. For example, they have to distribute their time between two languages (or more), they are exposed to different languages at different times in their lives at different levels of development, and their knowledge of their first, second, or even third language varies greatly. Bilinguals are, therefore, a very mixed group with even a greater range of word skills than monolinguals (see Grosjean, 2010).

### 2.3.2.1 Ethnicity, SES and the acquisition of vocabulary

Just like monolinguals, bilingual children learn their first language from people they share their home with. These people, whose first languages are other than that of the host country, are likely to trace their origins to various different cultures. In the multiple cultures throughout the world parents communicate with their children in different ways. In some cultures parents talk and read more for their children than in other cultures. For example, in Vietnam, parents have in general not been used to reading books for their children and that may therefore possibly apply to second language learners with Vietnamese home culture (A. D. K. Tran, personal communication, September 12, 2013). Furthermore, mothers in low-income African-American (of African origin living in America) families appear to talk less with their children than European-American (of European origin living in America) middle-class mothers (Heath, 1983; Hoff, 2006), which is reflected in a slower rate of vocabulary development of children belonging to the former group (Roberts, Burchinal, & Durham, 1999). However, a large number of learners with African-American background also belong to lower SES
families and the effect of ethnicity on their language development has proved to be an indistinguishable variable from their SES status.

Low SES status applies to a great number of immigrant groups all over the world. Many come from developing countries, or countries that have had a long history of unemployment, or even war, which may explain the fact that immigrants are frequently less educated than those of native origin (European Commission, 2011). Furthermore, immigrants are also more likely to be unemployed or have jobs of lower prestige and income than native-born citizens irrespective of their educational status. For example, in Iceland, the economic crisis in 2008 resulted in higher rates of unemployment among immigrants than among native-born Icelanders, and lower wages (Skaptadóttir, Wojtynska, & Ólafs, 2011). Furthermore, 22-year-old immigrants in Iceland are almost two times less likely to have finished college education than their Icelandic born peers (Garðarsdóttir & Hauksson, 2011).

Nonetheless, while quality language input is strongly related to word acquisition in the domestic language, its impact on second language word learning is not so simple. For example, the MEd study of Haraldsdóttir (2013) revealed that maternal educational level of her Ice2 preschool participants was not related with their Icelandic vocabulary skills or their ability to form past tense from verbs. Furthermore, Chen and colleagues (2012, p. 2013) found, that maternal education did not influence English vocabulary skills in either Chinese-English or Spanish-English fourth and seventh graders. The maternal educational level of their participants was high school for the Spanish children and college for the Chinese learners, thus the mothers of the latter group were more educated. As 98% of the Chinese parents and 92% of the Spanish parents spoke their first language at home the researchers suggest that the relation between maternal education and word skills was mediated by home language use. Chen and colleagues claim that future studies should compare vocabulary development of children whose parents have similar levels of education but different degrees of first language use at home.

2.3.2.2 Exposure to each language

Bilinguals’ word learning depends on whether one or more languages are spoken at home. If two different languages are spoken within the family, bilinguals most likely learn words for objects and experiences that are related to domestic life in both languages. For example, they will learn
words in both languages for *mummy* and *daddy*, and for *bath*, only if they hear these words in both languages. However, if the first language is the home language and the second language is spoken in kindergarten or at school, bilinguals tend to learn words related to domestic experience in the first language and words that are related to kindergarten or school in the second language. This is what Oller and colleagues have termed the distributed characteristic of bilinguals’ lexical knowledge (see Oller & Eilers, 2002; Oller, et al., 2007; Umbel, Pearson, Fernandez, & Oller, 1992). Bilinguals frequently know some words in one language but not the other, and vice versa (*singlets*), as well as knowing some words in both languages (*doublets*). In that way, bilinguals’ vocabulary acquisition is affected by the way or the circumstances in which they use their languages. Bilinguals’ vocabularies are also influenced by the amount of exposure they have to their two (or more) languages.

Among the youngest bilinguals, there is evidence suggesting that more exposure to one language results in more learning in that specific language, and the less time spent in the other language results in less acquisition in that language, demonstrating a model of competition between the languages. This was demonstrated in a study by Hoff and colleagues (2012) who found that high SES Spanish-English L2 children (with Spanish as their first language and English as their second language) around two years of age, could be classified in subgroups based on language(s) spoken at home. Their vocabulary increase in each language was in direct proportion to time spent using that language. This is in line with the findings of Quiroz and colleagues (2010) who showed that word skills in each language among L2 Spanish-English four- and five-year-old children (*n* = 50) were influenced by the amount of exposure to each language. The more English used at home, both in speech and reading, the more English word skills and less Spanish vocabulary, and vice versa. Oller and colleagues (2007) also found this model of competition among Spanish-English bilinguals in second and fifth grade (*n* = 620), demonstrating higher English vocabulary scores among participants who spoke both English and Spanish at home, than among those who only spoke Spanish at home. The former had more input in English as they were also using it at home. However, as both groups were using English much more than Spanish, and increasingly from grade two to grade five, their English vocabulary skills were improving and even reaching that of their English monolingual peers, but at the cost of their Spanish vocabulary skills.
2.3.3 The vocabulary gap between L1 and L2 children

As bilinguals distribute their times between two (or even more) languages they tend to have less exposure to each language than children who spend all their time using only one language. As a consequence, they tend to develop less vocabulary in each language than do monolinguals in their one language. Research indicates that the vocabulary gap between L1 and L2 learners starts to emerge early in life. For example, Hoff and colleagues (2012) showed that the scores of two and a half-year-old high SES Spanish-English L2 children on English vocabulary measures was similar to the scores of monolinguals aged one year and ten months. The bilinguals’ performance was, however, comparable to that of their monolingual peers when their scores in both languages were combined. This means that they were learning words at the same rate as the monolinguals, but their vocabularies were divided between their two languages.

Only three studies have been conducted on the vocabulary development of Icelandic second language learners. Haraldsdóttir (2013) conducted a MEd study on Ice2 children in their last two years of preschool ($n = 43$). The majority of these children had lived in Iceland since birth. Their word skills were tested with an Icelandic adaption of the PPVT measurement (V. Ólafsdóttir, 2011) (the same as used in the present study, see section 3.5.1.1). Furthermore, an abbreviated version of Ragnarsdóttir’s (2009) past tense test was administered. A comparison group of Ice1 children ($n = 111$) was included in the study. The researcher found that, already at this age, the mean scores of the Ice2 children on the PPVT were less than half of the mean scores of their Ice1 age peers. Even a greater difference emerged between the two groups in the ability to form past tenses of verbs.

Thordardottir and Juliusdottir (2012) explored Icelandic language proficiency among Ice2 first to tenth graders over a period of three years (Time 1, $n = 39$; Time 2, $n = 25$; Time 3, $n = 19$). Their participants had various first languages and a large range of length of residence in Iceland. To get an overview of the children’s general Icelandic language ability they completed an Icelandic version of the Test of Language Development for the primary and intermediate grades (TOLD-P and TOLD-I) (Simonardóttir, Guðmundsson, Skúlason, & Pétursdóttir, 1995; Simonardóttir, Guðmundsson, Skúlason, & Pétursdóttir, 1996). These tests, originally created in English (Hammill and Newcomer, 1997), have been translated, adapted, and semi-standardized for the Icelandic first language population and published. As an Ice1 comparison group was not implemented in the
study, the participants’ scores on the TOLD tests were compared to existing Ice1 norms. To assess the children’s vocabulary skills, a translated version of the PPVT (Dunn & Dunn, 2007) was used. This version has not been normed or validated with Icelandic learners (see section 2.2.1). The translation of the lower items was obtained from the Icelandic Diagnostic Center; a test that has been used extensively by clinicians in Iceland. The translation of the remaining items were completed by Thordardottir (2008). The findings of the study demonstrated that the increase in the children’s standard scores on both TOLD tests was not significant and the same applied to their raw scores on the translated PPVT test. However, a small but significant increase in mean performance on the vocabulary subtest of TOLD-P emerged between Time 1 to Time 2, and between all three Time points on TOLD-I. When the participants’ standard scores on TOLD were compared to Icelandic norms, most of them scored in a range indicating language impairment. However, the Ice2 children scored within normal range on a non-word repetition test, which the researchers interpret as an absence of such deficiency.

The vocabulary of Icelandic second language learners was the topic of my MA study (S. Ólafsdóttir & Ragnarsdóttir, 2010). The study included a group of Ice2 children in first (n = 49), second (n = 44), third (n = 38), and fourth grade (n = 42), with various first languages and length of residence in Iceland. Their Icelandic word skills were measured with an Icelandic adaption of the PPVT-4 (V. Ólafsdóttir, 2011) (the same as used in the present study, see section 3.5.1.1). Findings demonstrated that the first graders scored significantly lower than the second, third, and fourth graders, whereas no significant differences were detected between participants in the last three grades, when controlled for length of residence in Iceland. This indicates that Ice2 learners are not increasing their Icelandic word skills during these three school years. V. Ólafsdóttir (2011), on the other hand, found that this same Icelandic PPVT test differentiated very well between word skills of different age groups of Ice1 learners, aged four to eight (n = 117).

The MA Ice2 first to fourth grade participants, who were born in Iceland or had arrived to the country before two years of age, henceforth Ice2<2years learners (n = 60), were analysed separately as no score differences were detected within this group. The same pattern emerged as discussed above, as the Ice2<2years first graders had significantly lower scores than the Ice2<2years second, third, and fourth graders with no
significant difference between the other three age groups. From this it can be suggested that even Ice2<2years learners, who have spent all or most of their lives in Iceland, are not increasing their Icelandic word skills from grade two through grade four. The performance of the Ice2<2years children was then compared to that of Ice1 children in first grade (aged six) \( (n = 110) \) (participants of Ragnarsdóttir, et al., 2009). Findings demonstrated that the Ice2<2years first graders had significantly lower scores than the Ice1 age peers. Even the performance of the highest scorer among the Ice2<2years first graders was below the average score of the Ice1 peers (see Table 1). Moreover, the Ice2<2years learners in grade two, three, and four had mean scores lower than that of Ice1 learners in first grade (see Table 1).

**Table 1. Vocabulary raw scores of Ice2 and Ice1 children**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Number of participants</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice1 Grade 1</td>
<td>110</td>
<td>119.93</td>
<td>17.2</td>
<td>41</td>
<td>151</td>
</tr>
<tr>
<td>Ice2&lt;2years Grade 1</td>
<td>20</td>
<td>67.90</td>
<td>22.4</td>
<td>23</td>
<td>107</td>
</tr>
<tr>
<td>Ice2&lt;2years Grades 2, 3, and 4</td>
<td>40</td>
<td>104.03</td>
<td>23.1</td>
<td>41</td>
<td>152</td>
</tr>
</tbody>
</table>

These three studies with Ice2 learners are of great importance as they provide the first scientific evidence regarding the linguistic status of children who are learning Icelandic as a second language. They demonstrate that, unlike their monolingual peers, young students acquiring Icelandic as their second language make limited progress in the acquisition of Icelandic vocabulary in kindergarten and throughout their compulsory education. However, these studies have limitations that are important to address in future research. For example, in the study by Thordardottir and Juliusdottir (2012) the number of participants in each age group was very low and with large attrition, giving limited opportunities to explore progress over time. The main vocabulary assessment tool used in their research had been translated, but otherwise not adapted or standardized for the use with Icelandic-speaking children (see Skúlason, 2005). No information about the reliability, validity (see section 2.2.1), or the extent to which the test discriminates between the vocabulary knowledge of Icelandic speaking children, therefore, exists. The study of Haraldsdóttir (2013) and my MA
study (2010) were based on cross-sectional data and can, therefore, only give limited information about progress over time and the latter did not include monolingual Icelandic-speaking comparison groups for each age group. None of these three studies investigated whether the children’s vocabulary affected their reading comprehension or other aspects of academic achievement. In the current research, the aim is to address these issues.

Some studies suggest that if bilinguals continue to use both their languages far into adulthood they may overcome their initial deficiencies described above. For example, Oller et al (2007) showed that the English vocabulary gap between the L1 and L2 learners was much smaller among the fifth graders participating in their research than the second graders (half of whom spoke both English and Spanish at home). The researchers also refer to Bahrick, Hall, Goggin, Bahrick, and Berger (1994) who conducted a study with over 800 Cuban and Mexican immigrants in the U.S. The participants who continued to speak Spanish along with the English language had vocabulary recognition and oral comprehension scores in Spanish on a par or even better than Spanish monolinguals from age 30 onward. “Similarly, in English, the participants showed surprisingly good performance (nearly as good as monolinguals), on the same tests” (Oller, et al., 2007, p. 224). Bahrick and colleagues suggested that the bilinguals had overcome much of the early limited skills in English. It is important to point out, however, that Bahrick and colleagues’ research (1994) was carried out some time ago and does not include important information about their participants that could influence the interpretation of their data. For example, they did not specify their participants’ educational and social status, or their academic needs. They also maintained that their participants were fully capable of participating in English speaking society, whereas their English input was mostly generated from television and radio.

In spite of the fact that Oller and colleagues detected smaller vocabulary deficiency among L2 fifth graders than L2 second graders, their findings did not support their prediction that the L2 learners would eventually catch up with their L1 peers on vocabulary measures. In fact, as discussed in the next chapter, research findings differ regarding the influence of early vocabulary skills on subsequent growth of vocabulary, that is, whether L2 learners show a steeper vocabulary growth than L1 peers and manage to diminish the gap between the groups.
2.3.3.1 Vocabulary growth among L2 learners compared to L1 peers

Most studies on second language learners have demonstrated an increasing vocabulary difference between L1 and L2 learners across the school years. For example, Droop and Verhoeven (2003) compared the vocabulary skills of Dutch L2 third and fourth graders ($n = 122$) with that of L1 age peers ($n = 143$). Their L1 participants were of both high and low SES, whereas the L2 children all belonged to low SES families with Turkish and Moroccan as their first languages. The findings revealed that the vocabulary of the L2 children lagged behind that of the L1 children during the two years of the study, with a widening gap between the two groups. Three years later a cross-sectional study of Verhoeven and Vermeer (2006), consistently revealed an increased difference between native Dutch third to sixth graders and Mediterranean minority age peers on vocabulary measures. These L2 and L1 children were predominantly from low SES families.

A longitudinal study was carried out on Norwegian L2 children ($n = 90$) (Lervåg & Aukrust, 2010) who had Urdu as their first language. The central aim of the research was to explore the development of vocabulary among this group of children and investigate its role in their acquisition of literacy. The results showed that at the start of the study when the children were in grade two, the L2 children had a significantly smaller vocabulary than their L1 peers ($n = 198$). They also demonstrated slower word acquisition over the 18 months the study lasted. These two groups, on the other hand, demonstrated comparable decoding skills. This is consistent with Lesaux and colleagues’ finding (2010) that Spanish-English L2 learners ($n = 87$) in the US developed less vocabulary skills from grade four to five than did their L1 peers, both in their first and second language. Both groups, however, demonstrated equivalent decoding skills. The L2 participants were attending urban schools, and of low SES. In fact, early literacy learning, largely consisting of learning to decode, is manageable by the vast majority of children. This is not dependent on vocabulary knowledge and can be achieved with a very low vocabulary threshold.

The above quoted studies included L2 learners of low SES, but similar pattern has also been detected with L2 children of high SES. Roessingh and colleagues (2008; 2009; 2003; 2005), for example, followed L2 learners in Canada, mostly of high SES, throughout the school years and up to grade 12. They demonstrated that L2 learners increased their vocabulary skills to a lesser extent between grades than their L1 peers, revealing a widening gap.
between the groups. The L2 learners could easily develop basic word knowledge of high-frequency words but their poor word knowledge applied to low-frequency academic words. The vocabulary lag did not only emerge in their receptive word tasks but also in their productive language skills, as they were over dependent on high-frequency words and the gap in free productive vocabulary skills (see section 2.2.4) increased with each year over a period of eight years, up to grade nine (Roessingh & Elgie, 2009). However, the researchers have found that those L2 learners who have participated in a specifically designed ESL (English as a second language) program including strong focus on academic word learning, have bolstered their word skills significantly, which in turn has given a cause for optimism about their future academic prospects. The researchers conclude that without support, majority of L2 learners, irrespective of their age of arrival, will not meet academic demands in college and University. Also in Canada, Crossman and Pinchbeck (2012) showed with their study that well designed curriculum made accelerated language learning possible, even after grade 12.

In the Netherlands instructional programs have been launched (Appel & Vermeer, 1997) to speed up word acquisition of L2 learners and emphasis on explicit vocabulary instruction was implemented in the Curriculum for Primary School Education in 2006 (Ministerie van Onderwijs, 2006). It is possible that Dutch L2 children have benefited from an increased focus on word learning as, in contrast with the above quoted findings of an older study by Droop and Verhoeven (2003) with third to fourth graders, and Verhoeven and Vermeer (2006) with third to sixth graders, Verhoeven and van Leeuwe (2011b) detected a diminishing vocabulary gap between Dutch L2 learners (n = 394) and L1 peers (n = 1,293) from grade one through six. However, the vocabulary difference between the two groups did not disappear. The researchers concluded that there is a need for more and continued attention to word learning throughout the elementary school years (Verhoeven & van Leeuwe, 2011b, p. 1816). The L2 and L1 participants attended schools of mixed ethnic lower middle-class population, the majority of lower SES, the L2 children of Moroccan, Turkish, Surinam, and Antilles origin.

The studies quoted above tracked word acquisition of L2 learners with English, Dutch, and Norwegian as their second language. Each of these studies demonstrated a remaining vocabulary gap between L1 and L2 learners of the same age, even though the L2 children managed to improve
Their word skills each year. It is important to take into account the limitations of the above quoted studies. The participants in the study of Lervåg and Aukrust (2010) were very young and were followed over only 18 months, from grade two to grade three. Droop and colleagues (2003), and Lesaux and colleagues (2010) tracked development over only two school years, however with older learners, the former from grade three to four and the latter from grade four to five. Verhoeven and Vermeer (2006), on the other hand, tracked development over four years, from grade three and into the middle school years, up to grade six. All these studies tapped receptive word skills. Verhoeven and Vermeer also included writing tests in their study, however, which did not tap into free productive academic word skills, the use of tier 2 words, fundamental for the ability to express ideas and thoughts in expository texts (see section 2.9). Roessingh and colleagues included both receptive and free productive academic word skill measurements and followed their participants throughout the elementary school years. They, however, only compared learners' performance between grades, and therefore they did not investigate whether the second language learners demonstrated a different shape and rate of vocabulary growth. Their approach also did not give information about individual trajectories, whether those who started with larger vocabularies increased these at a faster rate than those with smaller initial vocabularies. Moreover, for the relevance of the current study, it is important to note that the languages, English, Dutch, and Norwegian, are inflectionally less complex than the Icelandic language in which verbs, nouns, adjectives, numbers, and pronouns are inflected with changes in both stems and endings. It can be suggested that Icelandic words are therefore more difficult for second language learners to recognize, understand, and subsequently add to their lexicon. However, as mentioned before, only one longitudinal study has been carried out to investigate the development of Icelandic language skills of Icelandic second language learners (described in section 2.3.3) (Thordardottir & Juliusdottir, 2012). On the basis of their findings, the researchers maintain that it may take young Icelandic L2 learners longer to develop general linguistic proficiency that are comparable to that of their monolingual peers, than L2 learners of other languages with simpler grammatical structure, such as English. In support of their argument, the researchers refer to Collier (1989) who demonstrated through her studies that it takes English L2 learners five to seven years, on average, to reach the English proficiency level of their L1 peers. In comparison, the only Ice2
participants who managed to reach the appropriate Ice1 age norm on TOLD in the study by Thordardottir and Juliusdottir had close to eight years of residence in Iceland, on average. These participants, however, made no gains on the translated version of the PPVT during the three years of the study. But, as there was no monolingual comparison group included in the study it is difficult to tell whether this result reflects the actual rate of development within this group or has something to do with the psychometric properties of the test itself (previously translated and tested with Ice1 children, Thordardottir, 2008) (see section 2.2.1).

In my cross-sectional MA study (described in section 2.3.3) (S. Ólafsdóttir & Ragnarsdóttir, 2010) on the word skills of Ice2 first to fourth graders there were no significant differences between PPVT vocabulary scores of second, third, and fourth graders when the effect of length of residence in Iceland was controlled. These results suggest that, unlike Ice1 children (V. Ólafsdóttir, 2011), Ice2 learners may not be enlarging their Icelandic word skills during the first few primary school years. This finding indicates a widening vocabulary gap between Ice1 and Ice2 learners during the first four years of schooling, although further studies, employing a longitudinal design and growth model analysis, are needed to confirm these findings.

As described above, a great number of studies indicate that many bilingual learners tend to suffer from a continuous lag on vocabulary measures throughout the school years compared to monolingual peers (see also Egelund, 2012; Centre for Literacy of Quebec, 2008; National Center for Educational Statistics, 2012). However, as discussed in the next chapter, evidence suggests that the picture may be more complicated than that. For example, research indicates that bilingual children with related first and second languages may have an advantage over other bilinguals, with more remotely related first and second languages. Thus, connecting cognates between the languages may well facilitate word learning.

2.3.4 Influence of cognates in first and second language
A cognate is a word or a word part that resembles, phonologically and/or orthographically, a word in another language. The English language shares many cognates with languages such as Spanish, French, Italian, and Portuguese. These cognates are mostly derived from Latin (Vadasy & Nelson, 2012). The Icelandic language shares numerous cognates with Faroese, Norwegian, Swedish, and Danish. All these languages, however,
belong to the Germanic languages and are part of the Indo-European language family. Examples of cognates between all these languages are the English words *shirt* and *skirt*. The former derives from the Old English word *sċyrte*, and the latter from Old Norse *skyrta*, which is the same as *skyrta* in Icelandic, *Schürze* in German and *Schort* in Dutch. However, both *shirt* and *skirt* derive from Proto-Germanic *skjurjön* (Oxford University Press, 2013). These Germanic languages together belong to the Indo-European language family, which includes most European languages, spoken in both West- and East Europe. Almost all the European languages are thus related. Only the Basque language spoken in the west Pyrenees, in the northeast of Spain and the southwest of France, and Maltese, which is a Semitic language, are unrelated to other European languages (see Crystal, 1997, pp. 298-303).

When words in two languages share a phonological and orthographic form, the learner may automatically detect these resemblances to facilitate memorization. According to the Parasitic model of vocabulary development (see Hall, 2002), even when a young child sees or hears a word in another language for the first time he automatically and unconsciously exploits prior word knowledge in his first or second language in order to understand and learn the word. L2 students are thus likely to learn words in a second language with less difficulty if the words share cognates with the first language (August, et al., 2005; Holmes & Guerra Ramos, 1995). This was demonstrated by the research conducted by Chen and colleagues (2012). The researchers found that Spanish-English bilinguals in grades four and seven were able to connect cognates between their two languages to facilitate the learning of English words, whereas this learning strategy was not available for the Chinese-English participants. The Spanish-English outperformed the other group on cognate items on a PPVT vocabulary test, whereas the two groups had similar scores on non-cognate items on same test. Moreover, while both groups had significantly lower scores than monolingual peers, the gap was smaller in cognate vocabulary between the Spanish-English and the monolinguals than that between the Chinese-English and the monolinguals on same test items. “In other words, cognate awareness is effective in reducing the gap in vocabulary development for Spanish-speaking ELLs” (Chen, Ramirez, et al., 2012, p. 2011).

In spite of the fact that the transfer of cognates between languages is often automatic and unconscious among young learners, recognizing cognates is sometimes not that easy. For example, the similarities between the Spanish word *amoroso* and the English word *amorous* are easily detected, whereas the
connections between the Spanish word *oscuro* and the English word *obscure* are more difficult to recognize (Dressler, 2000). Hancin-Bhatt and Nagy (1994) suggest from their findings that the ability to recognize cognates increases with age. Moreover, Jimenez, Garcia, and Pearson (1996) found that Spanish-English L2 sixth and seventh graders who were proficient English readers were better able to recognize and transfer L1-L2 cognates than the less proficient ones. However, cognates can occasionally lead to a false, negative transfer. If the phonologically and/or orthographically related words in both languages have totally different meanings, it can make word learning more difficult as learners may guess wrongly about their meanings or confuse word meanings between languages (Hall, 2002; Nation, 1990; Odlin, 1989).

Haraldsdóttir (2013) did not find a significant difference between Ice2 kindergartners with East-European \((n = 25)\) and non-European \((n = 14)\) first languages on the PPVT vocabulary test. The majority of her participants had lived in Iceland from birth. In my MA research (S. Ólafsdóttir & Ragnarsdóttir, 2010) (see section 2.3.3) Ice2 primary school children with East European first languages \((n = 35)\) scored higher on same PPVT test than those Ice2 learners whose first languages were not European \((n = 23)\). This kind of advantage, however, emerged only among those who were either born in Iceland or had arrived before the age of two.

Comparing these Icelandic findings with those obtained with English second language learners it is important to take into account that English words that share cognates with Spanish words are predominantly of low frequency, academic words falling into both tier 2 and tier 3 category in the English language (see section 2.1.1.1). These English words represent however high-frequency Spanish words. On the other hand, cognates between the various Indo-European languages are more of high frequency, words that belong to basic common human existence (e.g. numbers, family members, daily life) (Fortson, 2005). The findings of Haraldsdóttir and my MA study suggest that it is rather in the later stages of Icelandic second language learning that an advantage of this kind emerges among Ice2 learners. If cognates between Indo-European languages and the Icelandic language give Ice2 learners with European first languages a lasting advantage over the non-European Ice2 learners is still not known. In fact, the European languages not only share many cognates but also do they have similarities in their grammatical structure. The current study will reveal if a difference exists between European and non-European middle school
learners on vocabulary and reading comprehension measures, and how such gap develops over these years.

Although many studies have demonstrated a positive influence of recognizing cognates from the first language on second language word learning, longitudinal studies that compare L2 vocabulary acquisition of different language groups based on the relation between their first and second language are scarce, to say the least.

### 2.3.5 Reciprocal relation between first and second language vocabulary acquisition

Even when there is little correspondence between the first and the second language, first language proficiency can provide a solid foundation for second language acquisition, (Cummins, 1979; Roessingh, 2008). The theory of linguistic interdependence holds (Cummins, 1979) that when the L2 learner has learned an abstract, symbolic or an academic word in his first language, learning the same word in the second language for the same concept is easier, even when the words are phonologically or orthographically different. This argument has important educational implications. It implies that learning second language vocabulary should be easier for older children than for younger ones, as the older have more first language skills and more developed underlying cognitive ability (Cummins, 1979; Roessingh, 2008).

#### 2.3.5.1 Age of arrival and second language vocabulary acquisition

As discussed above, a number of research findings have demonstrated a time-on-task competition model between the languages among the youngest bilinguals. That is, that each language is learned in proportion to the amount of exposure to that language, with little or no cross-language influences (see section 2.3.2.2). This applies to many simultaneous bilinguals, who learn two languages concomitantly, or bilinguals who have little knowledge in their first language when they start to learn the second language. Young second language learners can, however, acquire basic, daily vocabulary and general language skills (BICS) very quickly and are able to attain native like pronunciation, fluency and correct speech (see Arnbjörnsdóttir, 2007). This applies to a large number of L2 children, who start to learn the second language at an early age when they are still in the period of first language acquisition. Their basic vocabulary skills are sufficient for requirements in
the first years of school. However, beginning in the middle school years when an increasing number of complex words need to be understood and learned, their vocabulary skills frequently start to lag behind that of their monolingual peers (Roessingh, 2008). Due to their less developed word skills, young arrivals tend to develop academic vocabulary skills at a slower rate than children who arrive to a new home country at an older age, especially for words that represent an abstract thought, such as symbolic words and metaphors (see Roessingh, 2008; Roessingh & Elgie, 2009). These children are at great risk of lacking linguistic skills necessary for higher order thinking, which can have serious consequences for their academic achievement. When these children also belong to families of low educational status, they are at an even greater risk.

Thus, positive transfer from first to second language vocabulary appears to apply mainly to L2 learners who have already established substantial first language word skills (see Snow & Kim, 2007). Roessingh (2008), for example, found that older L2 learners arriving in English speaking Canadian schools could use first language vocabulary skills as an important foundation in their second language word acquisition. So, even though they had smaller second language vocabularies than L2 age peers who had arrived at a younger age, they were capable of transferring meaning from their first language vocabularies to their second language word acquisition. An example of this kind is when a child has learned the word *advantage* in his first language, the concept behind this abstract word, and the situations in which it can be used, learning the equivalent second language word becomes easier. He only has to switch words. For that reason, older arrivals can learn new words in the second language at a faster rate than younger arrivals. However, the older arrivals are in a hurry, as they have to “beat the academic clock” (Roessingh, 2008, p. 89). While they spend their time learning the new language their age peers proceed in their studies, increasing their academic vocabulary and proficiency. The younger arrivals, on the other hand, tend to learn the second language words at a slower rate than their L1 peers and the older arriving L2 learners, and therefore need support from the very start to keep up with educational demands.

An advantage of older arriving learners was, however, not detected in the study of Thordardottir and Juliusdottir (2012) with Ice2 learners (*n* = 39) from age six to fifteen. When matched on length of residence in Iceland, participants who had arrived relatively late to Iceland (*n* = 8, age of arrival, ranging from 11.5 to 14 years) received lower raw scores and standard
scores on both vocabulary measures used than those who had arrived at an earlier age \((n = 8\), age of arrival ranging from six to nine years). Thus, the performance of the later arrivals was poorer than that of the younger arrivals, both in absolute scores and as compared to age appropriate norms. One possible reason for these results, the researchers suggest, is that these older arrivals were struggling at school with too difficult academic texts. Another explanation might lie in the inflectional complexity of the Icelandic language and the fact that the language used in learning materials in the upper grades tends to be more formal and complex than in earlier grades. However, it is important to take into account the very small number of participants, and that Thordardottir and Juliusdottir only compared different age of arrival groups. They did not investigate the influence of age of arrival on the rate of vocabulary growth, i.e. if those who arrived at an older age developed their Icelandic vocabulary skills at a slower or a faster rate than the younger arriving peers, or if age of arrival did not influence the rate of vocabulary growth.

### 2.4 The educational context of Ice2 learners

The findings of the only two studies that have explored Icelandic vocabulary acquisition among Ice2 learners strongly indicate that their word skills lag significantly behind that of their monolingual Icelandic-speaking peers and that they may even acquire their Icelandic language skills at a slower rate than learners learning English as a second language. As discussed above, the grammatical complexity of the Icelandic language has been suggested as a possible reason for these results. However, there is also a possibility that the educational context of Icelandic second language learners may play a role in the apparently limited vocabulary improvement among Ice2 school aged learners. For example, in the National Curriculum, under the specific sub-chapter on Icelandic as a second language \((Aðalnámskrá grunnskóla: almennur hluti 2011 og greinasvið 2013)\), it is stated that Ice2 learners should be capable of asking the meaning of words and idioms and be able to communicate in class, to understand the language in various settings, express themselves and communicate, taking account of social rules (pragmatics). No explicit reference is made to the importance of effective, evidence based methods in building up Icelandic vocabulary of Ice2 learners, involving engaging word-focused activities and frequent rehearsals (e.g. Laufer, 2009). This limited emphasis on vocabulary in the Icelandic curriculum is reflected in the teaching material intended for Ice2
learners. Special subject books for this group of learners (e.g. Gísladóttir, 2003; Jóhannsdóttir, 2001; Kristinsdóttir, 2010) include few or no rehearsals of target words. Furthermore, vocabulary learning was first specified in 2013 as a specific topic of study in training courses aimed at teachers of Icelandic as a second language (Ísbrú, 2013).

2.5 Summary

It is clear from the discussion above that there is longstanding evidence that bilingual learners tend to develop less vocabulary in each language than monolinguals in their only language, resulting in a widening gap between the two groups. For example, Lervåg and Aukrust’s Norwegian L2 participants (2010) with Urdu as their first language were found to have both smaller vocabularies and slower vocabulary growth between second and third grades than L1 children at the same age. The majority of studies with second language learners track development with very young learners and/or over a short period of time. The longitudinal studies of Roessingh and colleagues expand, however, over all the elementary school years (2008; 2009; 2003; 2005). By comparing achievement levels between grades, they have demonstrated an increasing gap between L2 learners of English in Canada (mostly of high SES) and L1 peers, in receptive word skills and also in free productive academic vocabulary use. Roessingh et al., however, did not take an advantage of the opportunity offered by longitudinal data to investigate growth directly, as growth model analysis was not used. For that reason they were unable to explore the both the shape and the rate of the students’ vocabulary growth. Their approach also did not give information about individual trajectories, i.e. if those with the highest initial vocabulary scores increased these at a faster rate than those with smaller initial word skills. Moreover, growth model analysis would have allowed them to detect factors that contribute to the rate of vocabulary growth, rather than on the achievement in different grades, i.e. influencing variables that drive change over time.

Research has shown that ESL instructional programs, including strong emphasis vocabulary activities, have significant positive effects on L2 word acquisition and academic prospects of L2 learners (Roessingh, 2008; Snow & Lawrence, 2011). For example, following an increased focus on explicit vocabulary instruction in the Dutch school system, the pattern has changed from a widening (Droop & Verhoeven, 2003) to a narrowing vocabulary gap (Verhoeven & van Leeuwe, 2011b) between Dutch L1 and L2 children.
Little emphasis has been placed on word learning in the Icelandic curriculum (*Abalmánskrá grunnskóla: almennur hluti 2011 og greinasvið 2013*). The only three studies that exist on the vocabulary acquisition of Icelandic second language learners suggest a large and widening vocabulary gap between Ice1 and Ice2 children. Haraldsdóttir (2013) demonstrated that her Ice2 participants in kindergarten had mean scores on the PPVT less than half of the mean scores of their Ice1 age peers. In the study by Thordardottir and Juliusdottir (2012) Ice2 participants made a modest improvement over three years on the vocabulary subtest of TOLD, but no increase in vocabulary scores was detected on a translated version of the PPVT, remaining significantly well below that of L1 learners. Furthermore, my MA study (S. Ólafsdóttir & Ragnarsdóttir, 2010) demonstrated that Icelandic second language learners in grades one to four who had lived in Iceland at least from the age of two, obtained significantly lower scores on the Icelandic PPVT vocabulary test than Icelandic monolingual speakers in first grade (participants of Ragnarsdóttir, et al., 2009). Furthermore, when controlled for length of stay in Iceland, the difference between Ice2 second, third, and fourth graders was not significant. These results indicate that Ice2 learners, unlike their Ice1 peers (V. Ólafsdóttir, 2011), may not be increasing their Icelandic word skills during their first few primary school years.

However, as discussed in section 2.3.3, these studies have limitations that need to be addressed. The study by Thordardottir and Juliusdottir (2012) included a low number of participants, making growth model analysis not applicable. Their main vocabulary measure had not been adapted or standardized for the use with Icelandic-speaking children (see section 2.2.1) and there was no monolingual comparison group. Furthermore, no account was made of possible influences of important background variables, such as the participants’ native language and their SES status (see sections 2.3.4.1 and 2.3.2.1 for detailed discussions about the effect of these variables). Although some of these shortcomings were met in my MA study (2010) and the research of Haraldsdóttir (2013), these studies employed a cross-sectional design. Consequently, we have little overview of how different aspects of the vocabulary knowledge of Icelandic second language learners develop. Such information is important as research suggests that attaining good knowledge of second language vocabulary is a complex process that is influenced by a host of factors that often differ across cultures, such as the structure of the language to be learned and the
educational context in which the learning takes place. For example, as discussed in section 2.3.3, the Icelandic language has a complex grammatical structure and may, therefore, be particularly difficult for second language learners to attain (Thordardottir, et al., 2002). Direct generalizations of research findings regarding second language learning across cultures may, therefore, be misleading.

One of the central aims of this study is to explore the development of Icelandic vocabulary skills of middle school aged second language learners in Iceland. The research will seek to address limitations presented in previous research in this field by employing a longitudinal design, tracking development from grade four and all the way up to grade eight, with a carefully matched monolingual comparison group. A range of vocabulary measures will be used to tap various dimensions of vocabulary knowledge that have been developed or adapted specifically for use with Icelandic-speaking participants. Growth model analysis will make it possible to compare Ice2 learners with their Ice1 peers in the shape and the rate of vocabulary growth. Moreover, this approach makes it possible explore individual trajectories, to detect the relations between initial word skills and the rate of growth. The influence of important background variables will be explored, on initial vocabulary skills and the rate of growth, such as the educational status of mothers, the relatedness of the Ice2 participants’ first languages with the Icelandic language, first language proficiency, and age of arrival. The study will also include a measurement tapping academic vocabulary use in a writing test, which has indeed rarely been included in literacy studies. Recalling a word for active use is the ultimate word knowledge and the academic tier 2 words play a key role for the ability to use language for learning. Transitioning over to an academic language use, both for understanding and productive use, becomes pivotal from grade four, for the acquisition and development of academic literacy (cf. Roessingh, Elgie, & Kover, 2013).

In the light of the strong relationship between word skills, reading comprehension, and general academic competence, providing detailed information about vocabulary acquisition is essential. Not only does the reader need to understand the majority of words in a text to understand its message (Laufer, 1989), he also needs to be able to detect the meaning of unknown words and eventually add them to his receptive and productive lexicon. Thus, reading comprehension both requires solid vocabulary and contributes to its development (e.g. Nagy, 2005; Stanovich, 1986).
Moreover, written texts include far richer variety of words than spoken language and, therefore, reading comprehension is one of the most important influencing factors in word acquisition. Vocabulary development is thus greatly mediated by literacy skills, and vice versa.

2.6 What is reading comprehension?

Understanding individual words in a written text is a prerequisite for its comprehension. A solid understanding of the text as a whole is also necessary to detect the meaning of unfamiliar words, which then leads to word learning. Reading comprehension thus both requires vocabulary skills and contributes to its development, indicating a reciprocal relationship.

Reading comprehension, however, entails much more than the understanding of single words in a text. The reader needs to be able to decode each word accurately and efficiently and understand its meaning within the context in which it appears, but he also has to be able to make inferences by going beyond what is stated explicitly in the text, making use of prior knowledge and experiences (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Basaraba, Yovanoff, Alonzo, & Tindal, 2013). Reading comprehension also requires a good sense of chronological order of facts and events, knowledge of text structure, and the ability to monitor on-going understanding and repair reading strategies if necessary (Ruffman, 1996). Thus, reading comprehension is a complicated process that challenges many kinds of cognitive and linguistic skills and various types of knowledge which “… can and should embrace all types of thinking, evaluating, judging, imagining, reasoning, and problem-solving” (Gates, 1949, p. 3; Hoover & Gough, 1990).

2.6.1 The simple view of reading

One of the most cited frameworks for the study of reading comprehension is The Simple View of Reading (SVR) (Gough & Tunmer, 1986; Hoover & Gough, 1990). The Simple View of Reading holds that the multiple aspects of reading comprehension can be divided into two separate parts: decoding and linguistic comprehension (reading comprehension = decoding x linguistic comprehension). Decoding is the ability to translate printed letters into sounds and link them together to make words. Linguistic comprehension (or listening comprehension), on the other hand, is an understanding of language in the form of vocabulary, sentence and text comprehension. These two key components are both indispensable aspects of reading comprehension as
reading is difficult with poor decoding skills and of little use without comprehension (Hoover & Gough, 1990). Neither of these skills, however, is sufficient on its own for solid comprehension to occur. Although these two aspects of reading comprehension are related, research indicates that learners may vary in their aptitude to acquire these skills. For example, while some learners with good language comprehension have difficulties in acquiring decoding skills, as is the case with individuals with dyslexia, others can master decoding skills without having good linguistic comprehension (see Catts, Adlof, & Weismier, 2006). In such cases, individuals may be suffering from specific reading comprehension difficulties (Cain, 2010; Catts, et al., 2006; Hoover & Gough, 1990). This distinction between decoding and linguistic comprehension is widely used as a theoretical framework for reading and seems to be applicable both to L1 and L2 learners (Chen, Geva, & Schwartz, 2012; Hoover & Gough, 1990; Verhoeven & van Leeuwe, 2011b). Furthermore, research suggests that decoding and linguistic comprehension skills are to some extent based on different underlying variables. Thus, while phonological skills and letter knowledge are fundamental for decoding skills, vocabulary is a key component of linguistic comprehension (Chen, Geva, et al., 2012, p. 1799; Biemiller, 2001; 2003; Biemiller and Slonin, 2001).

2.6.1.1 The changing relationship between decoding skills, linguistic comprehension and the development of reading comprehension

In the first stages of reading, when the main goal is to practice decoding skills, the texts that children are presented with are usually simple and easily understood. For that reason, individual differences in reading performance in the first years of schooling are more influenced by decoding skills than of linguistic comprehension (Catts, Hogan, & Adolf, 2005; Hoover & Gough, 1990). However, with age and more decoding proficiency, reading texts become richer and more complicated, and the goal of reading changes from ‘learning to read’ to ‘reading to learn’. Thus, linguistic comprehension takes over as the dominant influencing factor on reading comprehension (Aaron, Joshi, Gooden, & Bentum, 2008; Hoover & Gough, 1990), and its impact increases in strength with age. This changing relationship was, for example, demonstrated by Catts and colleagues (2005) who found that the amount of variance uniquely explained by word recognition in reading comprehension was 27% among eight-year-olds, 10% among ten-year-olds, and 2% among 14-year-olds. The corresponding values for listening comprehension, on the
other hand, were 9, 10 and 30% among 8-, 10- and 14-year-olds, respectively. Similar results have been obtained when vocabulary scores have been used as a proxy to linguistic comprehension. For example, results from the PISA 2009 assessment demonstrated that while decoding skills explained 7% of the variance in reading comprehension among 15-year-old Danish learners, vocabulary knowledge uniquely explained 39% of the variance in that same measure (Egelund, 2012, p. 83).

2.6.2 The effect of orthography on the acquisition of basic decoding skills

The development of decoding and reading fluency differs from one language to another, depending on their alphabetic writing systems. Alphabetic writing systems of various languages differ with regard to the consistency between spelling-to-sound relations. In shallow or transparent orthographies there is a regular correspondence between graphemes and phonemes. Examples of such languages are Spanish, Italian, and Finnish (Seymour, 2005; Ziegler et al., 2010). In these languages the majority of graphemes (individual letters or letter combinations) represent the same phoneme in the spoken language. Deep or opaque orthographies, on the other hand, have an inconsistent, bidirectional, one-to-many mappings between graphemes and phonemes. The orthographies of English and Danish, for example, are opaque with each grapheme representing a number of different phonemes (Ziegler, et al., 2010). A number of cross-language studies suggest that word decoding is more complicated and more difficult to obtain in deep orthographies than it is in shallow orthographic languages (Seymour, Aro, & Erskine, 2003; Ziegler, et al., 2010; Þráinsdóttir, Sigurðsson, & Lund, 2006). For instance, a large cross-language study including 16 West-European countries (Seymour, et al., 2003) demonstrated that the children who were learning to read in shallow orthographic languages, such as Finnish and Italian, developed basic reading skills significantly faster than children who were learning to read in more deep orthographic languages, such as English.

Hardly any cross-language comparisons that include children learning to read Icelandic have been made. One important exception is a large cross-linguistic study directed by Seymour and Duncan (2005). The aim of the study was to investigate the effect of orthographic depth on early reading development. Participants were children speaking eight different West-European languages (Þráinsdóttir, et al., 2006) whose phonological and
decoding skills were followed up through the first year at school. The results showed that, despite the fact that the Icelandic language has a rather transparent orthography (Lund, Sigurðsson, & Práinsdóttir, 2005; Práinsdóttir, et al., 2006), Icelandic children develop decoding skills at a slower rate and with more individual variation than do children learning to read orthographies that are classified as transparent, such as Finnish and Spanish (Lund, et al., 2005; Seymour, 2005; Práinsdóttir, et al., 2006). Moreover, the decoding performance of the Icelandic readers was greatly hampered by inconsistencies between graphemes and phonemes, and the children had more difficulties reading unfamiliar words (pseudo-words) than their age peers speaking languages of similar orthographic consistency (Práinsdóttir, et al., 2006). Given the relative transparency of the Icelandic orthography, these results are surprising and raise questions about the quality of early reading instruction in Icelandic schools (Birgisdóttir, 2011). However, the syllabic structure of Icelandic is quite complex with many words containing consonant clusters in word initial, medial and final positions, which may affect early decoding skills (Birgisdóttir, 2011; Lund, et al., 2005; Práinsdóttir, et al., 2006). For example, a recent longitudinal study by Birgisdóttir (2010) demonstrated that the decoding skills of first and second graders were not only affected by the orthographic complexity of words, but also by the complexity of their phonological structure. By the end of grade three (2011), however, neither factors seemed to pose any difficulties for the majority of the participants, although the poorest decoders still had difficulties in decoding such words, demonstrating reading fluency comparable with the best first graders. The researcher, on the other hand, did not explore the extent to which decoding skills influenced reading comprehension and its rate of growth, and she did not follow her participants into the middle school years.

The above results suggest that by grade three, most Ice1 learners have acquired sufficient decoding skills to enable them to divert their attention and effort to understanding the content of what they are reading. Little is known, however, about the development of early reading skills of Ice2 learners, nor how technical reading skills of both Ice2 and Ice1 children influence the rate of reading comprehension growth, i.e. if those who are still struggling by grade three or four have prevailing difficulties in reading comprehension.
2.6.3 The acquisition of phonological and decoding skills among L2 and L1 learners

One of the most important influencing factors in the acquisition of decoding skills is phonological awareness (Björnsdóttir, Einarsdóttir, & Simonardóttir, 2003; Hulme & Snowling, 2010; Scarborough, 1998). Phonological awareness refers to the ability to manipulate and reflect on the phonological aspects of language. Children with poor phonological awareness already in the preschool years have demonstrated long-lasting difficulties in both the development of decoding skills and also in reading comprehension (Hulme & Snowling, 2010; Scarborough, 1998). An Icelandic longitudinal study (Einarsdóttir, Björnsdóttir, & Simonardóttir, 2011; Björnsdóttir, Einarsdóttir, & Simonardóttir, 2013), for example, demonstrated that phonological awareness measured at five years of age, among Icelandic children, predicted academic development throughout the elementary school years and how well the participants fared educationally up to age 18.

Research suggests that children develop phonological awareness in relation to the demands implied by the phonological and orthographic structure of their language (Bialystok, 2007). This was demonstrated in a much-cited study by Caravolas and Bruck (1993), in which the phonological skills of Czech- and English-speaking monolingual beginning readers were compared. According to the authors’ hypothesis, as the Czech language has a much higher frequency of complicated consonant onset clusters (e.g. vzdělání) than English, and also a more regular spelling pattern, it should be easier for the young Czech learners to identify individual consonant phonemes in word-onset positions than for their English peers. Their predictions were confirmed; the Czech participants were significantly more proficient at breaking up consonant onset clusters in spoken non-words than were the English participants. Similar findings were obtained with Icelandic speaking children by Birgisdóttir (2003).

The above findings indicate that quite specific differences in phonological input may result in equally specific differences in the acquisition of phonological awareness. As bilingual children are exposed to more than one language, the question arises whether they may acquire more diverse phonological skills than monolinguals (Bialystok, 2007). Research indicates that some bilinguals have indeed shown better performance in some phonological tasks than monolinguals (Lesaux & Siegel, 2003). Their advantage, however, depends both on the phonological and orthographic
structure of their two languages. For example, Bialystok, Majumder, and Martin (2003) found that Spanish-English kindergartners and second graders outperformed monolingual peers on a phoneme counting task including English words. The researchers suggested that the Spanish-English children benefitted from the regularity of the Spanish orthography, providing them with better understanding of sound-symbol correspondences and more advanced phonological awareness. However, the researchers detected no differences between the monolingual English-speaking participants and the Cantonese-English and the French-English bilingual participants on the same task. It seems that these two groups did not benefit from the Cantonese and French orthography, as both have little phoneme-letter consistency.

However, while some bilinguals have outperformed monolinguals on some phonological tasks and sometimes other bilingual groups as well, their advantage tends to be largely confined to their first few years of formal schooling and then gradually disappear (Bialystok, 2007). Thus, rather than bilingualism itself, research indicates that the dominating influencing factors on the acquisition of phonological awareness and decoding skills are individual differences in reading ability (Geva, Wade-Woolley, & Shany, 1997), the quality of instruction (Bialystok, 2007, p. 69), and the orthographical system of the language (Bialystok, 2007, pp. 70-71). Thus, in spite of the fact that knowledge of two or more languages can in some cases enhance some aspects of phonological awareness, decoding skills are largely dependent on the complexity of each orthography and “… need to be relearned with each new writing system…” (Bialystok, 2007, p. 45; Bialystok, McBride-Chang, & Luk, 2005; Manis, Lindsey, & Bailey, 2004).

Despite an initial advantage demonstrated by some groups of bilinguals in solving phonological awareness tasks, research indicates that L2 and L1 children acquire decoding skills with comparable ease (see August & Shanahan, 2006; Droop & Verhoeven, 2003; Lervåg & Aukrust, 2010; Verhoeven & Vermeer, 2006). This has been demonstrated in a number of studies with second language learners of English, Dutch, and Norwegian with various first languages. For example, Droop and Verhoeven (2003) found that Dutch L2 third to fourth graders (n = 122) who were of Moroccan and Turkish origin and of both high and low SES, were on a par with L1 peers (n = 143) in decoding skills, and performed even better than L1 learners of low SES. With English L2 learners, Lesaux and Siegel (2003) found that by second grade the L2 children had acquired decoding and
spelling skills in the English language on a par or even better than L1 age peers. These findings are consistent with a study conducted by Oller and colleagues (2007) that demonstrated comparable basic reading skills of English L2 second to fifth graders (n = 620) of Spanish origin as to that of L1 learners’ of the same age. Lervåg and Aukrust (2010) also found no difference between the decoding skills of Norwegian L2 learners in second to third grade, with Urdu as their first language (n = 91) and L1 age peers (n = 198). The same applied to the results obtained by Nakamoto and colleagues (2007) who followed Spanish-English L2 participants (n = 261) from first to the sixth grade and found that their decoding ability remained comparable to a normative L1 sample throughout the research period. One study even indicated that L2 third graders in the UK were better decoders than their L1 peers (Burgoyne, Kelly, Whiteley, & Spooner, 2009). In explaining these results, the researchers suggest that, in contrast to the L1 learners, the effort of the L2 learners was more diverted to decoding than understanding, enabling them to read faster than the L1 learners.

The results discussed above are reflected in the conclusions made by The National Literacy Panel in the US (August & Shanahan, 2006) in their review of quantitative and qualitative research on the development of literacy in language-minority learners. According to their findings the decoding skills of L2 and L1 learners generally develop at a similar pace as these skills are often “taught well enough to allow language-minority students to attain levels of performance equal to native English speakers” (August & Shanahan, 2006, p. 4). The same, however, does not apply to vocabulary and reading comprehension (e.g. Biemiller, 2003). And it is mainly their vocabulary proficiency that influences later reading comprehension achievement or its rate of growth, and less technical reading skills (see section 2.7.1).

It is important to point out that although L2 learners generally perform well in assessments of decoding, there are however certain circumstances in which their word reading skills do not seem to match quite up to that of monolingual peers. Verhoeven and van Leeuwe (2011b) explored the growth of Dutch word decoding skills of a large group of Dutch L2 participants (n = 331) throughout the first six years of schooling and compared their performance to L1 counterparts (n = 2,487) of the same age. The tests were comprised of regular CVC (consonant vowel consonant) words, complex monosyllabic words with consonant clusters, and polysyllabic words. The Dutch L2 participants were all born in the
Netherlands with parents that originated from the Mediterranean countries. The results revealed that the ability of both L2 and L1 learners to decode words of increasing length and orthographic complexity improved with age (cf. Verhoeven, 1990). However, in contrast with the studies discussed above, the performance of the L2 children lagged significantly behind that of their L1 peers in decoding all three word-categories, whereas they eventually caught up in reading the simple CVC combinations and words with consonant clusters. They, however, continued to struggle with polysyllabic words. According to the researchers, the difficulties the L2 learners had with polysyllabic words could be explained by their word deficiency (i.e. poor vocabulary and morphological awareness). If the L2 children had known the meaning of these words, or their constituent parts, they would have been less likely to run into difficulties when trying to read them. Thus, the word reading difficulties of the L2 children were perhaps more related to their poor word skills than their decoding ability. However, applying to both the L2 and the L1 participants, the researchers found that the association between word decoding and reading comprehension prevailed, indicating the need for a continued attention to the speed and automaticity of word decoding, and also vocabulary instruction, throughout the elementary school years.

These findings are in line with studies of older English L2 students in college and University, as Laufer and Nation (2001) found that the fluency of L2 English University learners in reading words of a certain frequency level was greatly affected by their vocabulary size. The same applies to their L1 peers. Similarly, Laufer and Ravenhorst-Kalovski (2010) demonstrated that a small increase of key academic vocabulary improved the reading fluency of a group of English L2 college students, and importantly, enabling them to better concentrate on the content of the text and thereby enhancing their comprehension.

From the discussion above it can be concluded that young L2 and L1 children, who are learning to read in the same language, acquire decoding skills with comparable ease or difficulty depending on the orthographic structure of the language in which they learn to read. However, as word and decoding skills are related, and because L2 children tend to have poor word skills compared to their L1 peers, they have demonstrated prevailing difficulties in reading unfamiliar words.

No study has explored the development of decoding skills among Icelandic second language learners. Thus, we have little knowledge of their
progress in this aspect of literacy development and how it affects their reading comprehension. In the light of the results discussed above, and given that first and second language learners in Iceland are subjected to similar teaching methods and amount of training, Ice2 middle school learners should have comparable decoding skills to their Ice1 same age peers. It can however be expected that those Ice2 children with short residence in the country at the start of the study may still be struggling with decoding Icelandic words. Moreover, due to less word knowledge, Ice2 children should be more likely than their Ice1 peers to be perpetually hampered by unfamiliar words in a text. However, we do not know how individual differences in reading fluency, of both Ice2 and Ice1 children, influence the rate of reading comprehension growth, if those who struggle with technical reading skills by the start of middle school develop their reading comprehension skills at a slower rate than the more fluent readers (see section 2.6.2).

2.7 The influence of vocabulary on reading comprehension

As outlined in section 2.6.1.1, during the initial stages of reading development there is a strong link between children’s ability to decode words and their reading comprehension. However, with increasing word-reading skills, this relationship quickly diminishes and vocabulary takes over as one of the strongest influencing factors on reading comprehension skills and continues to be so throughout the life span. Children, however, enter school with very different levels of word skills, which in turn may enduringly influence their reading comprehension performance.

This was demonstrated in a much-cited study by Hart and Risley (1995, 2003) that showed that the vocabulary size of three-year-old English speaking children on welfare was less than half the vocabulary size of their age peers of higher SES. Analyses of the children’s linguistic environment revealed that poor vocabulary went hand in hand with limited verbal communications at home. The welfare children had only half as much word experience per hour at home as the working-class children, and less than one-third of the word experience received by children in the professional families. The study also showed that the children’s vocabulary size and use by age of three was strongly associated with their reading comprehension scores by the age of nine and ten. Thus, it seems that early linguistic deprivation has an enduring effect on subsequent educational performance.
A number of studies have demonstrated consistent findings (e.g. Baker, et al., 1995; Chall & Jacobs, 2003; Cunningham & Stanovich, 1997; Verhoeven, et al., 2011) showing that word skills of young children predict later reading comprehension and academic achievement. An Icelandic longitudinal study (Einarsdóttir, Símonardóttir, & Björnsdóttir, 2011) consistently revealed that language assessment in preschool, as measured with TOLD-2P (a test of language development) predicted Icelandic language scores on The National Testing Battery in grades four, seven, and 10. Moreover, the researchers found that the oral comprehension subtest of TOLD-2P correlated to the highest extent with these national tests.

The predictive power of vocabulary results mainly from the fact that children who start school with small vocabularies, experience difficulties in reading comprehension and tend to lag behind children with rich vocabularies throughout the school years. Furthermore, this gap tends to grow with time. This pattern of results can be explained by the reciprocal relationship between vocabulary and reading comprehension. Children who have poor word skills have difficulties with understanding texts that include many unknown words. As a consequence, their vocabulary expands at a slower rate than the vocabulary of children who have larger vocabularies, which then affects their reading comprehension development. Children with good lexical knowledge, on the other hand, understand more words in a text and find it easier to decipher the meaning of unfamiliar words. This provides them with a better understanding of the text, thereby increasing their word skills and reading comprehension faster. Furthermore, children with solid word skills and reading comprehension tend to read more, which then leads to more learning. Thus, word skills facilitate reading comprehension and reading comprehension enhances word skills (Stanovich, 1986).

The middle school years are extremely critical in this respect. This is the time when reading texts become increasingly complicated and include more and more low-frequency academic words. It is primarily by this age that children with well established decoding skills but poor word skills start to lag behind in their reading performance (Chall & Jacobs, 2003). This can have life-long consequences, as children who experience reading difficulties by the middle school years are more likely to drop out of school later in their academic careers than those without such deficiencies (Barrington & Hendricks, 1989; Snow, et al., 2001). If children’s word skill deficiency hampers their reading comprehension by age nine or ten, its impact will
only strengthen throughout the learning process. For example, results from
the PISA assessment in 2009 demonstrated that vocabulary knowledge was
by far the strongest predictor of Danish lower secondary school learners’
(12 – 15-year-old) reading comprehension skills (Egelund, 2012, p. 87). It
was stronger than SES, reading enjoyment, and orthographic coding.

2.7.1 Second language vocabulary strongly contributes
to second language reading comprehension

Catalysed by the tendency of L2 learners to have smaller vocabularies in
each language than their L1 peers in their only language, a number of
studies have investigated how this deficiency affects their reading
comprehension skills. The findings of these studies have demonstrated that
vocabulary is the key weakness of L2 learners, which perpetually affects
their reading comprehension performance (Lervåg & Aukrust, 2010;
Verhoeven, 2000).

For example, The National Reading Panel in the US (August &
Shanahan, 2006) concluded that there is no doubt that “well-developed oral
proficiency in English is associated with English reading comprehension …
Specifically, English vocabulary knowledge …”. And that vocabulary
“should be targeted intensively throughout the entire sequence” (August &
Shanahan, 2006, pp. 4-5).

Even in the earliest stages of reading acquisition when learners are still
acquiring decoding skills, young L2 children with poor word skills may
struggle with too many unknown words. For example, Gottardo and Mueller
(2009) found that decoding, vocabulary breadth, and syntactic awareness
influenced reading comprehension among L2 first and second graders, but
only decoding skills influenced the reading comprehension among their L1
peers. This indicates that, even when texts are simple, vocabulary deficit
may hamper the reading comprehension of L2 children. Consequently, L2
readers may have to struggle with both decoding and comprehension, while
their L1 peers can concentrate on practicing decoding skills with texts they
easily understand. This is line with Erdos and colleagues’ finding (2011)
with English-French L2 first graders in Canada. Both listening
comprehension (PPVT) and decoding predicted their reading
comprehension, revealing comprehension difficulties even with basic texts.
The researchers concluded that vocabulary is uniquely important for
learning to read in a second language.
The importance of second language vocabulary in second language reading comprehension has been clearly demonstrated in a series of studies by Roessingh and colleagues (2008; 2003; 2005). They have followed groups of Canadian L2 learners, mostly of high SES, throughout the first, middle, and high school years. They found that older arriving L2 children had in fact impressive English vocabulary knowledge, which they had acquired in the country of origin, but inferior to their L1 age peers. For the younger arrivals, on the other hand, the lag could only be attributed to lack of instructional attention in the Canadian mainstream class context, where they had spent most of their schooling experience. Insufficient word skills of these two groups of L2 learners then played a crucial role in the development of their reading comprehension skills. The reading comprehension of all participants remained behind that of L1 counterparts, which then directly influenced their academic achievement. These results indicate that assessments of vocabulary can be used to gain direct insights into L2 students’ educational performance.

Research indicates that the impact of L2 vocabulary on L2 reading comprehension remains strong throughout the entire learning process. For example, Laufer and Ravenhorst-Kalovski (2010) found that the English vocabulary levels of L2 college students ($n = 745$) in Israel attending an English course for an academic purpose explained 64% of the variance in their reading comprehension scores, and that even a small increase of key academic word knowledge improved their reading comprehension significantly.

In spite of the fact that the above quoted studies demonstrated that vocabulary scores influenced later reading comprehension skills, these researchers did not explore the influence of initial word skills on the rate of reading comprehension growth, whether those who started with higher word skills increased their reading comprehensions performance at a faster rate than those with poorer initial word skills, revealing a widening gap between the groups. If word skills only influence initial reading comprehension performance but not the rate of growth the gap between those with poor and rich word skills remains the same throughout the study period. A negative influence on the rate of growth means that the gap diminishes over time, and a positive influence on the rate of growth has the devastating implication of a widening gap between the groups.

Lervåg and Aukrust (2010), on the other hand, demonstrated that vocabulary deficiency (as measured by a Norwegian translation of the
PPVT test and a word definition measurement), among Norwegian L2 second to third graders was the main reason for their slow improvement in reading comprehension compared to their L1 counterparts. The reading comprehension of the children in the latter group was significantly better at the start of the project and they also made faster progress over the 18 months of the study. Furthermore, the researchers found that vocabulary skills were a more important predictor of reading comprehension growth among the L2 than the L1 learners, as vocabulary positively influenced not only their initial reading comprehension scores, but also further growth on both of the reading comprehension measures used (the NARA and the WRMT-PC reading comprehension test). The L1 children’s vocabulary knowledge, on the other hand, only predicted reading comprehension growth on one of the tests (the NARA test). It is important to take into account the short period of the study and therefore it does not give information about the impact of initial word skills on the rate of reading comprehension growth throughout the critical middle school years.

Fortunately, and not surprisingly, efforts have been made to respond to the frequent poor performance of L2 language learners on vocabulary and reading comprehension measures compared with L1 age peers, with explicit instructional methods (Roessingh, 2004, pp. 6-8; 2008; Snow & Lawrence, 2011). Roessingh (2008) has, for example, demonstrated a strong impact of ESL (English as a second language) programs on overall academic achievement among her Canadian-English speaking L2 participants who have been fortunate enough to receive such support. The ESL support included a strong focus on vocabulary, having the main purpose of the learners reaching their academic goals.

Another example of an effective literacy project is The Senior Urban Education Research Fellowship (SUERF) program (Snow & Lawrence, 2011), which was established in the US to address the needs of public urban school learners for effective, sustainable, and research based literacy intervention. The pupils in these schools consist of a large number of L2 students of African and Hispanic origin and L1 learners with limited English proficiency and have demonstrated a constant academic achievement lag. The research team launched a project, which they named SERP (Strategic Education Research Partnership) that was targeted for middle school learners. In order to define where the main literacy problems of the students lay, the researchers interviewed teachers and principals, observed classroom lessons, and reviewed test data from participating schools.
One universally noted challenge was vocabulary – students’ ignorance of the meaning of the words they encountered in the texts. These challenging vocabulary items – words used across content areas, words characteristic of written language and academic texts, words students from non-English-speaking or low-literacy homes were unlikely to have heard from their parents – were not typically taught (Snow & Lawrence, 2011, p. 8).

The Word Generation Program was a product of this pre-investigation. This instructional literacy intervention program included strong focus on all-purpose words that are useful across all content areas. It was implemented in certain treatment schools and learners in other schools were tested for comparison purposes. The researchers found strong and lasting vocabulary improvements for the L2 learners in the treatment groups, and that these gains were maintained and even enhanced during the following school year. Moreover, vocabulary gains of the L2 treatment group significantly predicted performance on the MCAS (Massachusetts Comprehension Assessment System). Vocabulary gains of the L1 learners with limited English proficiency, however, were not significant.

L2 learners in the US of Spanish origin appear to be one of the most studied groups in the field of second language literacy acquisition (e.g. Carlo, et al., 2004; Crosson, Lesaux, & Martiniello, 2008; Lesaux, et al., 2010). A majority of these learners were born in the US and are second or third generation immigrants. Therefore, many of them are simultaneous second language learners. The vocabulary deficiency among simultaneous L2 learners, or those who start the second language acquisition at a young age, results to a large extent from the fact that their times are distributed between two or more languages (Oller, et al., 2007) (see section 2.3.2.2).

2.7.2 First language skills and second language reading comprehension - influence of age of arrival

Literature suggests that young L2 learners develop vocabulary skills in their first and the second language relative to the amount and quality of exposure they have to each language, demonstrating a time-on-task competition model (see section 2.3.2.2). Older arriving L2 learners with considerable first language word skills, on the other hand, can benefit from their first language vocabularies in their second language word learning (see section 2.3.5.1). The same appears to apply to the acquisition of second language reading comprehension. Thus, vocabulary
skills in the first language do not seem to affect reading comprehension in the second language among the youngest learners.

For example, Erdos and colleagues (2011) found that the concurrent predictive value of first language linguistic comprehension measures administered to L2 English-French first graders in Canada was specific to their performance in measures of first language reading comprehension. The same pattern of results was found for comparable second language measures. Consistently, in a longitudinal study of the reading development of Spanish-English L2 children from kindergarten through grade two, Manis and colleagues (2004) demonstrated that second language vocabulary scores obtained in kindergarten predicted later reading comprehension in that language only. This was also Proctor and colleagues’ conclusion (2012) from their research with Spanish-English L2 second to fourth graders, who were tested at the beginning and the end of one school year, where first language measures had no added contribution to second language reading comprehension. With her Ice2 preschool participants, Haraldsdóttir (2013) did not find significant relations between first language skills, as reported by parents, with their performance on an Icelandic PPVT vocabulary test (the same as used in the current study) and also not with the children’s ability to form past tenses from verbs.

Reese, Garnier, Gallimore, and Goldenberg (2000), however, demonstrated with their research positive influence of first language literacy on later reading comprehension performance in the second language. They had followed Spanish-English children in the US throughout the middle school years. These bilinguals had acquired literacy skills in an English-dominant-environment, where ultimate proficiency in English was the desired outcome, as it was considered the foundation for later academic attainment. The researchers found that both Spanish literacy and English oral proficiency at the first year of kindergarten predicted English literacy on standardized tests in grade seven. The participants, who had been read to in the Spanish language, had an initial and continued advantage in the English literacy, and this was related to parents’ educational level, and even that of the grandparents, and this positive effect was lasting. However, it was their early oral proficiency in English that predicted their later performance on English reading comprehension tests. Notably, the educational level of the parents was very low, with average seven of years of schooling, ranging from zero to sixteen. The researchers suggested, that the more educated parents continued to provide their children with richer
reading experience, which then affected the children’s reading comprehension development in the second language.

However, the above quoted studies followed young learners over a very short time and did not demonstrate first language influence on the rate of second language reading comprehension growth. Lervåg and Aukrust (2010), on the other hand, showed only a small and a marginally significant independent influence of first language vocabulary to second language reading comprehension growth with their Norwegian L2 first to third grade participants. The researchers concluded that there is little support for the idea that first language vocabulary has a causal influence on the development of second language reading comprehension skills. It would have been interesting to track these participants’ reading comprehension performance for a longer period of time to detect if the influence of first language skills continually emerges as small and marginal, if it disappears, or if it becomes stronger with time. If first language skills positively and increasingly contribute to second language reading comprehension development, the fundamental skills for academic procedure, there is a reason to suggest that second language learners should be educated in their native language within the school setting.

Jessica Bell (2011) made a review on the impact of bilingual education on second language acquisition for UNESCO (The United Nations Educational, Scientific and Cultural Organization), and concluded that if children are growing up only in the first language, educational provisions should support them in becoming highly proficient in that language before engaging in academic work in the second language. She refers to Cummins (1984) suggesting that it takes six to eight years to become highly proficient, i.e. to reach CALP, in the second language (see section 2.1.1.1). On the other hand, Bell does not mention the number of evidence showing that L2 learners too frequently do not achieve this level, with serious consequences for the academic procedure (see section 2.7.1). In fact, the influence of first language instruction on second language acquisition is rather complex, and many influencing factors have to be taken into consideration, for example the influence of highly effective second language programs, including focus on academic vocabulary (Lesaux, Kieffer, Kelley, & Harris, 2015; Roessingh & Douglas, 2013; Snow & Lawrence, 2011).

It is indeed academic vocabulary instruction in the second language that is fundamental for reading comprehension in that language. For example,
Reese, Gallimore, and Guthier (2005) tracked English reading skills of Spanish-English bilinguals throughout their elementary and middle school years. These bilinguals had initially been instructed in their first language, in transitional bilingual programs. When the students made the transition to an English instruction, which took place for most of the learners between grades two and four, all the children experienced a dramatic decline in performance on standardized reading comprehension tests in English. Thus, the children moved to the English program by the age when educational demands increase and learners’ knowledge of academic words, in the second language, becomes pivotal (see section 2.1.1.1). The drop in reading achievement was not temporary and was not overcome by the majority of the students. One possible reason, the researchers suggested, was that the children had not received sufficient English language instruction as a part of their bilingual programs. It is indeed not only the amount and the quality of the language input that counts (see section 2.3.1.1), but also the quality language instruction (see section 2.3.1.2). Therefore, when bilingual programs are investigated the quality of instruction has to be taken into account, at least in the second language when that is under study.

The United Neighbourhood Organization (UNO) in the US chose English immersion over bilingual education … as it is effective in closing the performance gap between ELL’s and their peers nationwide, and is financially viable and scalable unlike the many bilingual transition programs that require untenable complements of teachers and resources and produce mixed results at best (Lesaux, 2013, p. 46).

On the other hand, first language impact on second language acquisition is more evident among older second language learners who have acquired considerable first language skills when they start to learn a second language. For example, Roessingh and colleagues (2008; 2002; 2003; 2005) have demonstrated in their studies that older arriving L2 learners benefit from having more first language word skills to build on than younger arrivals (see section 2.3.5.1), and they also have more background knowledge, which they have acquired in their first language. For them ‘less is more’ in the second language reading comprehension performance as they appear to be able to compensate for their vocabulary deficiency in their second language reading comprehension with more common underlying proficiency (termed CUP) (Cummins, 1981), more background knowledge and more first language word skills (Roessingh, 2008, pp. 100-101). Thus poor vocabulary
skills in the second language hinder their reading comprehension to a lesser extent than the younger arrivals. However, the greatest challenge for the older arriving L2 students is the fact that while they spend their times acquiring fundamental skills in the second language they may be left behind in the various subject materials, which in turn can hinder their educational progress (Roessingh & Kover, 2003).

Collier (1987/1988) studied the time it took 1,548 L2 learners in the US to acquire academic skills in all subjects and concluded that learners who were eight to twelve years old when they arrived seemed to be the most advantaged. This is because they already have some first language vocabulary skills to transfer to the second language and they still have time to catch up academically with their L1 peers. Even though those who arrive later have more developed first language vocabulary skills and background knowledge, they have less time. The youngest arrivals, on the other hand with little first and second language word skills, tend to develop reading comprehension skills at a slower rate than their monolingual peers. They lack the first language vocabulary, and notably also in second language vocabulary, necessary to facilitate the learning of academic, low-frequency words in the second language, which in turn hampers their reading comprehension with prevailing effects on their academic procedure.

In the only longitudinal study that has been conducted on the development of vocabulary among Ice2 learners (Thordardottir & Juliusdottir, 2012) the effect of age of arrival was not so clear-cut. According to the results, the older arriving participants (n = 8, age of arrival ranging from 11.5 years to 14) demonstrated less increase between grades in general language skills (TOLD) and in vocabulary (translated PPVT) than the younger arrivals (n = 8, age of arrival ranging from six years to nine), when matched on length of residence in Iceland. However, no measures of reading comprehension skills (or any other measures of academic achievement) were included in this study. Thus we have no information on how these two groups of Ice2 learners were faring academically. Furthermore, neither Roessingh nor Thordardottir and Juliusdottir explored the extent to which age of arrival influenced the rate of vocabulary and reading comprehension growth, as they only compared the achievement of different age of arrival groups between grades.
2.8  The development of reading comprehension among L2 and L1 learners - The reading comprehension gap between L1 and L2 learners

As discussed above, the vocabulary gap between L1 and L2 learners tends to remain and even grow throughout the school years (see section 2.3.3). This same pattern of results has also emerged in a number of other studies on reading comprehension measures, some revealing a diminishing gap, whereas most studies reveal a constant and a widening gap.

Lervåg and Aukrust’s longitudinal research (2010), that was discussed above (see section 2.3.3.1), revealed that Norwegian L2 children \((n = 90)\) had smaller L2 vocabularies in grade two, and they also made slower gains in both vocabulary and reading comprehension than a comparison group of L1 children \((n = 198)\) during a period of 18 months. This is in line with Lesaux and colleagues’ (2010) finding that Spanish-English L2 learners \((n = 87)\) in the US developed less oral language skills (including vocabulary) from grade four to five than their L1 peers, both in their first and their second language. The L2 learners also had poorer reading comprehension skills with slower growth during this period of two years.

Droop and Verhoeven (2003) detected an ever-widening gap between the vocabulary scores of Dutch L2 \((n = 122)\) and L1 \((n = 143)\) third and fourth graders, over two school years. The L2 children continued to lag behind their L1 peers in vocabulary and reading comprehension, the latter gap though remaining the same.

The participants in the above quoted studies were only followed over a short period of time. These studies do therefore not demonstrate the reading comprehension development throughout the critical middle school years, when children have started the longstanding process of reading to learn.

Nakamoto and colleagues (2007), on the other hand, followed Spanish-English L2 and L1 learners from grade one through grade six. The reading comprehension gap between the two groups started to increase by grade three, revealing a widening gap between the groups. Kieffer (2008) investigated data from students that had been followed from kindergarten through fifth grade. He found that while both groups slowed down in their reading comprehension development over the time of the study, the gap widened as the L2 children demonstrated even a slower growth than the L1 peers.
Mancilla-Martinez and colleagues (2011) participated in the above quoted SERP project in the US (see section 2.7.1) aimed at boosting academic performance of L2 learners (Snow & Lawrence, 2011). As a part of the pre-investigation they investigated reading comprehension growth of L2 learners \((n = 43 - 55)\) from grade five to grade seven, in an urban public school which serves 91% Spanish-English and 91% low-income population. They discovered that neither listening comprehension nor word reading ability assessed in the fifth grade predicted learners’ growth in reading comprehension up to grade seven. Those who lagged behind initially did not manage to accelerate their reading comprehension growth, not even by the seventh grade when their L1 age peers slowed down in their development. Instead, they continued to follow their path with the distance between the two groups remaining the same. The researchers conclude that this pattern of growth is not worrisome for those L2 learners who performed within the average range, as they have appropriate skills for grade-level expectations. However, for the learners with low comprehension in the fifth grade, the slowing rate of growth during the middle school grades is a cause for concern.

Our results showed that the rates of growth in reading comprehension did not vary across student ability level, such that students with initially low performance remained on a depressed and slowing trajectory through seventh grade. Thus these students suffer from consistently insufficient reading comprehension skills during this critical developmental period in which expectations for reading to learn content are high and increasing. Additionally, students’ ability to access grade-level material is limited by a slowing growth curve and a developmental plateau in skill development, such that their skills are stymied at inadequately low levels, preventing them from catching up with their peers (Mancilla-Martinez, Kieffer, Biancarosa, Gina, et al., 2011, p. 350).

As mentioned before, the Word Generation Program (2011) was launched in the US with significant positive gains, both on vocabulary growth and reading comprehension performance.

A six-year longitudinal research conducted by Verhoeven and van Leeuwe (2011b) demonstrated that Dutch L2 learners \((n = 394)\) in grade one and up to grade six initially lagged behind their same aged L1 peers \((n = 1.293)\) both in listening- and reading comprehension, however, the gap between these groups diminished during the course of the study, as the L2
learners made a faster growth than the L1 learners. This may have been the positive result of a strong focus on vocabulary as a contributing factor to reading comprehension among Dutch researchers for several years (cf. Appel, Kuiken, & Vermeer, 1995; Appel & Vermeer, 1998; Kuiken & Vermeer, 2005; Verhoeven, 1990; 2000). As has been discussed above, explicit vocabulary instruction was implemented as a specific goal in the Curriculum for Primary School Education in the Netherlands in 2006 (Ministerie van Onderwijs, 2006). Verhoeven and van Leeuwe found, however, that the gap between the two groups did not disappear during the research period. This indicates the need for more and continued attention to word learning. “Continued attention to the speed and automaticity of word decoding and lexical access throughout the elementary school years can thus be recommended for both L1 and L2 learners” (Verhoeven & van Leeuwe, 2011b).

From this it can be suggested that without instructional methods aimed at boosting word skills of L2 children, this group of learners tends to suffer from a continuous or rather a widening reading comprehension lag compared to L1 peers. However, in the Netherlands, Canada, and in the US specific focus on L2 vocabulary acquisition and instructional programs, with L2 word learning playing a vital role, have demonstrated significant positive effects on L2 reading comprehension. The remaining reading comprehension gap between L1 and L2 students that the above quoted studies have demonstrated merits further investigation, in which the development is followed over an extended period of time, throughout the middle school years and, the best, beyond.

In Iceland, vocabulary has gained trivial attention in the curriculum both for the subject Icelandic language and for Icelandic as a second language. This limited emphasis on vocabulary is reflected in the teaching material available for Ice2 and Ice1 children. The findings of the longitudinal research of Thordardottir and Juliusdottir (2012) do not give cause for optimism. The majority of the Ice2 children who participated scored far below the normal range for Ice1 children in general language skills (TOLD) and vocabulary (translated PPVT) throughout the research period. The effect of the children’s vocabulary scores on their reading comprehension skills was, however, not investigated. Thus, little is known about the influence of Icelandic vocabulary on the rate of reading comprehension growth, of both Ice2 and Ice1 learners, as this has not yet been investigated with specific Icelandic vocabulary measurements. This fact gives cause for
concern as the percentage of learners with poor reading comprehension in the 2009 PISA assessment was the highest in Iceland (17%) among Nordic countries (Egelund, 2012), and the percentage of Ice2 learners among the weak readers (11.6%) was almost three times higher than the percentage of Ice2 learners of the Icelandic population as a whole (4.5%) (Egelund, 2012, p. 29).

2.9 Measuring reading comprehension

When assessing reading comprehension it is important that the measurements measure what they are intended to measure and provide accurate information about the participants’ skills. Thus, the measurements must be both valid and reliable (see section 2.2.1). Content validity is met if all comprehension questions tap the ability to read and understand what is being read. The questions should not tap decoding ability only, or mainly comprehension, but should challenge both components. Moreover, test items should tap the readers’ understanding of the text and not their general knowledge, which they do not need to look up in the text (Hoover & Gough, 1990, p. 131).

Construct validity is likely to be obtained if reading comprehension measurements tap both decoding and comprehension skills. This depends, among other things, on the complexity and the familiarity of the content and on the prior knowledge of the readers. If the reader is unfamiliar with the content of the text, he may find it difficult to extract its message, even though he can read and understand all words and sentences (Roessingh, 2008). If background knowledge hinders the reader in extracting the message from the text, the test is not testing reading comprehension in a valid way.

There are various ways in which reading comprehension can be measured. In multiple-choice reading comprehension tests participants have to select which of three or more options is the correct answer. Although participants do not need to answer the questions with their own words, they need to compare more than two response options that sometimes can be misleading. Multiple-choice questions have the disadvantage that answers can to some extent be based on guessing given that there is a 20-30% chance of selecting the correct answer depending on the number of presented choices.

89
Open-ended questions, on the other hand, require written or oral responses. As the participant has to express his answers, more demands on productive language skills may underestimate the comprehension skills of many L2 learners, representing lower construct validity. Studies suggest that children who score poorly on open-ended comprehension questions perform significantly better on reading comprehension tests with closed questions (Burgoyne, et al., 2009; Spooner, Baddeley, & Gathercole, 2004).

Both these test formats offer the possibility to challenge different reading strategies. The least demanding comprehension questions are entirely formed from what is stated in the text, demanding the test-taker to recall or look for the answer in the text, challenging literal comprehension. An example of a reading comprehension question of this sort can be found in the NARA test, used in the present research. For example, the test-taker is asked what the boy and the girl in one of the stories in the test received from their uncle, both of which are explicitly stated in the text: the boy got a train and the girl skates. More difficult questions challenge inferential comprehension, requiring the reader to read between the lines, or interpret the writer’s meaning from information that is implicit in the text. Comprehension questions that require inference skills of the reader are included in the reading comprehension tests used in the present research. For example, in one story in Lesskilningspróf Námsmatssstofnunar (L.N., an Icelandic test used in this study) (Gunnarsson & Skúlason, 2009) it is stated that the name of the main person is that of a flower. The name is, however, not given. The participant is then asked to select a name for this person out of three possible choices where only one is of a flower, which additionally requires the reader to relate to his prior knowledge. If the reader does not know this flower (Ice. lilja) he has to guess the answer. This particular question could actually be quite difficult for second language learners to answer, as they are less likely to know Icelandic names of flowers than their monolingual Ice1 peers. The most difficult comprehension questions demand an evaluative comprehension, and are also included in the L.N. reading comprehension test used in this study. In such cases, the answer is not stated directly in the text, but the reader can draw the correct conclusion by reading between the lines, collect pieces of information from different places in the text and connect them to his or her background knowledge.
2.10 Summary and conclusions

In this chapter I have reviewed literature about the development of vocabulary of second language learners and how it is related to their reading achievement. At least two broad conclusions, regarding vocabulary and reading comprehension skills, can be drawn from the research evidence, each giving rise to new lines of research.

There is now substantial evidence indicating that bilingual learners develop less vocabulary in each language than do monolinguals in their only language, resulting in a remaining or even a widening gap between the two groups. However, research has also shown that instructional vocabulary programs have significant positive effects on second language word acquisition and academic prospects of L2 learners. Little attention has been paid to word learning in the curriculum and in Icelandic schoolbooks intended for Icle2 and Icle1 children. The findings of Thordardóttir and Juliusdóttir (2012) and from my MA study (S. Ólafsdóttir & Ragnar Ásgeirsdóttir, 2010) give cause for pessimism, as they suggest little or no vocabulary improvement of Icle2 compulsory school learners each school year. However, only one of these studies was longitudinal and did not include specific Icelandic vocabulary measurements. The other study was cross-sectional and employed only one Icelandic vocabulary test. Furthermore, neither of the studies included a comparison group of Icle1 age peers, and no account was made of possible influences of important background variables. Consequently, we have little overview of how different aspects of the vocabulary knowledge of Icelandic second language learners develop. Moreover, vocabulary studies with second language learners of other languages have included very young children, or not investigated developmental growth rate. Furthermore, the great majority of vocabulary research has not included measures of learners’ proficiency in utilising important academic lexical items, which are indeed the prime word skills, fundamental for educational success.

Research reviewed in this chapter suggests that there may also be different developmental pattern in the acquisition of literacy between first and second language learners. Studies with young readers of English, Dutch, and Norwegian languages (August & Shanahan, 2006; Droop & Verhoeven, 2003; Lervåg & Aukrust, 2010) indicate that L2 children generally develop decoding skills commensurate to their L1 peers (August & Shanahan, 2006, p. 4). However, following Dutch learners from grade one and up to grade six, Verhoeven and colleagues (2011a) found a
prevailing difficulty among the L2 learners in reading polysyllabic words, which they linked to their poor word skills. The researchers found that the relation between decoding and reading comprehension skills prevailed, and notably for both the L2 and the L1 children, meaning that some L2 and L1 children were continually struggling with reading fluency which perpetually hampered their reading comprehension. Icelandic is a highly inflected language with many multi syllable words. These kinds of words are difficult for Icelandic beginning readers to decode (Birgisdóttir, 2010), however, by second grade word length no longer seems to hamper decoding accuracy of Ice1 children. No research has explored the development of decoding skills of Icelandic speaking second language learners and in no study has the impact of reading fluency on the rate of reading comprehension growth for Ice2 and Ice1 learners been investigated.

Most literacy studies throughout the years with second language learners of other languages have been conducted with very young children. It is, however, particularly by grade four, when reading texts become increasingly complicated and include an increasing number of low-frequency academic words, that L2 learners start to lag behind their L1 peers in reading comprehension (Roessingh, 2008). Researchers have therefore the last decade increasingly followed children further, or up to grade five to seven. They have demonstrated a widening (Nakamoto, 2007; Kieffer, 2008; Roessingh, 2008; Droop & Verhoeven, 2003), or a constant reading comprehension gap between the groups (Mancilla-Martinez & colleagues, 2011). The reading comprehension gap that has been detected between L1 and L2 learners in the first few years of schooling and up to grade seven, merits further investigation. L2 and L1 children need to be followed over an extended period of time throughout the middle school years, and beyond. When reading comprehension acquisition is explored as a process that develops over time it is highly important that the rate of growth is treated as a dependent variable. Such information makes it possible to detect the extent to which initial decoding and vocabulary skills influence both initial reading comprehension and the rate of growth.

In fact, as is the case with second language vocabulary learning, second language reading comprehension is influenced by various skills and contextual factors, which means that generalization of results across cultures could be misleading. No research has investigated the development of reading comprehension among Icelandic second language learners and how it is affected by their Icelandic word skills. We have no knowledge what so
ever about normative shape and the rate of development in Icelandic vocabulary and reading comprehension skills among Ice2 and Ice1 compulsory school learners. We also do not know what linguistic and contextual factors drive change in these skills.

2.11 Aims and research questions

The central aim of this research is to explore the development of vocabulary and reading comprehension among Ice2 middle school learners, from grade four through grade eight, and the predictive power of vocabulary for their reading comprehension growth. Moreover, we seek to examine the extent to which theoretically important linguistic and contextual factors may influence the emergence of these skills. Finally, we aim to study academic productive language proficiency of this population and how this is predicted by earlier word skills.

This study is the first of its kind in Iceland and can contribute to our knowledge of Icelandic vocabulary and reading comprehension development during the middle school years. This is a period in which educational demands typically increase, as more complicated texts have to be read independently, providing a source for new knowledge. Such understanding will in turn be important for planning interventions to help prevent the development of reading comprehension problems of Ice2 learners. Inadequate reading comprehension can hinder students from achieving academic requirements and may lead to school dropout.

The study will seek to find answers to the following research questions:

1. Is there an initial and developmental difference in vocabulary and reading comprehension skills between Ice2 and Ice1 middle school learners?

2. To what extent do reading fluency and Icelandic word skills of Ice2 and Ice1 learners predict their reading comprehension skills and its rate of growth?

3. To what extent are Ice2 learners’ vocabulary and reading comprehension level and rate of growth influenced by their age of arrival, first language proficiency, literacy, first language relatedness with the Icelandic language, and maternal education?

4. Do Ice1 middle school learners use a higher number of higher order (tier 2) words than Ice2 peers, and how is this related to
the children’s productive language proficiency? To what extent are these predicted by earlier vocabulary skills?

2.11.1 Hypotheses

It is hypothesized that the Ice2 learners will fall farther behind Ice1 age peers on vocabulary measures throughout the study period, revealing a widening gap between the two groups. This proposition is based on previous findings with L2 learners of English (Roessingh, 2008), Dutch (Droop & Verhoeven, 2003), and Norwegian (Lervåg & Aukrust, 2010), and on the three prior studies with Ice2 children (Haraldsdóttir, 2013; Thordardottir & Juliusdottir, 2012; S. Ólafsdóttir & Ragnarsdóttir, 2010), the last two suggesting that Ice2 children are not increasing their Icelandic vocabularies between all grades. It is also important to note that the Icelandic language is more complex grammatically than English, Dutch, and Norwegian, and therefore it can take Ice2 children even longer than L2 of the other languages to increase their Icelandic word skills. Importantly, for example in the US, Canada, and the Netherlands positive outcomes have emerged from educational programs aimed at increasing L2 learners’ lexicon (Appel & Vermeer, 1997; Roessingh, 2008; Snow & Lawrence, 2011). However, explicit vocabulary instruction has not been practiced, in general, in Icelandic schools. Moreover, in the Icelandic curriculum vocabulary plays only a minor role (Ádalnámskrá grunnskóla: almennur hluti 2011 og greinasvið 2013), and its function as pivotal skills to reading comprehension is overlooked.

Based on The Simple View of Reading (Hoover & Gough, 1990), it is hypothesized that reading fluency will influence initial reading comprehension skills of both groups, the Ice2 and the Ice2 learners, whereas its impact will diminish throughout the study period, meaning a negative influence on the rate of growth, when Icelandic word skills take over (Catts, et al., 2005). Furthermore, it is suggested that initial Icelandic vocabulary measures will positively influence initial scores and the rate of growth in vocabulary and reading comprehension for both the Ice2 and the Ice1 group, that those with the largest word skills in the first year of middle school will make a faster improvement in these skills than those with poorer initial word skills (cf. Lervåg & Aukrust, 2010; Lesaux, et al., 2010; Verhoeven, 1990).
In addition, it is expected that age of arrival will positively influence the rate of growth in Icelandic vocabulary and reading comprehension (cf. Roessingh, 2008) (see sections 2.3.5.1 and 2.7.2). Taking into account the complexity of the relationship between first and second language skills (e.g. Lervåg & Aukrust, 2010; Proctor, et al., 2012), and the influence of first language instruction on second language acquisition (Proctor, August, Snow, & Barr, 2013) (see section 2.7.2), it is not expected that first language proficiency and instruction, as reported by parents, will influence Icelandic vocabulary and reading comprehension skills of our Ice2 participants (see section 2.7.2).

The European languages belong to the Indo-European language family. These languages share many cognates and also have some similarities in their grammatical structure. My MA study (S. Ólafsdóttir & Ragnarsdóttir, 2010) demonstrated a difference on the Icelandic PPVT vocabulary test between European and non-European primary school Ice2 learners, however only among those with the longest residence in the country, the European group outperforming the non-European group (see section 2.3.4). Nonetheless, hypothesizing if a difference will emerge during the middle school years and how it develops over time, we have very little evidence to build on.

In a number of studies, SES, frequently indexed by a single indicator as maternal education, has emerged as a strong influencing factor on children language development, mediated by quality language input (see section 2.3.1.1). The Ice2 participants all have both parents and/or foster parents with other first languages than Icelandic (see section 3.4), and are therefore likely to use their first languages predominantly at home (see sections 2.3.1.1 and 2.7.2). It is therefore not proposed that maternal education will influence their Icelandic word learning and reading comprehension, also taking account the findings of Haraldsdóttir (2013). On the other hand, based on earlier findings (Hart & Risley, 2003; Ragnarsdóttir, et al., 2009) it is expected that this background variable will influence Icelandic word skills and reading comprehension for our Ice1 learners, in favor of those with higher educated mothers.

Suggesting an answer to the last research question we suggest that Ice1 children will use a higher number of tier 2 words in their writings and that their level of writing proficiency will also be higher (see section 2.3.3.1). We also hypothesize that the number of tier 2 words will increase in line with higher writing proficiency level (cf. Roessingh, Elgie, & Kover, 2013).
Furthermore we propose that writing proficiency and the use of tier 2 words will be positively predicted by earlier receptive vocabulary skills (PPVT, Orðskilningur, and Orðalykill). This hypothesis is based on the fact that several encounters with words are necessary before the words are likely to come to mind during speech and writing. Thus, those with larger initial word skills are likely to develop not only larger receptive vocabulary, but also productive word skills, and more proficient writing skills.

This study has several strengths in testing the hypotheses. Firstly, it is longitudinal and tracks vocabulary and reading comprehension development over five years, from grade four and all the way up to grade eight. Secondly, the vocabulary measurements, PPVT, Orðskilningur, and Orðalykill, and the L.N. reading comprehension test are all norm-referenced measurements that have gone through a strict validation process. Thirdly, word skills are tested with different test formats, tapping different aspects of word learning which increases the reliability and the ecological validity of the study. Moreover, the inclusion of a free academic vocabulary measurement is highly important, as this represents the prime word knowledge of the lexical items that are fundamental for educational achievement. Finally, the majority of participants took part at all three time points, with attrition spanning from only 2% - 12% (see section 3.4).

In this study it is not practically possible to take into account various other factors that may be related with second language vocabulary and reading comprehension skills, like motivation, self-esteem, and social environment. This means that some variables that may be related with vocabulary and reading comprehension skills and development of Ice2 middle school learners are excluded from the analysis. Findings thus do not give a complete explanation or a complete model that explains fully the relation between outcome variables and what influences their emergence and rate of growth. The choice of variables for the current study is based on research findings in the field, in which these specific constructs have emerged as significant influencing factors on the acquisition of both the dependent variables, vocabulary and reading comprehension.

Statistical models serve as useful tools for scientists who are grounded in the philosophy of empiricism and analytic reductionism. Empiricism often favors observations derived from experimentation over passive observations, and rejects evidence based on intuition or reasoning alone. Analytic reductionism is the belief that we can gain understanding of a complex system by
understanding the parts of the system and how they interact (Schatzneider & Petscher, 2011, pp. 55-56).
3 Method

3.1 Research design and data collection

This research employed a longitudinal design in which two age groups of Ice2 participants were tested with equal intervals at three time points. The younger group was tested in grades four, five, and six, and the older group in grades six, seven, and eight. A comparison group of Ice1 age peers was also assessed at all three time points. Measures included tests of vocabulary, reading comprehension, reading fluency, and writing skills. The vocabulary and reading comprehension measures were administered at all three time points, but reading fluency and writing skills only once: reading fluency at Testing 1, to the fourth and the sixth graders and writing skills at Testing 3, to the sixth and eighth graders (see Table 2 for an overview of administration of measures). Background information about age of arrival in Iceland, parents’ estimation of first language proficiency and literacy knowledge, first language instruction, and mother’s educational status were also obtained (see Appendices 1, 2, 3, and 4).

Testing 1 was administered in 2012, Testing 2 in 2013, and Testing 3 in 2014 (see Table 2).
Table 2. Measurements administered at Testing 1, 2, and 3: to the younger group (grades four, five, and six) and to the older group (grades six, seven, and eight)

<table>
<thead>
<tr>
<th></th>
<th>PPVT</th>
<th>Orabjörn</th>
<th>Orskillingur</th>
<th>NARA</th>
<th>Lesskilningspóf</th>
<th>Námsmat L.N.</th>
<th>Reading fluency</th>
<th>Non-word repetition</th>
<th>Writing test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Testing 1:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ice2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 4 and 6</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ice1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 4 and 6</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Testing 2:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ice2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 5 and 7</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Ice1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 5 and 7</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Testing 3:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ice2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 6 and 8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ice1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 6 and 8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>X x</td>
</tr>
</tbody>
</table>

3.2 Preliminary inquiries

As some of the measurements are standardized or norm-referenced only or predominantly with Ice1 learners, four Ice2 learners in grade four and grade six were randomly selected for a trial testing. These preliminary administrations were intended to give an indication about whether the tests were suitable for this group of learners. Results demonstrated no floor or ceiling effects. These preliminary participants did not take part in the actual research.
3.3 Ethical issues

Permission was initially obtained from The Data Protection Authority (Icel. Persónuvernd) and from the local authorities of the schools’ districts. After having received letters of approval from these authorities the schools’ principals were contacted by phone. After a brief discussion where the intended research project was described and permission obtained for further process within the schools, permission letters were sent to the principals. When these had been signed and collected, contact persons within the schools were selected and contacted. Permission letters were then sent to interpreters for translation into the various first languages of the parents. When these had been received they were delivered to the children. From same schools, three schools were selected for the participation of Ice1 age peers, after approval had been obtained from the principals. When permission letters from Ice2 and Ice1 parents had been returned, signed and approved, the testing could be launched.

3.4 Participants

Participants of the study included 90 Ice2 learners, 49 were in grade four and 41 in grade six. A comparison group of 96 Ice1 learners also participated in the study; 41 in grade four and 55 in grade six (see Table 3).

The Ice2 learners were recruited from a total 27 schools, situated in areas in the southwest, northwest and northern parts of the country, each representing a wide range of social and economic status. Eight of the schools were selected within Iceland’s capital Reykjavik on the basis of their average scores (low, medium, high) in 2011 on the subtest of Icelandic in the national tests carried out by The National Testing Institute (Icel. Námsmatssístofnun) in grades four, seven, and ten, every year. Three schools are situated in Akureyri, which is Iceland’s biggest town and is located in the north, and two schools are located close to the main road between Akureyri and Reykjavik (which is a part of Iceland’s ring road). All children who were attending grades four and six in these schools and with both parents and/or foster parents whose first languages are other than Icelandic, received a permission letter for their parents. These letters were, when needed, in the parents’ first languages. Based on findings showing that L2 children tend to be over-represented as suffering from specific learning difficulties (Limbos & Geva, 2001; Samson & Lesaux, 2009), no Ice2 children were excluded on the basis of such diagnosis. Parents of 92 Ice2
children gave approval for the participation. From these 92 children one were excluded from the sample, as he was illiterate at both Testing 1 and Testing 2. Another participant had multiple errors on the non-word repetition test administered in Testing 2, indicating language impairment, (see section 3.5.5). He had scores on other measurements under and close to the minimum scores at Testing 1 and Testing 2. However, at Testing 3 he emerged as a statistical outlier (Stud. Residuals > 3) and was excluded from the study. Thus 90 Ice2 learners were included in the analysis at Testing 1 ($n = 49$ in the younger group and $n = 41$ in the older group), of whom 61 participated in my MA study ($n = 35$ in the younger group; $n = 26$ in the older group). Thus 49 participants in the younger group participated at Testing 1, all 49 participated in Testing 2, and 48 in Testing 3, the attrition only 2%. In total 41 participants in the older group participated in Testing 1, 39 participated in Testing 2, and 34 in Testing 3, giving 12% attrition.

The Ice1 comparison group was recruited from three schools in the southwest part of Iceland. These three schools had a low, medium, and high average score on the 2011 national test of Icelandic, and included some of the Ice2 participants. All learners attending grades four and six in these schools, whose parents and/or foster parents have Icelandic as their first language, received a permission letter for their parents. Only Ice1 learners with diagnosis of severe learning difficulties were excluded. In total 96 Ice1 participants participated in Testing 1 ($n = 41$ in the younger group and $n = 55$ in the older group). No statistical outliers were detected within the group ($Cook’s < 1$, Stud. Residuals < 3) and thus no one was excluded from the analysis. Nonetheless, three Ice1 learners scored under 95% on the non-word repetition test. These participants scored close to mean scores at all three time points on all measurements, and one was close to maximum scores. From the 41 participants in the younger group who participated in Testing 1, 37 participated in Testing 2, and 37 in Testing 3, with 10% attrition. In total 55 participants in the older group participated in Testing 1, 48 participated in Testing 2, and 48 in Testing 3, giving attrition rate of 12%.

As demonstrated in Figure 1, most of the participants have Polish as their first language, or 39 in total, and 13 participants have Philippine languages, which are here classified together. Three participants have two first languages.
In total 65 Ice2 children had European first languages and 25 non-European first languages (cf. research question 3, section 4.3).

Table 3 demonstrates means and standard deviation of background information for both age groups of the Ice2 and Ice1 participants.
Table 3. Ice2 and Ice1 participants’ background variables (means and standard deviations)

<table>
<thead>
<tr>
<th></th>
<th>Ice2 Grade 4</th>
<th>Ice2 Grade 6</th>
<th>Ice2 Combined</th>
<th>Ice1 Grade 4</th>
<th>Ice1 Grade 6</th>
<th>Ice1 Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 49</td>
<td>N = 41</td>
<td>N = 90</td>
<td>N = 41</td>
<td>N = 55</td>
<td>N = 96</td>
</tr>
<tr>
<td>Age of arrival</td>
<td>2.75 (2.85)</td>
<td>4.90 (3.61)</td>
<td>3.73 (3.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First language</td>
<td>3.77 (1.06)</td>
<td>3.85 (1.08)</td>
<td>3.81 (1.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First language</td>
<td>3.35 (1.67)</td>
<td>3.49 (1.51)</td>
<td>3.41 (1.60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>literacy</td>
<td>0.99 (1.62)</td>
<td>1.15 (2.24)</td>
<td>1.06 (1.91)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>instruction</td>
<td>2.86 (1.02)</td>
<td>2.85 (0.81)</td>
<td>2.85 (0.92)</td>
<td>3.68 (0.86)</td>
<td>3.36 (1.1)</td>
<td>3.49 (1.0)</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.86 (1.02)</td>
<td>2.85 (0.81)</td>
<td>2.85 (0.92)</td>
<td>3.68 (0.86)</td>
<td>3.36 (1.1)</td>
<td>3.49 (1.0)</td>
</tr>
</tbody>
</table>

As Table 3 reveals, the Ice2 participants arrived to Iceland at various ages. Half of the participants, or 51% (n = 46), arrived before the age of five, of whom 38% were born in Iceland (n = 34), 28% (n = 25) arrived at age five and six, and 21% (n = 19) arrived by age seven or later. In the analysis, age of arrival is used as a scale variable (see section 4.3). The younger group of Ice2 learners was significantly younger than the older group when they arrived in Iceland (t(75) = -3.085; p = .00), whereas the two age groups did not differ significantly on other background variables.

Parents’ estimation of first language proficiency and literacy were reported on Likert scale as in comparison with age peers living in the host country. Most of the Ice2 children, or 69% (n = 63), were estimated as well or very well proficient in their first language, whereas 8% with rather poor or poor first language skills. Moreover, 61% (n = 54) were estimated as well, or very well literate in their first language, whereas 13% (n = 12) were not able to read in their first language. First language instruction was

1 Significant mean difference between Ice2 age groups (p < 0.05)
2 First language proficiency estimated by parents, measured on a 1–5 scale, where 1 = poor, 2 = rather poor, 3 = neither good nor poor, 4 = good, 5 = very good
3 First language literacy estimated by parents, measured on a 1–5 scale, where 1 = poor, 2 = rather poor, 3 = neither good nor poor, 4 = good, 5 = very good
4 First language instruction was measured as hours per week
5 Maternal education was measured on a 1-5 scale, where 1 = not compulsory school, 2 = compulsory school, 3 = college, 4 = university, 5 = higher university degree

106
counted as the number of hours per week on average, as reported by the parents. In total 64% of the Ice2 participants \((n = 56)\) did not get any first language instruction, whereas 22% \((n = 19)\) received one to three hours per week, and 8% \((n = 7)\) five or more hours per week.

Maternal educational level differed significantly between the Ice2 and Ice1 learners \((t(179) = 4.464; p = .00)\). While 53% \((n = 46)\) of the Ice1 mothers had completed university degree, the same applied to only 17% \((n = 15)\) of the Ice2 mothers. Similarly, the proportion of mothers who had completed compulsory education (or less) was 31% among the Ice2 participants, but 17% of the Ice1 mothers. Furthermore, while four Ice2 mothers, or 4.4%, had not completed compulsory school education the share was 2.2% of the Ice1 mothers. Just as for the Ice2 children, the two age groups of Ice1 learners did not differ on maternal education.

### 3.5 Measurements

For measuring Icelandic word skills, three specific Icelandic vocabulary assessments were used: PPVT, Orðalykill, and Orðskilningur (see detailed descriptions below). These vocabulary tests challenge different aspects of word skills and include different test formats.

#### 3.5.1 Receptive word skills

Two vocabulary measurements tapped receptive word skills, PPVT and Orðskilningur.

##### 3.5.1.1 PPVT

The breadth of receptive vocabulary was measured with an Icelandic translation and adaption of the American vocabulary test Peabody Picture Vocabulary Test, PPVT-4 (Dunn & Dunn, 2007; V. Ólafsdóttir, 2011) (see selection of words in section 2.2.3.1). It was pre-tested with Ice1 children from age three to ten years and norm referenced for four to nine year old children. As the participants of the present study were at and beyond this age this measurement was only administered to the Ice2 children.

The test was administered individually. The child received a booklet with four coloured pictures presenting four options for each test item. A word was read out loud by the administrator and the child pointed at the picture that best represented the target word. The booklet consists of 165 test items, divided into 14 blocks, preceded by three training words. The first item, or the starting point, is determined by the chronological age of the
child. However, as the test was pre-tested and norm referenced with Ice1 children, the starting point was moved backwards, suggested by the child’s length of residence in Iceland. If the child made two errors in the first block, the starting point was moved backwards by one block. When the child had made more than six consecutive errors in one block, administration was ended.

As words are presented in their single forms and meanings requiring no productive output from the test-taker, this vocabulary test only taps the breadth of receptive word skills, representing construct validity (Read, 2000, pp. 94-99; Schmitt, 2010, pp. 181-182).

The inter-item reliability of this test is reported Cronbach’s alpha .97 which is consistent with that obtained in the present study (Cronbach’s alpha = .97 -.98). Validity was confirmed with correlation analysis (see section 2.2.1). PPVT correlated with Orðalykill in each grade (r = .61 -.76), however, only with Orðskilningur in grades six, seven, and eight (r = .33 -.76) (with unsatisfactory inter-item reliability in grades four and five for the Ice2, see section 3.5.1.2).

### 3.5.1.2 Orðskilningur (OSK) – a vocabulary subtest from the Icelandic testing battery for grade four- seven

Receptive word skills were also tapped in the vocabulary test Orðskilningur (Gunnarsson & Skúlason, 2009). This is a multiple-choice vocabulary test and was administered to groups of students. Participants were presented with sentences containing an underlined word. Their task was to select the correct meaning of that word from four possible options. Five of the 20 sentences in the test are idioms that need to be understood as whole units (see section 2.1.1.2).

Orðskilningur is based on the Icelandic Testing Battery for grades four and seven, for which words and sentences were originally randomly chosen. The test items were pre-tested and standardized with 900 participants in grade four, five, six, and seven, in randomly selected schools, in which whole classes participated, both Ice1 and Ice2 learners. The test was based on item response theory.

The inter-item reliability of this test was very low in the present study for the Ice2 participants in the first two grades, the lowest in fourth grade (Cronbach’s alpha = -.08), and fifth grade (a = .143), but satisfactory in sixth (a = .381), seventh (a = .494), and eighth grade (a = .392). The inter-item reliability was higher and satisfactory for the Ice1 in each grade (a =
.69 - .79). The unacceptable inter-item reliability in grades four and five for the Ice2 participants indicates that many of these fourth and fifth graders were guessing the answers or skipping the items to such an extent that the grand mean no longer represents systematic variation among the participants. However, higher inter-item reliability in grades six to eight indicates that their answers were increasingly based on guessing and thus increasingly reliable.

Validity was confirmed with correlation analysis. For Ice2, Orðskilningur did not correlate with PPVT in grades four and five, but with these scores in grades six, seven, and eight (r = .33 - .76). Orðskilningur correlated with Orðalykill in each grade (r = .30 - .65). For Ice1, Orðskilningur correlated with with Orðalykill in each grades (r = .63 - .82).

### 3.5.2 Receptive and productive word skills

One vocabulary test, Orðalykill, tapped both receptive and productive word skills.

#### 3.5.2.1 Orðalykill (OLY) - an Icelandic word definition test

Orðalykill (Gunnarsdóttir, et al., 2004) is a word definition test, which was administered individually. The test includes 55 items in order of difficulty. The participant was asked to give a definition for a word read out loud by the administrator. Testing was stopped when the child had made five errors in a row. The test thus taps both receptive and productive word skills, as it requires the child to produce an output (a synonym or other words that are related to the target word). Consequently, this test challenges an awareness of semantic relations (Proctor, et al., 2012), which is linked to vocabulary depth (Kieffer & Lesaux, 2012b) (see section 3.5.2.1 for further details of this test).

*Cronbach’s Alpha* for this measurement is reported .96. In the present study it was satisfactory for the Ice2 (.82 - .91) and higher for the Ice1 (.89 - .95), indicating inter-item reliability.

Validity of the test was confirmed with correlation analysis. Orðalykill correlated with PPVT for Ice2 in each grade (r = .61 - .76). For Ice2, Orðalykill correlated with Orðskilningur in each grade (r = .30 - .65). For Ice1 Orðalykill correlated with Orðskilningur in each grades (r = .63 - .82).
3.5.3 Free productive vocabulary: the writing test - What to do with the undeveloped space in the schoolyard?

The writing test is a discrete, free productive vocabulary measurement (see sections 2.2.2.2 and 2.2.4). Participants were asked to write a letter to the school principal and the teachers, proposing on how to use an empty space on the schoolyard. The number of tier 2 words used in the writing samples was counted. The choice of words that fell into the tier 2 category was based on the Icelandic word corpus, MÍM (Stofnun Árna Magnússonar í íslenskum fræðum, n.d.). Words that were beyond the most 1.100 frequent words were taken out and evaluated if they certainly could be classified as tier 2 words (not the words trampoline, skating, football, and others alike) (see a list of Icelandic tier 2 words used by the participants in appendix 7).

A holistic assessment of the quality of the writings was also included, on a Likert scale, excellent, proficient, adequate, and limited (see performance criteria-grid format, appendix 6). Those who did not meet requirements, writings with one sentence, no beginning and not an end, were rated limited. Writings including an opening of the statement, a middle chapter with support for argument(s) and an ending, without being convincing and not addressing the reader in any way, were rated adequate. Writings that were convincing, addressing the reader, including an opening, a middle chapter with support for argument(s) and an end, were rated proficient. Only writings that were outstanding, amusing, and convincing, including an opening, a middle chapter with good and solid supports for the arguments, and an ending, were rated excellent.

The writing test was administered to the sixth and the eighth graders at Testing 3. This was highly relevant taking into account that throughout the middle school years, after frequent encounters with tier 2 words in textbooks and schoolwork, these words should gradually emerge in learners’ productive language use (see section 2.1.1.1). Furthermore, by giving the task at Testing 3 it was possible to detect how word skills in fourth grade (the younger group) and in sixth grade (the older group) related with the use of tier 2 words in sixth and eighth grade and, additionally, predicted productive language skills in same grades.

Instructions were handed out or printed on the blackboard. All children received the same instructions (see appendix 5). In some cases, when there were only a few participants in a school, teachers or principals administered
the test, as it was too time demanding to administer it at the same time as the other tests. In order to obtain writing samples representing the children’s best efforts in retrieving and using their optimal productive word skills, the children were told that everyone, no matter how many, who wrote an excellent letter, with solid arguments for their suggestions and who addressed the reader in a convincing way, would receive a prize. (This promise will be fulfilled after defence.)

The prompt was selected from Roessingh and colleagues’ work (Roessingh, Elgie, & Pat, 2013). This test is a part of a large-scale assessment program administered to all students in central Alberta, known as The Highest Level Achievement Test (HLAT).

3.5.4 Reading skills

For the measurement of reading comprehension, two assessments were administered to Ice2 participants; a translation of the NARA test (Neale, 1997) and an Icelandic reading comprehension test from the Icelandic Testing Battery (Lesskilningspróf Námsmatsstofnunar, L.N.). The former was not administered to the Ice1 participants as it is intended for learners under the age of ten. These tests are in different formats but both include questions that challenge different reading strategies (see section 2.9).

3.5.4.1 Reading comprehension

Two reading comprehension tests were administered to the Ice2 participant, NARA and Lesskilningspróf Námsmatsstofnunar, and the latter test to the comparison group of Ice1 participants.

3.5.4.1.1 NARA – an Icelandic translation

This reading comprehension test is an Icelandic translation (by Ragnarsdóttir) of four stories of the NARA test (Neale Analysis of Reading Ability, NARA II). The original NARA test by Neale (1997) includes in total five stories that have been standardized for children from age six to twelve. The Icelandic translation was based on the British edition with support from the Norwegian translation by Lervåg & Aukrust (2010), and has not been adapted or validated for Icelandic speaking children (see section 2.2.1).

The test was administered individually where the child read short stories of increasing difficulty and length out loud. After each story the child was asked to give answers to a set of open-ended questions (ranging from four to
eight questions). The comprehension questions require both literal and inferential understanding (see section 2.9 for further details) (Cain & Oakhill, 2006), and require the reader to look for the answer within the text, demanding both decoding and comprehension skills, both of which contributes to the construct validity of the test (see section 1.2.1).

The internal consistency of this test was satisfactory in the present study (Cronbach’s alpha = .79 - .89). NARA correlated with L.N. scores for Ice2 (r = .52 - .71), representing validity of the measurement.

3.5.4.1.2 Lesskilningspróf Námsmatstofnunar - a reading comprehension subtest from the National testing battery for grades four to ten

The reading comprehension test Lesskilningspróf Námsmatstofnunar (henceforth L.N.) consists of three parts; one is intended for fourth to fifth graders, the second for children in grades six to seven, and the third for children in grades eight to ten (Gunnarsson & Skúlason, 2009). The test is based on a multiple-choice format and is administered to groups of children.

The two first parts comprise three separate texts; two narratives and one expository text, and the third part four separate texts; two narratives, one expository text, and a poem. Each text is followed by ten comprehension questions and the poem by six questions. The questions are of different levels of difficulty and require literal, inferential, and evaluative comprehension (see section 2.9). One of the narratives is common to the first two test parts, and the expository text in all three test parts, providing a sequence of scores from grade four through ten, making it possible to track improvements in the children’s scores across all grades (Gunnarsson & Skúlason, 2009).

The texts and the comprehension questions were selected from the Icelandic Testing Battery. These were pre-tested and standardized with 900 participants in randomly selected schools. Whole classes participated and included both Ice1 and Ice2 learners. The measurement was based on item response theory.

Inter-item reliability of this test is reported Cronbach’s Alpha .851. In the present study it was lower for the Ice2 (Cronbach’s Alpha = .43 - .74) than for the Ice1 (Cronbach’s Alpha = .70 - .80). For Ice2, L.N. scores correlated with NARA scores (r = .52 - .71), indicating validity of both tests.
3.5.4.2 Decoding skills - Reading fluency

Reading fluency measurements were included, as reading fluency (i.e. rate and accuracy) plays an important role in the middle school years when the role of reading changes from ‘learning to read’ to ‘reading to learn’ (see section 2.6.1.1) (cf. Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, & Snow, 2011).

For the younger group a sub-test of the screening test Leið til læsis was used (Snorradóttir, Sigurðardóttir, Sigurmundsdóttir, Skúlason, & Torfadóttir, 2010) and for the older group a norm-referenced reading fluency measurement (developed and widely used in educational programs by Steinunn Torfadóttir and Helga Sigurmundsdóttir). In both tests, the child is asked to read a short story out loud and final scores represent the number of correctly read words in a period of two minutes. Both tests were administered individually.

3.5.5 Screening for specific language impairment - A non-word repetition test

Each child was assessed individually, where the test-taker was asked to repeat a non-word spoken by the experimenter. Before the testing started two examples were given followed by a list of 25 word-like items and 25 non-word-like items. Each item included one to five syllables. The test was scored according to details given by Dollaghan and Campbell (1998) and involves counting the percentage of phonemes produced correctly. (Thordardottir & Juliusdottir, 2012, p. 8).

Thordardottir and Juliusdottir (2012) suggest from their findings that a non-word repetition test can be used as a screening tool for the presence of specific language impairment (Vance, 2008) in Ice2 learners. Language impairment can be defined as a developmental language disorder, and is not a term for language difficulties caused by bilingualism (Hulme & Snowling, 2010, pp.129-171.

High levels of diagnostic accuracy with bilingual children with SLI have been reported for Spanish by Girbau and Schwartz (2008) and for French by Elin Thordardottir and Brandeker (under review). However, substantially lower accuracy levels were reported by several studies focusing on L2 speakers in the US (Gutierrez- Clellen and Simon-Cereijido 2010; Kohnert,

In the study of Thordardottir and Juliusdottir, Ice2 participants did not underperform Ice1 children on this test and they had considerably higher scores than Ice1 children with language impairment, suggesting an absence of such discrepancy. The finding demonstrated that “…the ability to repeat Icelandic non-words did not require age-appropriate knowledge of the Icelandic language structures…such as vocabulary, syntax and morphosyntax” (Thordardottir & Juliusdottir, 2012, p. 20).

Same non-word repetition test as used in the study of Thordardottir and Juliusdottir was administered to both Ice2 and Ice1 participants. This measurement is based on a test by Dollaghan and Campbell (1998), but the Icelandic version was constructed by Thordardottir (2008).

3.6 Analytic approach

SPSS statistics (Statistical Package for the Social Sciences), and Mplus with maximum likelihood estimation (Muthén & Muthén, 2010), were used for analysing the data.

Descriptive statistics and reliability analysis provided information about the psychometric properties of each construct separately, giving the mean and standard deviation of scores for Ice2 and Ice1 age cohorts. Skewness and Kurtosis were used to provide information about the distribution of scores and the Sapiro Wilk test to inform whether the distribution of scores differed significantly from normality ($p < .5$). These represent scores for the younger group in grades four, five, and six, and for the older group in grades six, seven, and eight. Thus the two age groups are combined in grade six.

The longitudinal data from Testing 1, 2, and 3 were fitted into a growth curve model, also called latent variable growth model analysis. The growth model analysis made it possible to compare the rate and the shape of growth of vocabulary and reading comprehension skills over the middle school years between the Ice2 and Ice1 learners. The main question concerned whether the Ice2 learners improved their word and reading comprehension skills at the same rate as their Ice1 age peers, indicated by a constant gap between the groups, a slower growth, indicated by a widening gap between the groups, or an accelerated growth, indicated by a diminishing gap. This approach also gave information about individual trajectories, based on
initial scores: whether those who started with higher vocabulary and reading comprehension scores increased these at a faster rate than those who started with lower scores, at the same rate, or at a slower rate.

Still another advantage of growth model analysis is that it tolerates systematic missing values, which provides the possibility of an accelerated design as was employed in the present study. This approach made it possible to obtain a continuous developmental view from grade four through grade eight, from two different age groups (Duncan, Duncan, Strycker, Li, & Alpert, 1999). For this purpose data from the younger group were entered into the analysis for grade four (Testing 1; \( n = 49 \)), grade five (Testing 2; \( n = 49 \)), and grade six (Testing 3; \( n = 38 \)) and from the older group for grade six (Testing 1; \( n = 41 \)), grade seven (Testing 2; \( n = 39 \)), and grade eight (Testing 3; \( n = 34 \)); thus the two age groups overlapped in grade six (\( n = 89 \)). The two age groups of the Ice2 participants and the Ice1 participants did not differ significantly in grade six, except for the Ice2 participants on Orðskilningar, where the younger group outperformed the older (\( t(87) = 2.28 \); \( p < .05 \)).

The writing test was administered at Testing 3, to the sixth and the eight graders. Two-way ANOVA was conducted to compare Ice2 and Ice1 learners in their use of tier 2 words and in rubric scores. Furthermore, ANOVA and Bonferroni tests were used to compare the use of tier 2 words between writers of different proficiency levels. Correlation and regression analysis was used to analyse the extent to which these were related to and predicted by earlier vocabulary measures.
4 Findings

Before finding answers to the research questions descriptive analysis was conducted for each measurement, giving the mean and standard deviation of scores for Ice2 and Ice1 age cohorts separately. *Skewness* and *Kurtosis* were used to provide information about the distribution of scores and the *Sapiro Wilk* test to inform whether the distribution of scores differed significantly from normality (*p* < .05). These represent scores for the younger group in grades four, five, and six, and for the older group in grades six, seven, and eight. Thus the two age groups are combined in grade six.

4.1 Descriptive statistics

4.1.1 Means and standard deviation on vocabulary measurements for each age cohort

4.1.1.1 PPVT: Ice2

The mean scores and standard deviations for each age cohort of Ice2 participants in the PPVT vocabulary tests are presented in Table 4.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>49</td>
<td>99.00</td>
<td>24.50</td>
<td>44</td>
<td>134</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>115.88</td>
<td>22.10</td>
<td>54</td>
<td>148</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>116.57</td>
<td>24.64</td>
<td>42</td>
<td>159</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>123.64</td>
<td>21.27</td>
<td>79</td>
<td>155</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>131.53</td>
<td>18.77</td>
<td>99</td>
<td>155</td>
</tr>
</tbody>
</table>

The number of items on this test is 165, but no participants reached that level, as can be seen in Table 4. Ceiling effects were, however, detected among the eighth graders, which is evident from the limited variance of test scores in grade eight, and the distribution of individual scores in grade eight was significantly flat (*Kurtosis* = -1.174; *Shapiro Wilk*; *p* < .05).
Distribution of scores gave \textit{Skewness} and \textit{Kurtosis} $< |1|$ for other age cohorts.

\textbf{4.1.1.2 \textit{Orðskilningur: Ice2 and Ice1}}

The mean scores and standard deviations for each age cohort of Ice2 and Ice1 participants on the vocabulary measurement \textit{Orðskilningur} are presented in Table 5.

\textbf{Table 5. Ice2 and Ice1 participants’ mean raw scores and standard deviation on the vocabulary measurement \textit{Orðskilningur} for each age cohort}

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number</th>
<th>Ice2</th>
<th>Ice1</th>
<th>Ice2</th>
<th>Ice1</th>
<th>Ice2</th>
<th>Ice1</th>
<th>Ice2</th>
<th>Ice1</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>49</td>
<td>4.61</td>
<td>7.98</td>
<td>1.77</td>
<td>3.33</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>5.00</td>
<td>9.44</td>
<td>1.97</td>
<td>3.81</td>
<td>2</td>
<td>3</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>5.13</td>
<td>10.14</td>
<td>2.34</td>
<td>3.83</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>6.05</td>
<td>11.13</td>
<td>2.65</td>
<td>4.19</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>36</td>
<td>6.61</td>
<td>12.12</td>
<td>2.41</td>
<td>3.71</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>19</td>
</tr>
</tbody>
</table>

Possible range of scores on this test is 0-20, thus the table demonstrates that scores reached both the highest and the lowest level for the Ice1, and only the lower level for the Ice2 children. However, the distribution of individual scores within the group of Ice2 participants produced \textit{Skewness} and \textit{Kurtosis} $< |1|$ for each age cohort. The same applied to the distribution of individual scores within the comparison group of Ice1 learners, \textit{Skewness} and \textit{Kurtosis} were less than $|1|$ for each age cohort, except for the seventh graders but this did not differ significantly from a normal distribution ($\textit{Kurtosis} = -1.092; \textit{Shapiro Wilk}; p > .05$).

For the Ice2 group, the standard deviation increased from grade four to grade seven, meaning that the variation of individual scores increased between these grades. For the Ice1 learners, the standard deviation was also the highest in grade seven.

For the Ice2 children, the maximum score in each grade was a little higher than the average score for the Ice1. The Ice2 and the Ice1 groups are compared on these vocabulary scores in section 4.1.1.2, under research question 1.
4.1.1.3 Orðaleykill: Ice2 and Ice1

Mean scores and standard deviations for each age cohort of the Ice2 and the Ice1 participants on the vocabulary measurement Orðaleykill are presented in Table 6.

Table 6. Ice2 and Ice1 participants’ mean raw scores and standard deviation on the vocabulary measurement Orðaleykill for each age cohort

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice2</td>
<td>Ice1</td>
<td>Ice2 Ice1</td>
<td>Ice2 Ice1</td>
<td>Ice2 Ice1</td>
<td>Ice2 Ice1</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>41</td>
<td>3.12 17.00</td>
<td>2.98 6.95</td>
<td>0 6</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>37</td>
<td>5.16 21.59</td>
<td>4.30 8.80</td>
<td>1 6</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>92</td>
<td>5.33 23.75</td>
<td>4.76 8.81</td>
<td>0 5</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>48</td>
<td>6.87 26.98</td>
<td>5.97 10.33</td>
<td>1 6</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>48</td>
<td>8.03 28.60</td>
<td>6.05 10.01</td>
<td>0 8</td>
</tr>
</tbody>
</table>

Scores on this test can range from 0 to 55, and the table demonstrates that the Ice2 learners only reached the lower level, whereas the Ice1 learners reached neither the highest nor the lowest level. In fact, within the group of Ice2 participants, floor effects were detected. The distribution was positively and significantly skewed, giving Skewness > |1| for each grade, meaning that the majority of scores was clustered at the lowest part of the continuum (in grade four (Skewness = 1.486, Kurtosis = 1.724; Shapiro Wilk p = .00), grade five (Skewness = 1.222, Kurtosis = 0.301; Shapiro Wilk p = .00), grade six (Skewness = 1.258, Kurtosis = 1.147; Shapiro Wilk p = .00), grade seven (Skewness = 1.109, Kurtosis = 0.563; Shapiro Wilk p = .00), and in grade eight (Skewness = 1.107, Kurtosis = 0.871; Shapiro Wilk p < .05). Within the group of Ice1 participants, on the other hand, the distribution of scores gave Skewness and Kurtosis < |1| for each age cohort.

For both the Ice2 and the Ice1 participants, the standard deviation increased throughout the research period. This means that individual differences widened from grade four to grade eight.

For the Ice2 group, the maximum score in each grade is lower than the average score for the Ice1 learners. The two groups, Ice2 and Ice1, are compared on this vocabulary measurement in section 4.1.1.3, under research question 1.
4.1.2 Means, and standard deviation on reading comprehension measurements for each age cohort

As outlined in section 3.1 the two reading comprehension tests, the NARA (Neale, 1997) and Lesskilningspréf Námsmatstofnunar (L.N., Gunnarsson & Skúlason, 2009) were administered to the Ice2 participants at all three time points, and the latter was administered to the Ice1 group concomitantly.

4.1.2.1 The NARA reading comprehension test

Mean scores and standard deviations for each age cohort of the Ice2 participants on the NARA test are presented in Table 7.

Table 7. Ice2 participants mean raw scores and standard deviation on the reading comprehension test NARA

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>49</td>
<td>13.00</td>
<td>5.61</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>17.29</td>
<td>4.52</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>18.40</td>
<td>3.88</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>20.56</td>
<td>3.89</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>36</td>
<td>22.06</td>
<td>3.10</td>
<td>15</td>
<td>26</td>
</tr>
</tbody>
</table>

Scores on this test can range from 0 to 27 and the table demonstrates that scores reached the highest level in grade seven, and the maximum score is only one score lower in grades six and eight. The distribution of individual scores produced, nonetheless Skewness and Kurtosis < |1| for each age cohort, except for the fourth graders when there was a flat distribution with Kurtosis = -1.276, differing significantly from a normal distribution, Shapiro Wilk p < .05. This means that no participants scored high in grade four.

Ceiling effects were detected on this test, as the standard deviation decreased from grade four to grade five and again between grade five and grade six, with the two age groups combined, and was the smallest in grade eight (see Table 7).
4.1.2.2 Lesskilningspróf Námsmatsstofnunar, L.N.: Ice2 and Ice1

Mean scores and standard deviations for each age cohort of Ice2 and Ice1 participants on the L.N. reading comprehension test are demonstrated in Table 8.

Table 8. Ice2 and Ice1 participants’ mean raw scores and standard deviation on the reading comprehension test Lesskilningspróf Námsmatsstofnunar, L.N.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Ice2 Number</th>
<th>Ice1 Number</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>49</td>
<td>41</td>
<td>13,08</td>
<td>3.45</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>37</td>
<td>15,39</td>
<td>3.92</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>92</td>
<td>15,58</td>
<td>4.87</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>47</td>
<td>18,18</td>
<td>4.13</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>41</td>
<td>55</td>
<td>15,83</td>
<td>5.44</td>
<td>3</td>
<td>25</td>
</tr>
</tbody>
</table>

Possible range of scores on this test is 0 - 30 in grades four to seven, and 0 - 36 in grade eight. Table 8 demonstrates that for both the Ice2 and the Ice1 group neither the lowest nor the highest level was reached in any grade. However, the highest Ice1 scorer in grade five, six, and seven had 29 scores, with only one wrong answer. Distribution of individual scores produced Skewness and Kurtosis < |1| for each age cohort of both the Ice2 and the Ice1 group.

The standard deviation increased over these years for the Ice2 learners, revealing increasing individual differences, whereas remaining almost the same from grade four to grade eight for the Ice1 group.

For the Ice2 children, the maximum score in each grade is higher than the average score of the Ice1 peers. The Ice2 and the Ice1 groups are compared on this test in section 4.1.2.2, under research question 1.

4.1.2.3 The reading fluency test

The reading fluency test was administered at Time point 1, to the fourth and the sixth graders. This measurement was used as a time-invariant covariate in predicting reading comprehension scores and development, under research question 2 (see sections 4.2.1 and 4.2.2). Means scores and standard deviations for both age groups of the Ice2 and the Ice1 participants are demonstrated in Table 9.
For each group, Skewness and Kurtosis was < |1|, revealing no floor or ceiling effects.

As presented in Table 9, the Ice2 learners had poorer reading fluency than the Ice1 group, in grade four and grade six. To compare the Ice2 group with the Ice1 group in grades four and six, two-way ANOVA was conducted with grade (four and six) and language group (Ice2 and Ice1) as fixed factors. The main effect of language group was significant ($F (1) = 31.22$; $p = .00$) and the main effect of grade was also significant ($F (1) = 25.46$; $p = .00$). Interaction effect between language group and grade was not significant ($F (1) = 1.01$; $p > .05$). This means that the difference between the two language groups was significant, as was the difference between the two grades, and that the difference between grades four and six was the same for the Ice2 as for the Ice1 learners.

### 4.1.3 Screening for language impairment: a non-word repetition test

The Ice2 participants (apart from the one who was excluded from the analysis, see section 3.4) produced from 95% to 100% correct non-words on the non-word repetition task. Three of the Ice1 participants produced less than 95% correct non-words on same task (see section 3.4).

### 4.1.4 Correlation analysis

Table 10 demonstrates correlations between predictive and outcome variables within each grade for the two Ice2 and Ice1 age groups (younger and older).
Table 10. Correlation (Pearson R) between predictive variables and outcome measures

<table>
<thead>
<tr>
<th>Outcome/Measurement Grade 4</th>
<th>NARA Grade 4</th>
<th>NARA Grade 5</th>
<th>NARA Grade 6</th>
<th>L.N. Grade 4</th>
<th>L.N. Grade 5</th>
<th>L.N. Grade 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT</td>
<td>.59**</td>
<td>.53**</td>
<td>.53**</td>
<td>.36*</td>
<td>.63**</td>
<td>.42**</td>
</tr>
<tr>
<td>Orðskilningur</td>
<td>.13</td>
<td>.11</td>
<td>.04</td>
<td>.29*</td>
<td>.08</td>
<td>.21</td>
</tr>
<tr>
<td>Orðalykill</td>
<td>.56**</td>
<td>.51**</td>
<td>.43**</td>
<td>.48**</td>
<td>.47**</td>
<td>.42**</td>
</tr>
<tr>
<td>Fluency</td>
<td>.72**</td>
<td>.69**</td>
<td>.65**</td>
<td>.43**</td>
<td>.57**</td>
<td>.56**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome/Measurement Grade 6</th>
<th>NARA Grade 6</th>
<th>NARA Grade 7</th>
<th>NARA Grade 8</th>
<th>L.N. Grade 6</th>
<th>L.N. Grade 7</th>
<th>L.N. Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT</td>
<td>.82**</td>
<td>.67**</td>
<td>.54**</td>
<td>.70**</td>
<td>.48**</td>
<td>.70**</td>
</tr>
<tr>
<td>Orðskilningur</td>
<td>.48**</td>
<td>.39*</td>
<td>.36*</td>
<td>.51**</td>
<td>.40*</td>
<td>.51**</td>
</tr>
<tr>
<td>Orðalykill</td>
<td>.56**</td>
<td>.52**</td>
<td>.41*</td>
<td>.48**</td>
<td>.44**</td>
<td>.48**</td>
</tr>
<tr>
<td>Fluency</td>
<td>.16</td>
<td>.20</td>
<td>.29</td>
<td>.21</td>
<td>.27</td>
<td>.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome/Measurement Grade 4/6</th>
<th>L.N. Grade 4</th>
<th>L.N. Grade 5</th>
<th>L.N. Grade 6</th>
<th>L.N. Grade 6</th>
<th>L.N. Grade 7</th>
<th>L.N. Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orðskilningur</td>
<td>.52**</td>
<td>.62**</td>
<td>.43**</td>
<td>.66**</td>
<td>.62**</td>
<td>.66**</td>
</tr>
<tr>
<td>Orðalykill</td>
<td>.60**</td>
<td>.55**</td>
<td>.50**</td>
<td>.63**</td>
<td>.70**</td>
<td>.63**</td>
</tr>
<tr>
<td>Fluency</td>
<td>.57**</td>
<td>.51**</td>
<td>.37*</td>
<td>.57**</td>
<td>.47**</td>
<td>.57**</td>
</tr>
</tbody>
</table>

Table 10 reveals, the reading fluency scores of the two Ice1 age groups and the younger Ice2 group correlated moderately with their reading comprehension scores in all grades, and in fourth grade there was a strong relationship with NARA scores for the Ice2 group. For the older Ice2 group, on the other hand, the correlation between reading fluency and reading comprehension did not reach statistical significance in any grade.

The vocabulary measurements were moderately and strongly correlated with reading comprehension scores for both Ice2 and Ice1 participants. However, scores on Orðskilningur for the younger Ice2 group correlated only with L.N. scores in grade four, but not in grades five and six.

Table 11 demonstrates correlations between background variables and vocabulary measurements.

---

*6 ** \( p < .01; * p < .05 \)
Table 11. Correlations (Pearson R) between background variables and vocabulary measures

<table>
<thead>
<tr>
<th>Ice2 younger</th>
<th>Background/Measurement</th>
<th>Age of arrival</th>
<th>First lang. proficiency</th>
<th>First lang. literacy</th>
<th>First lang. instruction</th>
<th>Mothers’ education</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT Grade 4</td>
<td>- .21</td>
<td>.17</td>
<td>.04</td>
<td>.16</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>PPVT Grade 5</td>
<td>.02</td>
<td>.23</td>
<td>.08</td>
<td>.25</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>PPVT Grade 6</td>
<td>- .00</td>
<td>.25</td>
<td>-.03</td>
<td>.33*</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 4</td>
<td>.03</td>
<td>.14</td>
<td>-.09</td>
<td>.01</td>
<td>-.08</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 5</td>
<td>.13</td>
<td>-.03</td>
<td>.15</td>
<td>.21</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 6</td>
<td>.09</td>
<td>.18</td>
<td>-.03</td>
<td>.22</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 4</td>
<td>.08</td>
<td>.18</td>
<td>.13</td>
<td>.22</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 5</td>
<td>.03</td>
<td>.07</td>
<td>.06</td>
<td>.21</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 6</td>
<td>-.00</td>
<td>.09</td>
<td>.03</td>
<td>.23</td>
<td>-.04</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ice2 older</th>
<th>Background/Measurement</th>
<th>Age of arrival</th>
<th>First lang. proficiency</th>
<th>First lang. literacy</th>
<th>First lang. instruction</th>
<th>Mothers’ education</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT Grade 6</td>
<td>- .65**</td>
<td>-.32*</td>
<td>-.31*</td>
<td>-.16</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>PPVT Grade 7</td>
<td>- .39*</td>
<td>-.30</td>
<td>.19</td>
<td>-.15</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>PPVT Grade 8</td>
<td>- .55**</td>
<td>-.15</td>
<td>.11</td>
<td>-.12</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 6</td>
<td>- .21</td>
<td>.06</td>
<td>.05</td>
<td>-.04</td>
<td>-.03</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 7</td>
<td>-.22</td>
<td>-.16</td>
<td>.08</td>
<td>.07</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>Orðskilningur Grade 8</td>
<td>-.16</td>
<td>-.06</td>
<td>.11</td>
<td>-.12</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 6</td>
<td>- .55**</td>
<td>-.23</td>
<td>-.21</td>
<td>.04</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 7</td>
<td>-.40*</td>
<td>-.08</td>
<td>-.00</td>
<td>.01</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>Orðhálykill Grade 8</td>
<td>-.36</td>
<td>.01</td>
<td>.12</td>
<td>.16</td>
<td>.11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ice1 younger</th>
<th>Measurement/Background</th>
<th>Orðskilningur Grade 4</th>
<th>Orðskilningur Grade 5</th>
<th>Orðskilningur Grade 6</th>
<th>Orðhálykill Grade 4</th>
<th>Orðhálykill Grade 5</th>
<th>Orðhálykill Grade 6</th>
<th>Mothers’ education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Orðskilningur Grade 6</td>
<td>.25</td>
<td>.18</td>
<td>.16</td>
<td>.35*</td>
<td>.31</td>
<td>.40*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ice1 older</th>
<th>Measurement/Background</th>
<th>Orðskilningur Grade 6</th>
<th>Orðskilningur Grade 7</th>
<th>Orðskilningur Grade 8</th>
<th>Orðhálykill Grade 6</th>
<th>Orðhálykill Grade 7</th>
<th>Orðhálykill Grade 8</th>
<th>Mothers’ education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Orðskilningur Grade 6</td>
<td>.28*</td>
<td>.31*</td>
<td>.26</td>
<td>.34*</td>
<td>.38**</td>
<td>.34*</td>
<td></td>
</tr>
</tbody>
</table>
As Table 11 demonstrates, first language instruction correlated positively with PPVT scores for the younger Ice2 group in grade six. On the other hand, for the older Ice2 group, first language proficiency and literacy correlated negatively and significantly with PPVT scores in same grade. Age of arrival correlated negatively with PPVT scores for the older Ice2 group in all three grades, and with Orðalykill in grades six and seven. These were the only significant correlations between background variables and vocabulary measures for the Ice2 participants.

For the Ice1 participants, Orðskilningur and maternal education correlated only in grades six (older group) and seven, whereas maternal education and Orðalykill correlated in all grades, but not in grade five.

Table 12 demonstrates correlations between background variables and reading comprehension measures.

### Table 12. Correlations (Pearson R) between background variables and reading comprehension measures

<table>
<thead>
<tr>
<th>Ice2 younger</th>
<th>Ice2 older</th>
<th>Ice1 younger</th>
<th>Ice1 older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background/Measurement</td>
<td>Age of arrival</td>
<td>First lang. proficiency</td>
<td>First lang. literacy</td>
</tr>
<tr>
<td>NARA Grade 4</td>
<td>.08</td>
<td>.09</td>
<td>.05</td>
</tr>
<tr>
<td>NARA Grade 5</td>
<td>.09</td>
<td>.20</td>
<td>.29*</td>
</tr>
<tr>
<td>NARA Grade 6</td>
<td>.09</td>
<td>.19</td>
<td>.04</td>
</tr>
<tr>
<td>L.N. Grade 4</td>
<td>.11</td>
<td>.09</td>
<td>.05</td>
</tr>
<tr>
<td>L.N. Grade 5</td>
<td>.06</td>
<td>.19</td>
<td>.26</td>
</tr>
<tr>
<td>L.N. Grade 6</td>
<td>.10</td>
<td>.03</td>
<td>.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ice2 older</th>
<th>Ice1 younger</th>
<th>Ice1 older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background/Measurement</td>
<td>Age of arrival</td>
<td>First lang. proficiency</td>
</tr>
<tr>
<td>NARA Grade 6</td>
<td>-.45**</td>
<td>-.17</td>
</tr>
<tr>
<td>NARA Grade 7</td>
<td>-.25</td>
<td>-.18</td>
</tr>
<tr>
<td>NARA Grade 8</td>
<td>-.16</td>
<td>-.13</td>
</tr>
<tr>
<td>L.N. Grade 6</td>
<td>-.41**</td>
<td>-.05</td>
</tr>
<tr>
<td>L.N. Grade 7</td>
<td>-.21</td>
<td>-.09</td>
</tr>
<tr>
<td>L.N. Grade 8</td>
<td>-.41**</td>
<td>-.05</td>
</tr>
<tr>
<td>Mothers’ education</td>
<td>.24</td>
<td>.22</td>
</tr>
</tbody>
</table>

125
Table 12 demonstrates that first language literacy correlated positively with NARA reading comprehension scores in grade five for the Ice2 learners. First language instruction also correlated positively with NARA scores in grades five and six and with L.N. scores in grade six for the younger Ice2 group. However, first language instruction correlated negatively and significantly with L.N. scores in grades six and eight for the older Ice2 group. Age of arrival correlated negatively with NARA scores in grade six and with L.N. scores in grades six and eight for the older Ice2 group. These were the only significant correlations between background variables and reading comprehension scores for the Ice2 participants. For the Ice1 group, maternal education only correlated with L.N. scores in grade seven.

Finally, correlation analysis was conducted on the relations between background variables for the Ice2 group and demonstrated in Table 13.

Table 13. Correlations (Pearson R) between background variables for the Ice2 group

<table>
<thead>
<tr>
<th>Background/Measurement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age of arrival</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. First lang. proficiency</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. First lang. literacy</td>
<td>.39**</td>
<td>.40**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. First lang. instruction</td>
<td>.08</td>
<td>.14</td>
<td>.23*</td>
<td></td>
</tr>
<tr>
<td>5. Mothers’ education</td>
<td>-.08</td>
<td>.30**</td>
<td>.04</td>
<td>.13</td>
</tr>
</tbody>
</table>

Table 13 reveals that the older the children arrived in Iceland the more literate they were, however age of arrival and first language proficiency were not significantly related. Moreover, the more first language instruction the children had obtained the more literate they were. First language literacy and proficiency were also correlated. Finally, the higher the educational level of the mothers the more proficient the children were in their first language.
4.2 Research question 1

To explore whether there was an initial and a developmental difference in the vocabulary and reading comprehension skills of the Ice2 and the Ice1 learners, growth model analysis was conducted, with an accelerated design (see section 3.6.). This approach made it possible to compare the progress made by the two groups on the vocabulary and reading comprehension measurements from grade four and all the way up to grade eight. In growth model analysis the two growth factors, the intercept (the starting point) and the slope (the rate of growth) were treated as dependent variables, making it possible to compare the Ice2 and the Ice1 groups on these two parameters. Thus, in addition to initial scores, the rate of growth could be compared between the Ice2 and the Ice1 children to see whether one group demonstrated a slower growth than the other and thus how the gap between them developed throughout the middle school years. Moreover, growth model analysis made it possible to compare the shape of development between the two groups, whether there was a linear growth (stable increase across grades) or a non-linear growth (unstable increase between grades). These models also gave information about the covariance between the intercept and the slope, i.e. how individual differences developed over time; whether those with the highest scores initially increased these at a faster rate than those with the lowest scores (positive covariance), at the same rate (zero or not significant covariance), or at a slower rate (negative covariance). When a negative covariance occurs in growth model analysis a fit model can be difficult to obtain, as will be presented in the sections below.

This analysis was first conducted on the participants’ vocabulary scores, obtained at all the three time points. The measurements included the PPVT for the Ice2 group, and Orðskilningur for both groups. Growth model analysis on Orðalykill was, however, not possible as the distribution of scores for the Ice2 group differed significantly from a normal distribution in each grade, which was caused by floor effects (see section 4.1.1.3). It was not relevant to use logarithm of scores as in fact this measurement was too difficult to capture individual differences among the Ice2 participants.
4.2.1 Vocabulary growth patterns

4.2.1.1 PPVT

Figure 2 demonstrates the progress in mean scores made by the Ice2 participants between the five time points on the multiple-choice vocabulary test PPVT (see section 3.5.1.1), with the two age groups (younger and older, see section 2.6) overlapping in grade six.

As Figure 2 demonstrates, a sharp improvement in the Ice2 participants’ performance on the PPVT was observed between grades four and five, however, the same did not apply to the remainder of the study period. The line is flat between grades five and six, indicating no or limited progress, and there was only a small increase in mean scores from grade six to grade eight. This pattern was, however, most likely caused by ceiling effects that were detected on this test (see section 4.1.1.1).

In the growth model analysis, a fit model was difficult to obtain. The best model fit was represented by a non-linear growth curve, with factor loadings for two time points freed. The increase was thus not stable across the time of the study, which is consistent with Figure 2. Residual variances for grade four and seven were negative and had to be fixed at zero (see Geiser, 2013). This model gave the poor fit: Chi-Square (6) = 35.048 $p = .00$; $CFI = .891$; $TLI = .891$; $RMSEA = .232$ (90% C.I. = .161-.309; $p = .00$);

Figure 2. The development of Ice2 participants’ mean raw scores on the vocabulary measurement PPVT
SRMR = .276, with a mean intercept of 9.490 ($p = .00$) and mean slope of 1.014 ($p = .00$) (scores were divided by 10 to minimize variances which is necessary to obtain a fit model, the scores spanning from 42 – 159) (Muthén & Muthén, 2010). The covariance between the intercept and the slope was negative and significant ($s$ with $i = -0.930; p = .00$). This means that those with the highest PPVT scores initially demonstrated a slower increase in mean scores between grades four and eight than those with lower scores at the start of the study. The poor fit of the model was most likely caused by the negative covariance between the intercept and the slope, which indeed can make a fit model difficult to obtain (Muthén & Muthén, 2010).

Thus, the results of the growth model analysis gave the best fit with two time points freed, meaning that the increase was slow between two grades, as is consistent with Figure 2 demonstrating a flat line between grades five and six and slow increase from grade six to grade eight.

### 4.2.1.2 Orðskilningur

Figure 3 demonstrates the progress in mean scores made by the Ice2 and the Ice1 participants across the five time points on the vocabulary test Orðskilningur. In this test the participants were asked to select the correct meaning of a word or an idiom, from four possible options (see section 3.5.1.2). As before, the younger and older age groups of both the Ice2 and the Ice1 learners overlap in grade six (see section 3.6).
Figure 3 shows that the Ice2 children made very slow increase in mean scores from grade four to grade eight, and in particular between grades four and six. Their scores at the start of the study were well below the Ice1 participants’ scores and the gap between them increased as the study progressed. Thus, even though the scores of the Ice2 learners increased from grade six to grade seven, it did not suffice to stabilize or, better, narrow the gap between the two groups. The figure also demonstrates that the growth patterns for the Ice2 and the Ice1 learners from grade four to grade eight were not identical. Therefore, the two groups were not combined in one growth model analysis, but were analysed separately (Muthén & Muthén, 2010).

For the Ice2 learners, a good model fit was difficult to obtain. The best fit was represented by a non-linear growth curve, with factor loadings for one time point freed. Residual variances in grade four and grade eight were negative and had to be fixed at zero (Geiser, 2013). This model gave the poor fit Chi. Square (7) = 14.869; p = .0377; CFI = .761; TLI = .795; RMSEA = .112 (90% C.I. = .025-.191, p = .094); SRMR = 0.239), with mean intercept of 4.340 (p = .00) and mean slope of 0.457 (p = .00). Negative covariance between the intercept and the slope was significant (s with i = -0.838; p = .00). Thus, for the Ice2 group, the gap between those with the highest scores and those with the lowest scores initially decreased from grade four to grade eight.

For the Ice1 learners’ scores on Orðskilningur, the best model fit was represented by a linear growth curve, giving a fairly good fit: Chi. Square (6) = 14.462, p = .0249; CFI = .945; TLI = .945; RMSEA = .121 (90% C.I. = .041-.203, p > .05); SRMR = .129, with mean intercept 7.873 (p = .00) and mean slope 1.118 (p = .00). A linear growth curve means that there was a steady increase in mean scores for the Ice1 learners between all grades. The covariance between the intercept and the slope was negative, but not significant (s with i = -0.300; p > .05). This means that those Ice1 children with the highest scores initially increased these at the same rate as those with the lowest initial scores.

The better fit for the Ice1 group than the Ice2 group was most likely due to the negative significant covariance between the intercept and the slope
for the Ice2 group, whereas the negative covariance for the Ice1 group was not significant. A negative covariance between the intercept and the slope can indeed make a fit model difficult to obtain in growth model analysis.

Thus, on Orðskilningur the Ice2 group demonstrated a non-linear growth curve from grade four to grade eight, indicating that they did not make a steady improvement between grades. The Ice1 learners, on the other hand, demonstrated a linear growth curve, which means that their progress was constant across all grades. The mean intercept for the Ice2 group (4.340/s.e. 0.24) was lower than that of the Ice1 group (7.873/s.e.0.46) representing lower initial scores ($t(182) = -65.34; p = .00$). The mean slope for the Ice2 group (0.457/s.e. 0.11) was also lower than that of the Ice1 group (1.118/s.e.0.14) ($t(182) = -14.64; p =.00$) representing a slower growth for the former group from grade four to grade eight. This indicates a widening gap between the two groups, as represented by the growth patterns in Figure 3.

Taken together, results of the growth model analysis demonstrated a non-linear developmental pattern for the Ice2 group, with one time score freed, which is consistent with Figure 3 demonstrating a flat line between grades five and six. In contrast, the children in the Ice1 group made a significant increase between all grades.

4.2.1.3 Orðálykill

Figure 4 demonstrates the progress in mean scores made by the Ice2 and the Ice1 participants from grade four to grade eight on Orðálykill, which is a word definition test, thus tapping both receptive and productive word skills (see section 3.5.2.1).
Figure 4 demonstrates a large gap between the mean scores of the Ice2 and the Ice1 participants from grade four to grade eight. The Ice1 learners made steady progress between grades four and six, while there was a limited increase in the Ice2 learners’ mean scores across the same time period, as demonstrated by almost a flat line between grades five and six. The figure also demonstrates a widening gap between the two groups over the time of the study.

As stated before (see section 4.1), this vocabulary measurement was not submitted to growth model analysis due to the significant difference from a normal distribution detected in each age cohort for the Ice2 children.

4.2.2 Reading comprehension growth patterns

Two reading comprehension measures were administered in the study at all three time points. These were the NARA test, which was administered to the Ice2 participants only, and the L.N. test administered to all participants.

4.2.2.1 NARA

Figure 5 demonstrates the increase in mean scores made by the two Ice2 age groups (younger and older) on the NARA reading comprehension test, with the two age groups overlapping in grade six (see section 3.6).
Figure 5. The development of Ice2 mean raw scores on the reading comprehension test NARA

As Figure 5 reveals, the Ice2 participants made progress across all grades, however with slow increase between grades five and six, and with the sharpest improvements between grades four and five, and between grades six and seven. However, as discussed in section 2.7.2.1, there were some ceiling effects detected on this test, although the distribution did not differ significantly from normal in each grade.

Consistent with the growth lines presented in Figure 5, a non-linear growth curve provided the best model fit with factor loadings for one time point freed. The residual variance for grade eight was negative and had to be fixed at zero (Geiser, 2013), which gave a fairly good fit $\chi^2$ (6) = 17.663 $p < .05$; $CFI = .928$; $TLI = .928$; $RMSEA = .147$ (90% C.I. = .070-.229, $p < .05$); $SRMR = .126$, giving a mean intercept of 13.689 ($p = .00$) and a mean slope of 2.334 ($p = .00$). Negative covariance between the intercept and the slope was significant ($s$ with $t = -4.625$; $p = .00$), meaning that the gap between those with the highest initial scores and those who started with the lowest scores decreased from grade four to grade eight.

Thus, results of the growth model analysis revealed that the developmental pattern across grades four to eight was not linear, with one
time score freed, as is consistent with Figure 5, demonstrating slow increase between grades five and six.

4.2.2.2 L.N.

The developmental pattern in mean scores for each language group on the L.N. reading comprehension test is presented in Figure 6, with combined age groups in grade six (see section 3.6).

![Figure 6. The development of mean raw scores on the reading comprehension test, Lesskilningspróf Námsmatstofnanar L.N.](image)

As Figure 6 reveals, the increase of mean scores for both groups was not stable across the time of the study. Both demonstrated an increase in mean scores from grade four to grade five, but a nearly flat growth line emerged between grades five and six. After that time point, a sudden shift seems to have occurred with an accelerated growth between grades six and seven, followed by an extreme drop in mean scores between grades seven and eight. Thus, the reading comprehension scores of the Ice2 and Ice1 learners demonstrated almost an identically shaped growth curve across grades, with the Ice2 group constantly lagging behind.

To examine the difference between the Ice2 and Ice1 participants initially and in rate of growth on L.N. scores from grade four to grade eight,
growth model analysis was conducted with combined groups of Ice2 and Ice1 participants. This was possible as the two groups demonstrated identical developmental patterns. A dummy variable for the Ice2 group entered into the model in Step 1 informed about the difference between the two groups on the intercept (in grade four) and on the slope (rate of growth). An interaction effect entered into the model in Step 2 informed whether the developmental pattern was the same for both groups.

In spite of the sudden and strong shift downwards in growth lines of both groups from grade seven to grade eight, the shape of the growth curve from grade four to grade eight cannot be viewed as curvilinear, as there was not a constant slowing rate of growth from grade four to grade seven, when the line dropped down (Ice2: increase from grade four to five = 2.31 scores, from five to six = 0.19 scores, and from six to seven = 2.6 scores; Ice1: increase from grade four to five = 1.66 scores, five to six = 0.28 scores, from grade six to seven = 1.33 scores). In growth model analysis for a linear shape, correlation problem occurred between scores in grade six and grade eight (sample correlation $r = 1.00$), which means that both grades could not be included in the analysis. (This problem also emerged when growth model analysis was attempted with a quadratic shape). Scores of the eighth graders were thus taken out to demonstrate growth from grade four to grade seven.

Growth model analysis on the L.N. scores, with the combined groups of Ice2 and Ice1 learners, from grade four to grade seven gave a good fit for a non-linear growth curve, with factor loadings for two time points freed. The residual variances for grade four and grade seven were negative and had to be fixed at zero (Geiser, 2013). This model gave a good fit Chi-Square (3) = 4.922; $p > .05$; $CFI = .989$; $TLI = .985$; $RMSEA = .059$ (90% C.I. = .000-.148; $p > .05$); $SRMR = .128$, giving mean intercept = 16.193 ($p = .00$) and mean slope = 1.335 ($p = .00$). Negative covariance between the intercept and the slope was significant ($s$ with $i = -4.553; p = .00$). This means that for the combined groups of Ice2 and Ice1 learners, the gap between those who started with higher scores and those who started with lower scores decreased from grade four to grade seven.

The Ice2 and Ice1 groups were compared on the L.N. test by entering Ice2 as a dummy variable into the model in Step 1. The difference between the Ice2 and the Ice1 children was negative and significant on the intercept ($i$ on Ice2 = -6.138; $p = .00$) but positive and not significant on the slope ($s$ on Ice2 = 0.529; $p > .05$). Thus, the Ice2 participants started with lower scores, but the difference in the rate of growth between the two groups was
not significant, revealing a constant gap between the groups. Interaction covariate entered into the model in Step 2 did not exert a significant effect (\(s_{\text{inter}} = -0.049; p > .05\)), representing an identical development for the Ice2 and the Ice1 groups on the L.N. reading comprehension test, as is consistent with Figure 6.

To analyze the decline made by the Ice2 and Ice1 learners between grades seven and eight, repeated measures ANOVA was conducted. The repeated measures factor was grade (seven and eight) and the between-subjects factor was language group (the Ice2 and Ice1 groups). There was a significant main effect of grade \((F(1, 84) = 15.05; p = .00)\) with partial \(\eta^2 = .15\), which means that the two groups decreased their mean scores significantly between grades six and seven. The between-subjects factor was also significant \((F(1, 84) = 29.35; p = .00)\), whereas an interaction effect was not significant \((F(1, 84) = 1.29; p > .05)\). This means that the decrease from grade seven to grade eight was the same for the two groups, the Ice1 group outperforming the Ice2 group to the same extent in both grades.

Taken together, the results of the growth model analysis revealed that the L.N. reading comprehension scores of the Ice2 and Ice1 groups followed the same developmental pattern from grade four to grade eight. Both groups made a non-linear growth from grade four to grade seven, as the best fit was obtained with two time scores freed. Figure 6 demonstrates a flat line between grades five and six, and a slower increase between grades four and five than between grades six and seven. The gap between the two groups remained the same during the time of the study, as both groups increased these scores at the same rate from grade four to grade seven, and made the same decrease from grade seven to grade eight.

Growth model analysis in the L.N. model for the Ice2 and the Ice1 learners demonstrated a negative covariance between the intercept and the slope, which also emerged in the NARA model for the Ice2 group. This means that those who started with lower reading comprehension scores increased these scores at a faster rate than those who started with higher scores. Thus, on both tests, a regression to the mean emerged from the data. Those Ice2 and Ice1 learners with best reading comprehension skills in grade four were not able to accelerate these skills over time, as they demonstrated a slower improvement than those with poorer initial reading comprehension skills. Furthermore, in both models the mean improvement was unstable, from grade four to grade eight on the NARA test, and from grade four to grade seven on the L.N. test.
A negative covariance between the intercept and the slope causes problems in obtaining fit models in growth model analysis, which was the case in the current study as demonstrated below. A fit model can also be difficult to obtain when there are large individual variances in developmental patterns.

4.3 Research question 2

To seek answers to the second research question on the extent to which reading fluency and Icelandic word skills of the Ice2 and the Ice1 learners predicted their reading comprehension skills and its rate of growth, growth model analysis was conducted using the base models for the reading comprehension measurements as presented in section 4.1.2. The growth factors represented estimated initial reading comprehension scores (the intercept) and the rate at which these developed from grade four to grade eight (the slope) on the NARA test, and from grade four to grade seven on the L.N. reading comprehension test (see section 4.1.2.2). The two growth factors, the intercept and the slope, were then treated as dependent variables, which offered the possibility to predict their emergence (Muthén & Muthén, 2010). This was performed by entering time-invariant covariates, reading fluency and vocabulary scores (measured in grade four) into the models. In this way it was possible to obtain information about the extent to which initial reading comprehension scores (in grade four) and the rate at which these developed over the middle school years were predicted by the children’s initial reading fluency and vocabulary scores.

Two reading comprehension measures were used as outcomes in these analyses: The NARA test, which was only administered to the Ice2 group, and the L.N. test, which was administered to all participants (see section 3.5.4.1 for a full description of these tests). Two measurements assessing word skills were used as covariates in the growth models: Orðskilningur (see section 3.5.1.2) was administered to all participants, but the PPVT (see section 3.5.1.1) was only administered to the Ice2 group. As mentioned before, the scores on Orðalykill (see sections 3.7.1.3 and 4.1) for the Ice2 participants demonstrated floor effects in grade four (as in other grades) and, therefore, this vocabulary test could not be entered into growth model analysis as a time-invariant covariate. The reading fluency measure (see section 3.5.4.2) was administered to both the Ice2 and the Ice1 participants and scores in grade four entered into the growth models.
As a fit model was difficult to obtain for both reading comprehension tests (see sections 4.1.2.1 and 4.1.2.2), stepwise analysis was not applicable. Reading fluency and vocabulary measures were for that reason entered each one separately into the reading comprehension models. The results of the growth model analysis on the influence of reading fluency and vocabulary on the intercept and the slope in the NARA and in the L.N. reading comprehension models are summarized in Table 14. In the next two sections, when results from growth model analysis are presented, references to this table will be made.

Table 14. The influence of time-invariant measurements entered separately into the NARA and L.N. reading comprehension models on the intercept and the slope

<table>
<thead>
<tr>
<th>Reading fluency influence on intercept / slope</th>
<th>Vocabulary influence on intercept / slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>NARA Ice2</td>
<td></td>
</tr>
<tr>
<td>Fluency: 0.689* / 0</td>
<td>PPVT: 0.549* / 0.147*</td>
</tr>
<tr>
<td></td>
<td>Orðskilningur: -0.182* / 0.333**</td>
</tr>
<tr>
<td>L.N. Ice2</td>
<td></td>
</tr>
<tr>
<td>Fluency: 1.583** / -0.321**</td>
<td>PPVT: 0 / 0.159**</td>
</tr>
<tr>
<td></td>
<td>Orðskilningur: -0.164* / 0.249**</td>
</tr>
<tr>
<td>L.N. Ice1</td>
<td></td>
</tr>
<tr>
<td>Fluency: 0.649** / 0</td>
<td>Orðskilningur: 0 / 0.143**</td>
</tr>
</tbody>
</table>

4.3.1 The influence of initial reading fluency and vocabulary scores on initial NARA scores and its rate of growth

Growth model analysis was conducted on the NARA test with scores on the reading fluency assessment and the two vocabulary measures (PPVT and Orðskilningur) as time-invariant covariates. All these tests were administered in grade four. Each measure was entered separately into the model in order to see how each of them influenced the children’s initial NARA scores (the intercept) and the rate at which these developed from grade four to grade eight (the slope). These are estimated parameters. The

---

7 * p < .05; ** p < .01
influence of a time-invariant covariate on the intercept represents its mean influence on initial reading comprehension scores and the influence of a time-invariant covariate on the slope represents its mean influence on the rate of growth.

The plot between the Ice2 learners’ scores on the NARA test and the reading fluency test, the PPVT, and Orðskilningur, gave Cooks $< |1|$ within each age cohort, revealing no statistical outliers. This means that excluding any participant from the analysis would not have altered results.

Reading fluency entered into the NARA model positively and significantly influenced the intercept ($i$ on fluency $= 0.689; p < .05$), but not the slope ($s$ on fluency $= 0.084; p > .05$). This means that fluent readers outperformed the slower readers initially on the NARA test, but the reading comprehension gap between the fluent and the less fluent fourth graders remained the same from grade four to grade eight. Thus, fluent Ice2 readers in fourth grade continued to outperform the less fluent ones throughout the time of the study on the NARA test, as reading fluency did not influence the rate of growth (see Table 14).

The Ice2 learners’ scores on the PPVT vocabulary measurement positively and significantly influenced the intercept ($i$ on PPVT $= 0.549; p < .05$) and also the slope ($s$ on PPVT $= 0.147; p < .05$). This indicates that children with higher PPVT vocabulary scores in fourth grade scored higher on the NARA initially and showed faster progress over the time of the study than children who had lower PPVT scores in grade four.

The Ice2 learners’ scores on Orðskilningur entered into the NARA model were negatively related to the intercept, ($i$ on Osk $= -0.182; p < .05$), but were positively and significantly linked to the slope ($s$ on Osk1 $= 0.333; p = .00$). The negative initial relationship between the NARA test and Orðskilningur was surprising, as reading comprehension and vocabulary tend to be highly related (see Table 10. section 4.1.4, these two measurements were only significantly related for the older group). One possible explanation for this result is that the reliability of Orðskilningur at the lowest grade level was extremely low among the Ice2 participants, (see section 3.5.1.2). We see, on the other hand, positive and significant influence of Orðskilningur on the slope. This indicates that systematic variation in the Orðskilningur vocabulary scores that correlates with the reading comprehension test emerges in the older grades, which is consistent with correlation figures presented in Table 10 (section 4.1.4).
In summary, the more fluent Ice2 learners’ outperformed the less fluent ones on the NARA test in grade four, and remained ahead throughout the period of the study. Moreover, those Ice2 learners with high scores on the PPVT vocabulary test in fourth grade outperformed their peers with lower vocabulary scorers on the NARA test and also demonstrated a faster reading comprehension growth. High initial scores on Orðskilningur were also linked to a faster rate of growth in reading comprehension, but not to a higher initial performance.

4.3.2 The influence of initial reading fluency and vocabulary scores on initial L.N. scores and their rate of growth

The plot between the participants’ scores on the L.N. reading comprehension test, and reading fluency and vocabulary skills (as measured by PPVT, and Orðskilningur) gave Cooks < |1| within each age cohort of the separate groups of Ice2 and Ice1 participants, revealing no statistical outliers. This means that excluding any of the participants would not have altered the results.

As stated in section 4.1.2.2, the Ice2 and Ice1 groups demonstrated an identical growth pattern on the L.N. reading comprehension test and could, therefore, be analysed in one growth model, representing development from grade four to grade seven. However, due to a negative covariance between the intercept and the slope emerging in the models with the combined groups of Ice2 and Ice1, a fit model was difficult to obtain. For that reason, stepwise analysis, in three steps, was not applicable which would have made it possible to enter a dummy variable for Ice2 learners into the L.N. model in Step 2 with each above measurement as a covariate at Step 1, and an interaction covariate in Step 3.

Thus, growth model analysis was conducted for the Ice2 and Ice1 groups separately to compare the two groups on the influence of the time invariant covariates, reading fluency and Orðskilningur, on initial L.N. reading comprehension scores and the development of those scores. Furthermore, vocabulary scores in fourth grade on PPVT for the Ice2 group were also entered into the L.N. model.

For the Ice2 learners, reading fluency entered into the L.N. model exerted a positive and significant influence on the intercept \( (i \text{ on fluency } = 1.583; \ p = .00) \) and a negative and significant influence on the slope \( (s \text{ on fluency } = -0.345; \ p = .00) \).
fluency = -0.321; \( p = .00 \)). For the Ice1 group, reading fluency entered into the model also exerted a positive and significant influence on the intercept (\( i \) on fluency = 0.649; \( p = .00 \)), but not a significant influence on the slope (\( s \) on fluency = 0.012; \( p > .05 \)).

The inclusion of Ice2 group’s scores on Orðskilningur in fourth grade in the L.N. model exerted a negative and significant influence on the intercept (\( i \) on Osk = -0.164; \( p < .05 \)) but a positive and significant influence on the slope (\( s \) on osk = 0.249; \( p = .00 \)). The initial negative relationship between Orðskilningur and L.N. was unexpected (see also Table 10, section 4.1.4, significant relations between the two tests emerged in grade four, but not in grades five and six for the younger group, and for the older group the relations were significant in grades six, seven, and eight) but, as noted above, the reliability of Orðskilningur in grades four and five was low (see section 3.5.1.2). On the other hand, the positive and significant influence of Orðskilningur on the slope indicates that systematic variation in the vocabulary scores on Orðskilningur that correlates with the reading comprehension test emerges in the older grades.

For the Ice1 group, the inclusion of scores on Orðskilningur in the L.N. model did not exert a significant influence on the intercept (\( i \) on Osk = 0.261; \( p > .05 \)) but exerted a positive and significant influence on the slope (\( s \) on Osk = 0.143; \( p = .00 \)).

Finally, as PPVT was only administered to the Ice2 participants, PPVT scores in fourth grade were entered into the L.N. model for the Ice2 group. Its influence on the intercept was not significant (\( i \) on PPVT = 0.102; \( p > .05 \)), but there was a positive and significant influence on the slope (\( s \) on PPVT1 = 0.159; \( p = .00 \)).

To summarize, for the Ice2 group, PPVT scores positively influenced both initial scores on the NARA test and the rate of growth, whereas these vocabulary scores only influenced the rate of growth on the L.N. measurement (see Table 14). This means that the initial performance gap between higher and lower PPVT scorers on the NARA test increased during the time of the study. On the other hand, Ice2 children who scored higher on the PPVT test did not outperform the lower PPVT scorers initially on the L.N. test but increased these reading comprehension scores at a faster rate from grade four to grade seven, a gap emerging and increasing during the research period.
The children’s initial scores on Orðskilningur positively influenced the rate of growth on L.N. scores, both for the Ice2 group and the Ice1 group. The same applied to the Ice2 group’s scores on the NARA. This indicates that those Ice2 and Ice1 learners with higher initial vocabulary scores made more progress in reading comprehension than their peers with lower initial vocabulary scores and the gap between them increased over time.

Reading fluency positively influenced the intercept in both reading comprehension models, which means that fluent Ice2 readers scored higher on both the NARA and the L.N. test initially, and the fluent Ice1 readers scored higher on the L.N. test initially. On the other hand, for the Ice2 children, fluent readers increased their L.N. reading comprehension scores at a slower rate from grade four to grade seven, allowing the less fluent Ice2 fourth graders to catch up with their peers. A different picture, however, emerged for the Ice1 group. Reading fluency did not influence their rate of growth, meaning that the fluent Ice1 fourth graders continued to outperform the less fluent ones throughout the middle school years. The same pattern of results emerged in the analyses of the Ice2 group’s scores on the NARA test; the fluent fourth graders continued to outperform the less fluent readers, initial reading fluency not influencing significantly the rate of growth on the NARA test, the gap remained the same.

4.4 Research question 3

Growth model analysis was conducted to obtain information about the extent to which initial level and the growth of vocabulary and reading comprehension measurements, administered at all three time points, were influenced by background variables, including age of arrival, first language proficiency, first language literacy level, amount of first language instruction, origins of the first language (European vs. non-European), for the Ice2 group, and maternal education for both groups. These background variables were entered into the base models for vocabulary and reading comprehension measurements as presented in section 4.1, under research question 1.

4.4.1 The influence of background variables on the intercept and slope in the vocabulary models

Background variables were entered into the vocabulary models as time-invariant covariates. Stepwise analysis was not applicable with the data obtained from the vocabulary measurements, as a fit model was difficult to obtain (see section 4.1.1). Therefore, each background variable was entered
separately into the models. This approach informed about the extent to which each time-invariant covariate influenced the children’s initial vocabulary scores (the intercept) and the rate at which these developed from grade four to grade eight (the slope).

The Ice2 group’s age of arrival (AoAr) was first entered into the base model for the vocabulary test PPVT. There was a significant negative influence on the intercept \( (i \text{ on } \text{ AoAr} = -0.335; \ p < .05) \), whereas a significant positive influence on the slope \( (s \text{ on } \text{ AoAr} = 0.324; \ p = .00) \). Age of arrival entered into the base model for Orðskilningur exerted no significant influence on the intercept \( (i \text{ on } \text{ AoAr} = -0.154; \ p > .05) \), but a significant positive influence on the slope \( (s \text{ on } \text{ AoAr} = 0.193; \ p = .00) \). This means that children who were older when they arrived in Iceland made faster improvement in solving these two vocabulary tests during the course of the study than did children who arrived here at a younger age.

When the Ice2 group’s first language proficiency, first language literacy level, and amount of first language instruction were each entered into the base models for the PPVT and Orðskilningur vocabulary measurements an acceptable fit emerged, but none of these variables had a significant influence on the intercept or the slope. Centering produced the same results, i.e. when the intercept was moved from one grade to another significant influence did not emerge in any grade for any of these background variables. However, it is important to note that first language proficiency and literacy were based on parents’ estimates, which may have influenced the accuracy of these measures.

To explore the influence of first language instruction as reported by parents in more details, a dummy variable for those who had not received any first language instruction at Testing 1 was entered into the base model for both vocabulary measurements, PPVT and Orðskilningur. In this way it was possible to compare this group with those who had received first language instruction, irrespective of average number of hours per week. For both tests, the difference between these two groups was not significant, neither on the intercept, nor on the slope. Centering revealed that this applied to all age cohorts, meaning that when the intercept was moved from one grade to another, a significant difference between the two groups did not emerge in any grade.

When the time invariant covariate, mothers’ educational level was entered into the base models for PPVT and Orðskilningur for the Ice2 group, the influence was not significant on the intercept or the slope. Centering gave same
results. On the other hand, when this same analysis was conducted on the Ice1 group’s scores on Orðskilningur, mothers’ educational level produced a significant and positive influence on the intercept ($i$ on meduc = 1.130; $p < .05$). This means that for one unit positive change in mothers’ educational level the students’ intercept (scores in grade four) increased by 2.38 scores (1.130/s.e. 0.47). Mothers’ educational level, however, did not influence the rate of growth ($s$ on meduc = -0.071; $p > .05$).

The performance of Ice2 children with European and non-European first languages were then compared by entering a dummy variable for the difference between the two groups into the base models for the two vocabulary tests. On the PPVT test, the difference between the two groups was not significant on the two growth factors, the intercept and the slope. Centering demonstrated that this applied to all age cohorts, except for the seventh graders ($i$ on European = 0.482; $p < .05$), and eighth graders ($i$ on European = 0.570; $p < .05$) where the difference was significant in favour of children with a European first language.

The same kind of analysis for their scores on Orðskilningur revealed no significant effect on the intercept or on the slope. Centering was used to find out if the difference between the two language groups was significant on the intercept for any age cohort. Significant differences were detected between the two language groups, again in favour of the European first language group, on the intercept in grade five ($i$ on European = 0.831; $p < .05$), grade six ($i$ on European = 1.111; $p < .05$), grade seven ($i$ on European = 1.391; $p < .05$), and in grade eight ($i$ on European = 1.671; $p < .05$).

In sum, older arriving Ice2 learners increased their vocabulary scores at a faster rate than younger arriving Ice2 peers on both vocabulary tests, the PPVT and Orðskilningur, from grade four to grade eight. The difference between the European and non-European Ice2 learners was not significant either on the intercept or on the slope on PPVT or on the vocabulary test Orðskilningur. On the other hand, centering revealed that the European group scored significantly higher on Orðskilningur in grades five, six, seven, and eight, and on PPVT in grades seven and eight.

For the Ice1 group, mothers’ educational level influenced initial scores (in grade four) on the vocabulary test Orðskilningur but not the growth. Thus, those Ice1 children with higher educated mothers started higher initially and remained at a higher level compared to those with the lower educated mothers. On the other hand, mothers’ educational level did not influence either initial vocabulary scores of the Ice2 group or the development of those scores.
The level of first language proficiency and literacy, as estimated by parents did not influence vocabulary scores of the Ice2 participants or the development of those scores. The same applied to first language instruction as reported by parents at the start of the study.

4.4.2 The influence of background variables on the intercept and the slope in the reading comprehension models

As for the vocabulary test, background variables were entered as time-invariant covariates into the base models for the reading comprehension tests. As explained before (see section 4.2), stepwise analysis in three steps (covariate in Step 1, a dummy variable for the difference between the Ice2 and the Ice1 learners in Step 2, and an interaction effect in Step 3) was not applicable with the data obtained from the two reading comprehension measurements, and therefore each background variable was entered separately into the models. This approach informed about the extent to which each background variable influenced the children’s initial reading comprehension scores (the intercept) and the rate at which these developed from grade four to grade eight (the slope) on the NARA, and from grade four to grade seven on the L.N. test (see section 4.1.2.2).

Background variables were first entered into the NARA model for the Ice2 group, but this test was not administered to the comparison group of Ice1 learners (see section 3.5.4.1.1). When the Ice2 learners’ background variables (mothers’ educational level, first language proficiency, literacy level, and the amount of first language instruction) were entered into the NARA model an acceptable fit emerged, but none of these variables had a significant influence on the intercept or the slope. Centering produced the same results, i.e. when the intercept was moved from one grade to another significant influence did not emerge in any grade for any of these background variables. However, it is important to note that first language proficiency and literacy were based on parents’ estimates only, which may have influenced the accuracy of these measures.

To explore the influence of first language instruction as reported by parents in more detail, a dummy variable for those who had not received any first language instruction was entered into the base model for the NARA test. In this way it was possible to compare this group with those who had received first language instruction, irrespective of average number of hours per week. The difference between these two groups was not
significant, neither on the intercept nor the slope. Centering revealed that this applied to all age cohorts, meaning that when the intercept was moved from one grade to another, significant difference between the two groups did not emerge in any grade.

The European and non-European first language groups were then compared by entering a dummy variable for the difference between the two groups into the NARA model. The difference between the two language groups was significant on the intercept, i.e. in grade four, in favour of the European group (\(i\) on European = 3.450; \(p < .05\)) but not on the slope (\(s\) on European = -0.633; \(p > .05\)). Centering demonstrated significant differences between the two language groups in grade five (\(i\) on European = 2.844; \(p < .05\)); grade six (\(i\) on European = 2.208; \(p = .00\)); grade seven (\(i\) on European = 1.571; \(p < .05\)); however, in grade eight the difference between the two language groups was not significant (\(i\) on European = 0.935; \(p > .05\)).

When age of arrival was entered into the NARA model, negative influence on the intercept was not significant (\(i\) on AoAr = -0.052; \(p > .05\)) but significant positive influence emerged on the slope (\(s\) on AoAr = 0.312; \(p = .00\)).

Background variables were then entered into the base models on the L.N. reading comprehension test, for the Ice2 group and the Ice1 group (see section 4.1.2.2). When the Ice2 group’s age of arrival was entered into the L.N. model, there was no significant influence on the intercept (\(i\) on AoAr = -0.056; \(p > .05\)), but a significant positive influence on the slope (\(s\) on AoAr = 0.187; \(p = .00\)). This means that Ice2 learners who were older when they arrived in Iceland, made a faster improvement on L.N. test during the course of the study than did children who arrived here at a younger age.

When the Ice2 group’s first language proficiency, first language literacy level, and amount of first language instruction were entered into the L.N. model an acceptable fit emerged in each case, but none of these variables had a significant influence on the intercept or the slope. Centering produced the same results, i.e. when the intercept was moved from one grade to another significant influence did not emerge in any grade for any of these background variables. However, as before, it is important to note that first language proficiency and literacy were only based on parents’ estimates.

To analyse the influence of first language instruction in more detail, a dummy variable for those who had not received any first language instruction at Testing 1 was entered into the L.N. model, to compare this
group with those who had received first language instruction, irrespective of average number of hours per week. The difference between the two groups was not significant on the intercept or on the slope. Centering revealed that this applied to all age cohorts, both when entered as a dummy variable and as a time invariant covariate for average number of hours per week as reported by parents, i.e. when the intercept was moved from one grade to another significant difference between the two groups did not emerge in any grade.

The two first language groups (European and non-European first languages) were then compared by entering a dummy variable for the difference between the two groups into the L.N. model. The difference between the two language groups was not significant on the intercept ($i$ on European $= 0.630; p > .05$) and not on the slope ($s$ on European $= 0.259; p > .05$). Centering demonstrated that this applied to all age cohorts, the fifth, the sixth, and the seventh graders. Mothers’ educational level was also entered into the base models on the L.N. test, for both the Ice2 group and the Ice1 group, and did not influence the intercept and not the slope.

In sum, for the Ice2 group, age of arrival positively and significantly influenced the rate of growth in NARA and L.N. reading comprehension scores, meaning that those children who arrived older in Iceland made a faster improvement in their reading comprehension scores from grade four to grade eight on the NARA test and from four to grade seven on the L.N. measurement than those who arrived younger. This is consistent with the vocabulary growth models for PPVT and Orðskilningur, in which age of arrival positively and significantly influenced the slope in both models. The background variables, first language proficiency, literacy, and instruction influenced neither initial scores nor the growth on NARA and L.N. reading comprehension scores for the Ice2 group, and this also emerged in the vocabulary growth models, for scores on PPVT and Orðskilningur.

For both Ice2 learners and the comparison group of Ice1 learners, mothers’ educational level did not influence their initial L.N. scores or how fast these scores developed over time, the same applied to the NARA test for the Ice2 group.

However, relatedness with the Icelandic language emerged as a positive influencing factor on reading comprehension scores. The Ice2 participants with European first languages scored significantly higher on the NARA test initially, and continued to outperform the non-Europeans on the NARA test from grade four to grade seven. These findings are consistent with those
demonstrating significant differences between the two language groups on vocabulary measures in grades five, six, seven, and eight on Örskilningur, and on the PPVT test in grades seven and eight, always in favour of the European group. The two language groups did, however, not differ significantly in any grade on the L.N. test, or in the rate of growth. The difference between the two groups, always in favour of the European first language group, emerging on the two vocabulary measurements and on the NARA test, is nonetheless important to highlight.

To explore if there was another factor, than the relatedness between their language pairs, giving the European group an advantage compared to the non-European group, the available background variables of the two groups were compared, using independent samples t-tests. Those with European first languages had significantly better first language literacy skills than the non-European participants, as estimated by their parents (t (88) = 3.34; p = .00); and this group also received more first language instruction than the non-European group (t (85) = 2.99; p = .00). Differences on other background variables were not significant between these two language groups. It is important to note, however, that first language literacy level and first language instruction did not influence vocabulary and reading comprehension scores in any grade or in rate of growth (see section 4.3), and the same findings emerged when the two language groups were analysed separately. Comparing the two groups on reading fluency (see section 4.1.2.3) a difference did not emerge between the two language groups, in grade four (t (1) = -1.30; p > .05), or in grade six (t (1) = 0.81; p > .05).

4.5 Research question 4

The Ice2 and the Ice1 children were compared on the writing test ‘What to do with the empty space on the schoolyard?’ administered at Time point 3, to the sixth and the eighth graders (see detailed description of this free vocabulary measurement in section 3.5.3). We explored if the Ice1 learners used a higher number of tier 2 words than the Ice2 peers, and how this was related with their productive language proficiency. We also analysed the extent to which these were predicted by earlier vocabulary skills. The writing measurement thus gave two variables: one for the number of tier 2 words used by the writers and another for holistic scoring based on rubrics.

Words that fell into the category of tier 2 words were counted for each individual (see appendix 7). Ratings for productive language proficiency
were based on a four level Likert scale: limited, adequate, proficient, and excellent (see appendix 6). When writings fell between rubrics, plusses and minuses were used, adding or subtracting 0.125 from the rubric scores. This provided a scale variable that could be used as a dependent variable in ANOVA, correlation, and regression analysis.

Table 15 demonstrates the mean number and the range of tier 2 words in the children’s writings, as well as the distribution of participants in rubric scores, both for the combined and the separate age groups of Ice2 and Ice1 participants. One Ice2 participant did not respond to the requirements of the task, apparently because he did not understand the instructions. This participant had lived in Iceland all his life. Not all participants took this writing test: 35 of the 48 Ice2 sixth graders and 33 of the 37 Ice1 age peers, 29 of the 34 Ice2 eighth graders and 42 of the 48 Ice1 age peers (see section 3.5.3).

Table 15. Holistic scores of Ice2 and Ice1 sixth and eighth graders and the use of tier 2 words within each proficiency level

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage</th>
<th>Number</th>
<th>Percentage</th>
<th>Max number of tier 2 words</th>
<th>Min number of tier 2 words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Younger: Grade six</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited</td>
<td>15</td>
<td>43</td>
<td>36</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Adequate</td>
<td>19</td>
<td>54</td>
<td>46</td>
<td>0.32</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Proficient</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>100</td>
<td>100</td>
<td>0.20</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td><strong>Older: Grade eight</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>14</td>
<td>35</td>
<td>33</td>
<td>0</td>
<td>0.07</td>
</tr>
<tr>
<td>Limited</td>
<td>10</td>
<td>35</td>
<td>52</td>
<td>48</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Adequate</td>
<td>15</td>
<td>20</td>
<td>52</td>
<td>48</td>
<td>0.33</td>
<td>0.60</td>
</tr>
<tr>
<td>Proficient</td>
<td>4</td>
<td>2</td>
<td>14</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>42</td>
<td>100</td>
<td>100</td>
<td>0.52</td>
<td>1.17</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Combined age groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited</td>
<td>25</td>
<td>39</td>
<td>35</td>
<td>35</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Adequate</td>
<td>34</td>
<td>35</td>
<td>53</td>
<td>46</td>
<td>0.32</td>
<td>0.60</td>
</tr>
<tr>
<td>Proficient</td>
<td>4</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>0</td>
<td>6</td>
<td>8</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>65</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>0.34</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 15 demonstrates that the highest number of tier 2 words was 10, used by an Ice1 eighth grader, whereas the Ice2 learner with the highest number used four tier 2 words, and was also an eighth grader. Moreover, the proficient and the excellent letters included a higher number of tier 2 words than the adequate and the limited ones. Table 15 also shows that 14% of the Ice2 eighth graders were proficient writers and none of the younger ones. Moreover, excellent Ice1 writers all belonged to eighth grade, or 14%. However, while only 5% of the older Ice1 participants wrote proficient letters the share was 18% among the younger ones. This suggests that there was a difference between the two age groups of Ice2 learners, but hardly between the two age groups of the Ice1 learners. On the other hand, Table 15 demonstrates that for the combined age groups of the Ice2 and the Ice1 children, 6% of the Ice2 learners were proficient or excellent writers, whereas the share was 19% among the Ice1 writers.

To explore the pattern of these results further, two-way ANOVA was used to compare the Ice2 group with the Ice1 group as well as the two age groups, with grade (four and six) and language group (Ice2 and Ice1) as fixed factors. Main effect of grade was not significant \( (F(1, 138) = 2.04; p > .05) \), whereas main effect of language group was significant \( (F(1, 138) = 6.98; p < .05) \). An interaction effect, language group x grade, was included in the analysis to demonstrate if the difference between the sixth and the eighth graders was the same for the Ice2 as for the Ice1 group. However, interaction effect was not significant \( (F(1, 138) = 0.30; p > .05) \). These results mean that the Ice1 participants surpassed the Ice2 participants in the use of tier 2 words, and within both groups the eighth graders did not outperform the sixth graders.

Two-way ANOVA also demonstrated that the Ice2 learners had lower rubric scores than the Ice1 group \( (F(1, 135) = 6.48; p = .00) \), but the eighth graders did not outperform the sixth graders \( (F(1, 135) = 2.05; p > .05) \), neither within the Ice2 group nor the Ice1 group, as interaction effect, language group x grade, was not significant \( (F(1, 135) = 0.02; p > .05) \).

Thus, the older Ice2 and Ice1 learners neither used a higher number of tier 2 words nor were they more proficient writers than the younger Ice2 and Ice1 learners. The Ice1 group, on the other hand, surpassed the Ice2 group both in the use of tier 2 words and in rubric scores.

Two-way ANOVA was used to compare writers of different proficiency levels in their use of tier 2 words. When both rubric scores (limited, adequate, proficient, and excellent) and language group (Ice2 and Ice1) were used as
fixed factors the difference between the Ice2 and the Ice1 groups was not significant \((F(1) = 0.45; p > .05)\). The two groups were thus analysed separately. For the Ice2 group, main effect of rubric scores was significant \((F(2) = 46.82; p = .00)\). However, \textit{Bonferroni} tests revealed significant difference in the use of tier 2 words between the proficient and adequate writers \((p = .00)\), but not significant between the adequate and limited writers \((p > .05)\). For the Ice1 group, main effect of rubric scores was significant \((F(2) = 49.23; p = .00)\). \textit{Bonferroni} demonstrated that the difference between the excellent and proficient writers was significant \((p = .00)\), and between proficient and adequate writers \((p = .00)\), but not significant between adequate and limited writers \((p > .05)\). Thus the difference between the different proficiency levels was the same for both the Ice2 group as for the Ice1 group, significant between the three highest levels, but not between the two lowest levels.

As demonstrated in Table 16 the mean number of tier 2 words was very low for each group, and the distribution gave \textit{Skewness} and \textit{Kurtosis} \(>|1|\) \((\textit{Shapiro Wilk}; p < .05)\) in each case, and could for that reason not be entered into regression analysis. Therefore, we could not seek an answer to the research question on the extent to which vocabulary measures in grade four and grade six predicted the use of tier 2 words in grade six and eight. However correlation analysis provided information about the relation between those variables. The correlation between vocabulary measures and the use of tier 2 words and rubric scores are demonstrated in Table 16:

\textbf{Table 16. Correlations (Pearson R) between vocabulary measurements, in grade four and six, and the use of tier 2 words and rubric scores in grade six and eight}

<table>
<thead>
<tr>
<th></th>
<th>PPVT (Ice2)</th>
<th>Orðskilningur</th>
<th>Orðalykill</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ice2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade six</td>
<td>Mean no. of tier 2 words</td>
<td>.19</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>Mean rubric</td>
<td>.34*</td>
<td>-.09</td>
</tr>
<tr>
<td><strong>Ice1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade six</td>
<td>Mean no. of tier 2 words</td>
<td>.13</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>Mean rubric</td>
<td>.08</td>
<td>.36*</td>
</tr>
<tr>
<td><strong>Ice2</strong></td>
<td>Mean no. of tier 2 words</td>
<td>.45*</td>
<td>.38*</td>
</tr>
<tr>
<td>Grade eight</td>
<td>Mean rubric</td>
<td>.34</td>
<td>.39*</td>
</tr>
<tr>
<td><strong>Ice1</strong></td>
<td>Mean no. of tier 2 words</td>
<td>.40**</td>
<td>.23</td>
</tr>
<tr>
<td>Grade eight</td>
<td>Mean rubric</td>
<td>.32*</td>
<td>.25</td>
</tr>
</tbody>
</table>

\(\ast p < .05; \ast\ast p < .01\)
Table 16 demonstrates that for the younger group of the Ice2 children, only scores on Orðalykill in grade four correlated significantly with the use of tier 2 words in grade six, notably, in both cases the scores/numbers are very low. On the other hand, and importantly, for the older group of the Ice2 children, all vocabulary measurements in grade six correlated with the use of tier 2 words in grade eight. For the Ice1 group, no vocabulary scores in grade four correlated significantly with the use of tier 2 words in grade six, whereas only Orðskilningur in grade six correlated with the use of tier 2 words in grade eight.

Table 16 reveals that only one vocabulary test administered in grade four and six correlated with rubric scores in grades six and eight for both the Ice2 and the Ice1 learners. For the younger Ice2 group, only PPVT scores in grade four correlated with rubric scores in grade six, and for the older Ice2 group only vocabulary scores on Orðskilningur correlated with rubric scores in grade eight. For the younger Ice1 group, Orðalykill in grade four correlated significantly with rubric scores in grade six, and scores on Orðskilningur in grade six correlated significantly with rubric scores in eighth grade.

Regression analysis was then conducted with rubric scores as a dependent variable and the vocabulary measurement that correlated with these scores as an independent variable for each age cohort of the Ice2 and the Ice1 learners. This approach provided information about the extent to which vocabulary measures in grades four and six predicted rubric scores in grades six and. The results are presented in Table 17.

Table 17. Regression analysis on rubric scores for Ice2 and Ice1 participants

<table>
<thead>
<tr>
<th>Younger Ice2 Step</th>
<th>predictor</th>
<th>$\Delta R^2$</th>
<th>$B$</th>
<th>$Se_B$</th>
<th>$F$ change</th>
<th>Older Ice2 Step</th>
<th>predictor</th>
<th>$\Delta R^2$</th>
<th>$B$</th>
<th>$Se_B$</th>
<th>$F$ change</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT grade 4</td>
<td>.09</td>
<td>0.01*</td>
<td>0.00</td>
<td>4.46*</td>
<td>4.46*</td>
<td>Orðalykill grade 4</td>
<td>.10</td>
<td>0.04*</td>
<td>0.02</td>
<td>4.71*</td>
<td>4.71*</td>
</tr>
<tr>
<td>Orðalykill grade 4</td>
<td>.10</td>
<td>0.04</td>
<td>5.07*</td>
<td>4.43*</td>
<td>4.43*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17 demonstrates that vocabulary scores explained from 7.7% (older Ice1 group) to 11.9% (older Ice2 group) of the distribution of rubric scores, thus the least of those scores for the older Ice1 group and the most
for the older Ice2 group. On the other hand, for the younger Ice2 and Ice1 groups, vocabulary scores explained more of the variance in rubric scores for the Ice1 group than the Ice2 group. For the younger Ice2 group, PPVT scores in grade four explained 9% of the variance in rubric scores in grade six, and for the younger Ice1 group, vocabulary scores on Orðalykill explained 10.4% of the variance in rubric scores in grade six.

Taken together, for the younger Ice2 group, scores on Orðalykill in grade four correlated highly with the use of tier 2 words in grade six, whereas for the younger Ice1 group, no vocabulary measurements in grade four correlated with the use of tier 2 words in grade six. Nonetheless, vocabulary scores explained a little more of the variance in rubric scores for the younger Ice1 group than the younger Ice2 group.

All vocabulary scores in sixth grade correlated highly with the use of tier 2 words in eighth grade for the Ice2 group, and for this group vocabulary scores explained the most of the distribution in rubric scores. For the Ice1 age peers, scores on Orðskilningur in grade six correlated with the use of tier 2 words in grade eight, whereas these scores explained the least of the variance in rubric scores for this group compared to the younger Ice1 group and the two Ice2 age groups.
5 Discussion

The central aim of the research was to explore the development of vocabulary and reading comprehension among Ice2 learners, and the predictive power of vocabulary for their reading comprehension growth. Moreover, we sought to examine the extent to which theoretically important linguistic and contextual factors would influence the emergence of these fundamental skills for Ice2 learners. A comparison group of Ice1 age peers was included in the study throughout the period, spanning from grade four through grade eight. This is the first Icelandic study of its kind and extends previous work in this area with L2 learners of English (Roessingh, 2008), Dutch (Droop & Verhoeven, 2003), and Norwegian (Lervåg & Aukrust, 2010). In line with these studies the results of the current research demonstrated a widening vocabulary gap between Ice1 and Ice2 learners. On the other hand, the Ice2 group followed the path of their Ice1 age peers in reading comprehension development, although lagging constantly behind.

In fact, the results of the present study also converge with numerous findings on speakers of other languages that have demonstrated that vocabulary is a strong influencing factor on reading comprehension development (Blachowicz, et al., 2006; Davis, 1944, 1972; Snow, et al., 1994), as Icelandic word skills measured in grade four positively influenced the rate of growth in reading comprehension for both the Ice2 and the Ice1 children throughout the middle school years. In this chapter I discuss the findings based on the research questions. Limitations and implications for further studies are also presented.

5.1.1 Is there an initial and developmental difference between Ice2 and Ice1 learners in vocabulary and reading comprehension skills?

We hypothesized that the Ice2 learners would lag behind their Ice1 peers on vocabulary measures already in the first year of middle school and that the gap would widen during the time of the study (see section 2.11.1). This proposition was based on studies with L2 learners of English (Roessingh, 2008), Dutch (Droop & Verhoeven, 2003), and Norwegian languages (Lervåg & Aukrust, 2010) and on two prior studies with Ice2 learners (Thordardottir & Juliusdottir, 2012; S. Ólafsdóttir & Ragnarsdóttir, 2010).

The data provided support for our hypothesis: the Ice2 learners started in grade four with lower Icelandic word skills and they lagged farther and
farther behind all the way up to grade eight. Growth model analysis, based on an accelerated design, revealed that on the vocabulary measurements, administered at the three time points to the younger (grade four to grade six) and to the older group (grade six to grade eight), the Ice2 learners demonstrated a development that was not stable across time.

On the multiple-choice vocabulary test, PPVT, tapping receptive word skills, the Ice2 learners demonstrated a slow improvement from grade five through grade eight. The results extend my MA cross-sectional study with Ice2 children, in grades one, two, three, and four, of which in total 61 participated in the current study, which revealed no difference between the second, third, and fourth graders on the same PPVT test (S. Ólafsdóttir & Ragnarsdóttir, 2010). However, ceiling effects were detected in the current study in grade eight, when the scores were not normally distributed, which indicates that the vocabulary test was not capturing individual differences in word skills among the oldest Ice2 participants. Nonetheless, these two studies demonstrate that Ice2 children make slow improvement on this Icelandic PPVT vocabulary test from grade two and up to grade seven. It is important to note that this vocabulary measurement was pre-tested with Ice1 children from age three to ten and norm referenced for age four to nine, and includes concrete lexical items suitable for children at this age. For that reason, the test does not tap learners’ knowledge of abstract words and idioms associated with educational demands from grade four and onward (see section 2.1.1.1).

The receptive vocabulary measurement Orðaskilningur, on the other hand, includes lexical items and idioms that can be classified as tier 2 words. An understanding of words of this kind is important for academic literacy (Beck, et al., 2002, p. 16). On this multiple-choice test the Ice2 group started below the Ice1 comparison group in grade four and increased these scores very slowly, specifically between grades four and six. Moreover, the poor inter-item reliability that emerged on this test in grades four and five for the Ice2 learners indicates that many of these children were only randomly choosing the correct answers in these two grades. The Ice1 group, on the other hand, made steady progress from one grade to the next, and therefore the vocabulary gap between the two groups widened from grade four to grade eight.

Both the Ice2 and the Ice1 children demonstrated an uneven increase between grades on the vocabulary test Orðalykill. This was the most demanding vocabulary measurement, requiring participants to describe in
their own words the meaning of a target word, challenging receptive vocabulary, semantic awareness, and productive language skills. Just as Órðskilningur, Órðalykill consists of tier 2 words, the lexical items that emerge predominantly in written texts of all kinds and increasingly throughout the middle school years. Already in fourth grade, the Ice2 group had lower mean scores than the Ice1 peers, and in each grade the Ice2 learners’ scores remained clustered within the lowest part of the continuum, while no such pattern emerged for the Ice1 learners. In other words, the Ice2 participants never managed to become fully fledged on this test. The highest Ice2 scorer in each grade was below the average of the Ice1 peers, and the difference between the Ice1 and the Ice2 groups widened throughout the middle school years.

Thus, on both vocabulary measurements that were administered to the Ice2 and the Ice1 participants at all time points, Órðskilningur and Órðalykill, the former group started below the comparison group and the gap widened over the study period. These findings are even more worrying when taking into account the finding that those Ice2 fourth graders with higher vocabulary scores on PPVT and Órðskilningur increased them at a slower rate than the lower initial vocabulary scorers. There seems to be a barrier there, hindering Ice2 learners in their Icelandic word acquisition. On the other hand, this did not apply to the Ice1 comparison group, on Órðskilningur, as no significant difference emerged in the rate of growth between the different initial vocabulary scorers, all developing their word skills at the same rate. This fact is also worth considering as those Ice1 children who started with low vocabulary scores in grade four remained on a depressed trajectory.

The finding that a gap in Icelandic word knowledge of these two groups at the start of the study can be considered as unremarkable, as the Ice2 children arrived in Iceland at various ages. However, the fact that the difference between the two groups became larger each year is a matter of serious concern and indicates that the school system is not providing Ice2 learners with fundamental skills in the Icelandic language for academic development in Icelandic speaking schools.

Our findings extend previous studies with second and third grade learners in Norway (Lervåg & Aukrust, 2010), and third to fourth graders in The Netherlands (Droop & Verhoven, 2003), revealing a widening vocabulary gap between L1 and L2 children already in the first few years of schooling. Our results also add to, and are in line with, those obtained in
other research, in which L2 and L1 children have been followed into the middle school years. Verhoeven and Vermeer (2006) demonstrated an increased vocabulary difference between L1 and L2 third to sixth graders in the Netherlands, and Lesaux and colleagues (2010) between L1 and L2 learners in the US from grade four to five. Furthermore, Roessingh and colleagues (2008; 2009; 2003; 2005), who followed L2 learners of English in Canada throughout the school years and up to grade 12, have demonstrated that the vocabulary gap between the two groups increased from one grade to another. On the other hand, after an increased focus on explicit vocabulary instruction in the Dutch school system, the pattern has changed from a widening to a narrowing vocabulary gap (Verhoeven & van Leeuwe, 2011b) between Dutch L1 and L2 children from first to the sixth grade. Our results thus extend all these findings, as we demonstrated vocabulary development over five years, up to grade eight, and, unlike Roessingh and colleagues, we compared the shape and the rate of vocabulary growth between Ice2 and Ice1 children, and individual trajectories.

We proposed that the gap between the Ice2 children and the Ice1 comparison group would also widen in reading comprehension, based on findings with L2 children of English, (Roessingh & Kover, 2003), Dutch (Droop & Verhoeven, 2003), and Norwegian (Lervåg & Aukrust, 2010). Droop and Verhoeven, and Lervåg and Aukrust tracked development of primary school children, and Roessingh and Kover compared achievement levels between grades, from the first to the twelfth. As mentioned before, in our study reading comprehension development was tracked from grade four and up to grade eight, and the growth rate was investigated as the dependent variable. The Ice2 learners in our study followed the same pattern of development as the Ice1 age peers from grade four through grade eight on the L.N. reading comprehension test. The improvement of both groups was slow from grade four through grade seven, in particular between grades five and six. Moreover, between the seventh and the eighth grade, mean scores of both groups dropped dramatically. The gap between the two groups remained the same, as both groups demonstrated identical rate of growth on the test from grade four to grade seven and the drop in mean scores was the same between the seventh and the eighth grade. This developmental pattern is discouraging, as the Ice2 children did not manage to accelerate their growth when the Ice1 age peers slowed down. The growth curve is in line with that emerging in a study with English L2 and L1 learners in the US.
from grade five to grade seven (Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, et al., 2011), in which both the L2 and the L1 children demonstrated a slowing rate of growth in reading comprehension, and in particular by seventh grade. Thus the reading comprehension gap between the L1 and the L2 children remained the same, as was the case for our Ice2 and Ice1 participants. Our study extends further than that of Mancilla-Martinez and colleagues, as we demonstrated development over five years, up to grade eight, revealing that the gap between the Ice1 and the Ice2 groups remained the same during the entire study period. Notably, in both studies multiple-choice reading comprehension tests were used, and studies suggest that L2 children perform significantly better on reading comprehension tests of that kind than on others with open-ended comprehension questions (Burgoyne, et al., 2009; Spooner, et al., 2004).

On the reading comprehension measurement NARA, only administered to the Ice2 group, in which open-ended questions are answered aloud, the increase of mean scores was also unstable throughout the study period. It is important to take into account, that this test was translated from a Norwegian test, standardized for children from age six to about ten to eleven (see section 3.5.4.1.1). As the child has to answer questions in the NARA test orally, the test taps into the learner’s proficiency in reading aloud, and expressing his answers in speech, in addition to reading comprehension skills, which may explain slow improvement of the Ice2 children. However, on this test, individual variation in scores decreased from grade four to grade eight, indicating ceiling effects, i.e. that this test did not fully capture reading comprehension skills of the best performers.

Both the NARA and the L.N. reading comprehension test challenge different reading strategies, as both include literal and inferential comprehension questions (see section 2.9). The inferential questions are more linked to academic thinking, a challenge children face in the different subject areas, when they add new knowledge to prior knowledge through the act of reading. However, on the L.N. test, participants are also challenged with evaluative questions, requiring them to collect information provided here and there in the text, read between lines, and draw conclusions. Inferential and evaluative reading strategy is more associated with academic proficiency. Moreover, while the NARA test includes short stories of increasing difficulty, the L.N. measurement includes stories and expository texts, the latter, like inferential and evaluative questions, is more challenging and tightly connected with academic requirements. It is indeed
by fourth grade that increasing demands are placed on children’s ability to extract knowledge from expository texts (see section 2.9), and to use more advanced reading strategies, when the reading material of the various content areas becomes more complicated and decontextualized. While our findings suggest that Ice2 and Ice1 middle school learners, in general, need more instructional support to boost their reading comprehension skills, the slow improvement on this test is of most concern for those with the poorest reading comprehension skills, who did not manage to reach the strong ones even when they slowed down.

Moreover, we found that on the NARA test, for the Ice2 group, and on the L.N. measurement for both the Ice2 and the Ice1 learners, those with the highest initial reading comprehension scores increased these at a slower rate than those who started with the lowest scores. Thus, on both reading comprehension tests there was a tendency to mediocrity. The fact that individual differences narrowed over time and that the growth curve slowed down, at least as regards to the L.N. test in which no ceiling effects were detected, indicates that there was a barrier there that even the best initial Ice2 and Ice1 performers could not exceed.

For our Ice2 participants, the developmental pattern in Icelandic vocabulary and reading comprehension skills should be considered with the limitation in mind that two of the Icelandic vocabulary measurements, PPVT and Orðalykill were pre-tested with Ice1 learners only. The third vocabulary measurement, Orðskilningur and the reading comprehension measurement L.N., are based on the combined results of Ice1 and Ice2 participants in national tests in Iceland. The same does, however, not apply to the NARA reading comprehension test, which was translated for the purpose of this study. The need for specific measurements for Ice2 learners is particularly evident from the fact that all three vocabulary measurements did not appear to fully capture Icelandic word skills of this group, PPVT in grade eight, Orðskilningur in grades four and five, and Orðalykill in all grades. Moreover, on the NARA test the variation of scores diminished from grade four to grade eight, which indicates ceiling effects. In fact no Icelandic vocabulary and reading comprehension measurements that have been pre-tested and norm referenced with Ice2 learners separately exist.

5.1.2 To what extent do reading fluency and Icelandic word skills of Ice2 and Ice1 learners predict their
reading comprehension skills and its rate of growth?

Based on the simple view of reading (Hoover & Gough, 1990) we proposed that reading fluency would influence reading comprehension in fourth grade, whereas these technical reading skills would lose their impact throughout the study time and vocabulary take over with an ever increasing influence (Catts, et al., 2005) (see section 2.6.1.1). As regards to the influence of reading fluency our proposition applied mainly to our Ice2 participants. While fluent Ice2 and Ice1 readers in fourth grade had higher reading comprehension scores (NARA and L.N.) in the same grade than the less fluent peers, the influence of reading fluency on the rate of growth was not the same for the Ice2 and Ice1 children on the L.N. reading comprehension measurement. For the Ice2 group, those with higher initial reading fluency scores, having outperformed the slow readers in reading comprehension in grade four, increased these reading comprehension scores at a slower rate than the slower initial readers. This means, that the Ice2 children who were still struggling with decoding Icelandic words in grade four managed to reach reading fluency sufficiently during the time of the study as they made a faster improvement in reading comprehension than the fluent Ice2 fourth graders.

For the Ice1 children, on the other hand, fluent readers in fourth grade outperformed the slow readers in same grade on the L.N. test, whereas the fast and slow readers developed their L.N. reading comprehension scores at the same rate from grade four to grade seven, as the slower initial Ice1 readers were not able to catch up with the fluent ones over these years in reading comprehension. While the orthography of the Icelandic language is rather transparent, it is however more complex than that of, for example, the Finnish language (see section 2.6.2). Young children who learn to read in Finnish have in general acquired reading fluency by the end of the first school year (Lund, et al., 2005; Seymour, 2005; Práinsdóttir, et al., 2006). Verhoeven and van Leeuwe (2011b), however, found correlation between decoding skills and reading comprehension of both L2 and L1 Dutch learners from grade one through grade six. Birgisdóttir (2011) demonstrated in her studies with Ice1 children that by the end of grade three the poorest readers still had difficulties in decoding orthographically complex words. Furthermore, the longitudinal study of Einarsdóttir, Björnsdóttir, and Símonardóttir (2011) revealed that weak performance on phonological awareness tasks (HLJÓM-2) in kindergarten predicted Icelandic language
skills in grades four, seven, and 10. Our findings echo and extend these findings with Dutch L2 and L1 and Ice1 children, as they demonstrate that slow Ice1 readers in grade four, struggle with reading comprehension all the way up to grade eight.

It is indeed by fourth grade that children need to have mastered decoding skills and acquired reading fluency, when they no longer should be learning to read but rather proceeding to the act of reading to learn (Catts, et al., 2005; Chall & Jacobs, 2003). For both the Ice2 and the Ice1 children, the findings of the present research suggest an improvement in reading fluency skills as the sixth graders outperformed the fourth graders. Nonetheless, the Ice2 children demonstrated poorer reading fluency than the Ice1 counterparts in grade four and also in grade six, the difference was not significantly smaller for the older than the younger group (see section 4.1.2.3). Studies suggest that L2 and L1 children acquire decoding skills with comparable ease, this emerging in a number of studies with children in English, Dutch, and Norwegian schools (see August & Shanahan, 2006; Droop & Verhoeven, 2003; Lervåg & Aukrust, 2010; Verhoeven & Vermeer, 2006) (see section 2.6.3).

However, the Ice2 children appeared to master reading fluency eventually, as the less fluent readers caught up with the fluent ones on the L.N. reading comprehension test. Moreover, for the Ice2 children, it is not surprising that they demonstrated poorer reading fluency than their Ice1 peers in fourth and sixth grade, at Testing 1 (see section 4.1.2.3) as these Ice2 children arrived in Iceland at various ages (see section 3.4) and it may take them some time to master decoding and reading fluency skills in the new language.

On the other hand, reading fluency did not demonstrate the same influence on the rate of growth on the NARA test as on the L.N. reading comprehension measurement for the Ice2 children, as reading fluency did not influence the rate of growth on the NARA reading comprehension test. This means that fluent Ice2 fourth graders, outperformed the slow readers in fourth grade, whereas fast and slow readers developed NARA reading comprehension scores at the same rate, the slow readers remained behind the fast readers throughout the study period. The prevailing reading comprehension gap between slow and fast readers on the NARA test, as compared with the diminishing gap between these groups on the L.N. test is worth considering. The NARA test requires the child to read the text aloud and answer comprehension questions orally, while in the L.N. test, the texts
are read and answers given silently. The NARA test thus challenges to a
greater extent technical reading skills than the L.N. test. The findings of the
current study indicate that especially when reading aloud reading fluency
tends to hamper Ice2 children in reading comprehension performance.
Future studies should track the relationship between reading fluency and
reading comprehension of Ice2 children throughout the compulsory school
years, and with both types of tests. The influence of reading fluency on
reading comprehension development of Ice2 children merits more research.

The findings of the current study indicate that especially when reading aloud reading fluency
tends to hamper Ice2 children in reading comprehension performance. The predictive power of Icelandic word skills of Ice2 children on their
reading comprehension development was manifested with our results. The study makes it clear that Icelandic vocabulary plays a major role in
academic development for this population. In fact, for both the Ice2 and the
Ice1 students, those fourth graders with higher vocabulary scores (PPVT for
the Ice2 group and Orðskilningur for both the Ice2 and the Ice1 learners)
demonstrated a faster improvement on both reading comprehension
measurements, for the Ice2 group on NARA from grade four to grade eight
and on L.N. for both the Ice2 and the Ice1 group from grade four to grade
seven. Moreover, for the Ice2 group, higher initial PPVT vocabulary scorers
also scored higher initially on both the NARA test and on the L.N. test. For
the Ice1 group, higher initial vocabulary scorers on Orðskilningur also
scored higher initially on L.N. However, for the Ice2 group, initial
performance on the vocabulary test Orðskilningur did not positively
influence initial reading comprehension scores (NARA and L.N.), but this
was most likely due to the poor reliability of Orðskilningur in grade four, as
stated above.

The findings of a positive influence of vocabulary on both initial
reading comprehension scores and the rate at which these scores developed
over time, demonstrate a widening reading comprehension gap throughout
the middle school years between those Ice2 and Ice1 children with poor and
rich word skills. It appears that Matthew’s effect (Stanovich, 1986) applies
to both groups, Ice2 and Ice1 middle school learners, a developmental
pattern in which the rich get richer and the poor get poorer. This detrimental
tendency can be explained by the fact that children who struggle with texts
including too many unknown words improve their reading comprehension
skills slowly. Ice2 and Ice1 children with larger Icelandic vocabulary, on the
other hand, improve their reading comprehension skills at a faster rate. The
middle school years are particularly critical in this respect, as studies
suggest that children who struggle with reading comprehension during these
years are the most likely to be early school leavers (Barrington & Hendricks, 1989; Snow, et al., 2001, p. 4). For that reason, our results have the significant message: Icelandic vocabulary skills of Ice2 and Ice1 children by the end of primary school can inform about future educational development. These findings are in line with those obtained from the longitudinal study of Einarsdóttir, Björnsdóttir, and Simonardóttir (2011) who found that Icelandic language skills (as measured with TOLD-4, including vocabulary measures) of Ice1 five-year-olds predicted Icelandic language scores on the Icelandic Testing Battery in grades four, seven, and 10.

Our findings, first of all, thus reveal that Ice2 middle school children suffer throughout the middle school years from poor Icelandic vocabulary skills, they make limited progress in acquiring these skills, compared to their Ice1 age peers, and it is their vocabulary deficiency in the Icelandic language, but less their technical reading skills, which results in slow improvement in reading comprehension. The current research extends prior studies on Icelandic second language learners in which no improvement or a very slow increase emerged in their Icelandic word acquisition between grades (Thordardottir & Juliusdottir, 2012; S. Ólafsdóttir & Ragnarsdóttir, 2010), as we demonstrated the consequences of this unfortunate vocabulary progress on their reading comprehension development. Furthermore, we extend the work of Roessingh and colleagues who followed L2 and L1 learners of English from the first to the twelfth grade, and studied the predictive power of vocabulary for later achievement, as we demonstrated that Icelandic word skills positively drive the rate of growth in reading comprehension throughout the important and critical middle school years.

In fact, we demonstrated the predictive power of Icelandic vocabulary for Ice2 middle school learners on the rate of reading comprehension growth, with vocabulary measurements including concrete vocabulary associated with basic interpersonal communication skills (BICS) (Cummins, 1981) (PPVT) and vocabulary that is more connected with higher order thinking, fundamental for the ability to read for learning (Orðskilningur). The strong influence of both high- and low-frequency vocabulary on the rate of growth in reading comprehension was detected in the comprehension of short stories (NARA) and in the understanding of texts that challenge academic reading strategies connected with academic requirements in the different subject areas, i.e. expository texts (L.N.). Taken together, Icelandic word knowledge of Ice2 children in the initial stages of language
acquisition, as well as higher order word skills, predict the rate of growth in reading comprehension at a lower level and at an advanced academic level, throughout the middle school years.

The question, however, remains whether theoretically important contextual factors influence these two fundamental skills of our Ice2 participants, Icelandic vocabulary and reading comprehension and the rate of growth.

5.1.3 To what extent are Ice2 learners’ vocabulary and reading comprehension level and rate of growth influenced by their age of arrival, first language proficiency, literacy, first language relatedness with the Icelandic language, and maternal education?

We proposed that age of arrival would positively and significantly influence the rate of growth in Icelandic vocabulary and reading comprehension (cf. Roessingh, 2008) (see sections 2.3.5.1 and 2.7.2). Our findings confirmed this, as our results demonstrated that those Ice2 children who were older when they arrived in Iceland made faster improvements than those who arrived at a younger age or were born in Iceland, on the two vocabulary tests, PPVT and Orðskilningur, and on the two reading comprehension tests, NARA and L.N. This finding supports the theory of linguistic interdependence (Cummins, 1979) and extends prior studies demonstrating that older learners, with more first language word skills, learn words in another language easier, and increase their reading comprehension scores more effectively than the younger arriving second language learners (Roessingh & Kover, 2003) (see section 2.3.5.1). We demonstrated positive influence of age of arrival on the rate of vocabulary and reading comprehension growth. The older arriving Ice2 children made faster improvement on these skills during the time of the study than the younger arriving peers. Equipped with larger first language vocabularies than the younger arriving Ice2 children, and having also acquired more background knowledge, and reading experience, the older arriving Ice2 children develop reading comprehension skills in the Icelandic language faster (see section 2.7.2). This means that the younger arriving Ice2 children need support even if they have spent all their lives in Iceland. This is consistent with my cross-sectional MA study (S. Ólafsdóttir & Ragnarsdóttir, 2010) in which Ice2
children in grades one, two, three, and four, who were born in Iceland or had arrived before the age of two were behind six year old Icel children on the Icelandic PPVT vocabulary test, the highest sixth grade Ice2 scorer below the average of the Ice1 age peers. Our findings are, however, not in line with the findings of Thordardottir and Juliusdottir, (2012) with Ice2 learners, who demonstrated that the performance of the later arrivals was poorer than that of the younger arrivals, on a translated PPVT and a vocabulary subtest of TOLD (see section 2.3.5.1). It is important to note that the subjects in their study were very few \( n = 39 \), of age ranging from six to fifteen, and with 35% attrition from Time point 1 to 3. Moreover, the approach in analyzing the data was totally different in the two studies, as in the present study the influence of age of arrival on the rate of growth in vocabulary and reading comprehension was explored, whereas the different age of arrival groups were compared in the other study.

In spite of our finding that older arriving Ice2 middle school learners advanced their Icelandic vocabularies and reading comprehension skills at a faster rate than the younger arriving peers, we did not find an influence of first language proficiency, literacy, and instruction, as reported by parents, on Icelandic vocabulary and reading comprehension skills or the rate at which these developed over the time of the study. First language proficiency and literacy, however, were only based on parents’ estimates, in comparison with age peers living in the host country. Therefore, any assumptions based on these results should be made with caution. However, first language instruction, reported by parents at the start of the study as average hours per week, did not exert an influence on initial scores or the elevation of scores on any vocabulary or reading comprehension measurement. The same findings emerged when those Ice2 participants who had received first language instruction were compared with those who had not acquired any. The relation between first and second language skills is complex, and for the youngest children, spending their time in more than one language, findings suggest rather a competition model between the two languages (Oller, et al., 2007). In regards to first language instruction, when for example bilingual programs are explored, the quality of instruction is fundamental, at least in the second language when the second language acquisition is under study (Reese, et al., 2005) (see section 2.7.2). Nonetheless, as vocabulary and reading comprehension skills of the Ice2 learners were measured in the Icelandic language only, the findings provide partial information about their linguistic and literacy skills, taking into
account the distributed nature of bilinguals’ vocabularies (see section 2.3.2.2). Further studies are thus needed with Ice2 learners that use comparable measurements in Icelandic and their first languages, follow them throughout the compulsory school years, and take account of the nature and the quality of both first and second language instruction.

Moreover, mothers’ educational level did not influence vocabulary and reading comprehension skills of the Ice2 children, which is in line with the findings obtained by Haraldsdóttir (2013) with her Ice2 children in kindergarten. Our results are, however, not consistent with those of Reese, Garnier, Gallimore, and Goldenberg (2000) (see section 2.7.2). Their L2 participants with more educated parents, and even grandparents, had been more read to in the first language and demonstrated an initial and continued advantage in the English literacy. Notably, the educational level of the parents was very low, with average seven years of schooling, ranging from zero to sixteen in total. It can be suggested that no education versus some education makes a tangible difference in this respect, that those with no education are the least likely to read for their children and to possess books in general, and that it is the early and the continued domestic experience with books that is influential in the acquisition of literacy in the second language. In our study the educational level of both groups, Ice2 and Ice1, ranged from not having completed compulsory school to having a higher university degree. Consistent with our findings, Chen and colleagues (2012, p. 2013) found, that maternal education did not influence English vocabulary skills in either Chinese-English or Spanish-English fourth and seventh grade L2 learners. The average maternal educational level of their participants was high school and college. Moreover, our findings give a reason to suggest that quality language input in either language is an influencing factor on vocabulary development in that language only, as maternal education influenced Icelandic vocabulary skills of the Ice1 children, and throughout the study time, those of more educated mothers constantly surpassing the others. However, this background variable did not influence reading comprehension skills of the Ice2 and the Ice1 groups, indicating that more than quality domestic language experience is needed for the complex reading comprehension proficiency to grow, and underscores the responsibility of teachers and the educational policy.

A possible influence of the relatedness between the first languages of Ice2 learners and the Icelandic language was explored by comparing those with European and non-European first languages. The Ice2 participants with
European first languages outperformed those with non-European first languages on PPVT vocabulary scores in grades seven and eight, and on Orðskilningur in all grades, except in grade four (inter-item reliability was not satisfactory in this grade, see section 3.5.1.2). On the NARA test the difference, still in favor of the Europeans, was significant in all grades, but not in grade eight (probably due to ceiling effects, see section 4.1.2.1). On the L.N. reading comprehension measurement, however, a difference between the two groups did not emerge in any grade. Thus, a difference between the two language groups emerged throughout the study period on vocabulary measurements and on the NARA reading comprehension test, and always the European group surpassed the other. The difference between the two groups did, however, not emerge in the rate of growth on any vocabulary or reading comprehension measurement, suggesting a stable gap. In my MA study (S. Ólafsdóttir & Ragnarsdóttir, 2010) a difference between the two language groups was detected only among those Ice2 first to fourth graders with the longest residence in Iceland, as those with European first languages scored higher than the non-European peers on the Icelandic PPVT vocabulary test. Chen and colleagues (2012) found that Spanish-English bilinguals outperformed Chinese-English in grades four and seven on cognate items only on an English PPVT vocabulary test. Cognates between the various European languages on the other hand apply predominantly to high-frequency words that belong to daily activities and second language learners master quite easily (see section 2.1.1.1 and 2.3.4). It is therefore an interesting finding that the difference between the two language groups first emerged on PPVT in grade seven. This vocabulary measurement includes words that represent word skills of Ice1 children up to age nine, and for that reason these words are of rather high frequency, however administered by increasing difficulty. The findings of the current study and of the MA research indicate that the advantage of the European learners is longstanding compared to those with non-European first languages. Importantly, the European languages also share language structures, and it is possible that similarities of that kind help Ice2 learners with European first languages to develop Icelandic language skills. Further research is needed in order to detect what specific similarities between the European languages facilitate the learning of the Icelandic language and in comparison with Ice2 learners with unrelated language pairs. Indeed, the findings of my two studies with Ice2 learners suggest that Ice2 learners with non-European first languages need extra support and surveillance, while
both groups of Ice2 learners need specific Icelandic vocabulary instruction to speed up their Icelandic vocabulary and reading comprehension development.

In spite of the fact that in the present study the Ice2 participants were divided into only two language groups, European and non-European, these children have a large variety of first languages (see section 3.4). Choosing only Ice2 learners with two or just a few different first languages is difficult in Iceland, taking into account that the number of Ice2 children in each school is very small. Furthermore, the choice of including all Ice2 learners, regardless of their first languages, is based on my experience as a teacher of Icelandic as a second language. I have observed that those Ice2 children that suffer the most from slow progress in Icelandic language skills are frequently those who belong to the smallest language groups and whose first language differs the most from Icelandic. Excluding these children from this study was, for me, unthinkable.

5.1.4 Do Ice1 middle school learners use a higher number of tier 2 words than Ice2 peers, and how is this related to the children’s productive language proficiency? To what extent are these predicted by earlier vocabulary skills?

Throughout the middle school years children should be increasing their knowledge of the vocabulary that is associated with higher order thinking, the tier 2 words. These words start to play a major role by fourth grade, when children proceed from ‘learning to read’ to the act of ‘reading to learn’. An understanding of such words is not only fundamental for reading comprehension, but the ability to use words in this category in speech and writing is also critical for the development of academic literacy. Furthermore, using a word in free language use, correctly and accurately, represents the highest level of knowing a word, it is to own the word (see section 2.1.3). As knowing and using tier 2 words is fundamental for academic literacy, it is important to raise the lexical bar (Corson, 1984), i.e. support and encourage learners to use their full range of lexical resources, by providing them with vast opportunities to practice the use of these words. This encouragement goes hand in hand with writing practices. In fact, students’ proficiency in expressing their thoughts, in an organized expository language output, is limited if they do not have these lexical
resources. The two skills, the use of advanced vocabulary and the productive sophisticated language proficiency, are therefore pivotal for literacy development. When the learners have acquired these skills they are not only consumers of academic texts but also producers. These are the ultimate language skills, and closely associated with academic proficiency.

We could confirm our hypotheses: The Ice2 learners in sixth and eighth grade used fewer tier 2 words in their writings than the comparison group of Ice1 age peers. This is consistent with the vocabulary lag that emerged between the two groups on the controlled vocabulary measurements (when vocabulary test items are chosen beforehand, see section 2.2.4) that were administered at all time points. Furthermore, in the proficiency of organizing their letters, convincing the reader, and giving support for their arguments, i.e. in holistic rubric scores, the Ice1 children surpassed the Ice2 children.

On the other hand, the difference in the use of tier 2 words between writers of different proficiency levels, excellent, proficient, adequate, and limited, was the same for the Ice2 children as for the Ice1 comparison group. Consistent with studies of children’s writings in Canada (Roessingh, 2012a), excellent writers used the highest number of tier 2 words, followed by the proficient ones, whereas the adequate and limited letters included the lowest number of lexical items of this kind. However, unlike Roessingh, we found that those with adequate writing skills did not use a higher number of tier 2 words than those with limited writing skills. In other words, becoming a proficient and even more an excellent writer makes the difference in the use of academic words, for Icelandic middle school learners.

In addition to the lag of the Ice2 children on both skills, our finding gives even more cause for concern, that for both the Ice2 and the Ice1 students, the eighth graders did not significantly outperform the sixth graders in the use of tier 2 words and in writing proficiency level. This indicates that for both language groups there is no improvement in these skills between the two grades. Nonetheless, our findings revealed that the difference between the Ice1 and Ice2 children in the use of tier 2 words and in rubric scores was the same in grades four and six, suggesting a constant gap between the groups. In Canada a widening gap has been detected between L1 and L2 learners in their writing proficiency, L1 learners increasing these skills at a faster rate than L2 learners (Roessingh & Elgie, 2009) (see section 2.3.1). Notably, their participants had been followed over eight years, up to grade nine. Future studies, with Ice2 children should track
these skills over the compulsory school years and in comparison with Ice1 learners.

In another Canadian study (Roessingh, 2012a), 28% of third grade learners were rated as proficient or excellent writers. In the present study 6% of the sixth and eighth grade Ice2 participants were either proficient or excellent writers, and the share was 19% among the Ice1 sixth and eighth graders. The Canadian and the Icelandic participants performed the same task with the same instructions, and in both studies same rubrics for holistic scoring were used (see section 3.5.3). Thus, the proportion of proficient and excellent writers among Canadian third graders was higher than that of Ice2 and Ice1 sixth and eighth graders.

Based on the fact that several encounters with words are needed before the learner is likely to retrieve the word for active use, we sought to detect the extent to which the use of tier 2 words in the writings was related to earlier receptive vocabulary skills, as measured with controlled vocabulary tests. Thus we explored how vocabulary measurements in grade four related to the use of tier 2 words in grade six (younger group) and how vocabulary measurements in grade six related to the use of tier 2 words in grade eight (older group).

Importantly, the use of tier 2 words correlated with all vocabulary measurements only for the older group of the Ice2 participants. This group appeared to have stretched their lexical resources to a greater extent than the younger Ice2 group and the two age groups of the Ice1 learners. It can be suggested that the Ice2 eight graders were doing their utmost to complete the task. In other words, these youngsters, who differed in initial vocabulary skills, demonstrated the highest lexical bar (Corson, 1984), i.e. their use of tier 2 words represented the best their initial word skills. The vocabulary measurements that correlated with the use of tier 2 words for this group tapped the breadth of receptive word skills, of high (PPVT) and low-frequency words (Orðskilningur and Orðalykill), productive vocabulary and semantic awareness (Orðalykill). Vocabulary scores on these tests increased in line with a higher number of the use of tier 2 words, the latter, however, only ranging from zero to four.

On the other hand, for the younger Ice2 group, only scores on Orðalykill correlated with the number of tier 2 words, and for the younger Ice1 group, no initial vocabulary measurements correlated with the use of words in this category. For the older group of the Ice1 participants, the use of tier 2 words only correlated with scores on Orðskilningur.
Thus for the younger Ice2 group and both age groups of the Ice1 children, the relationship between the use of tier 2 words and earlier vocabulary measures was not straightforward, but rather conflicting. For those with the smallest initial vocabulary, they may have been using all their lexical skills and therefore not using the advanced vocabulary simply because their word skills were poor. On the other hand, for those with best initial performance on the controlled vocabulary tests, it can be suggested that they were not doing their utmost to fulfill the task. A possible reason is that they had not practiced those skills, that they had not been encouraged and obtained frequent opportunities to use a rich variety of words in their language output.

However, it is important to note in this respect that studies have revealed that free productive vocabulary progresses far more slowly than both receptive vocabulary and also productive word skills when measured with controlled productive vocabulary tests (Laufer, 1998) (see section 2.1.3). It cannot be taken for granted that an increase in word comprehension and on controlled productive vocabulary measurements will lead to a more varied and academic word use in free writing samples, at least, if learners are not encouraged and trained to do so.

In spite of the fact that all earlier vocabulary measurements correlated with the use of tier 2 words for the older Ice2 group, only one initial vocabulary measurement correlated with and predicted rubric scores in grades six and eight, for both age groups of the Ice2 and the Ice1 learners. Thus the relation between earlier vocabulary measures and the holistic, rubric score level, was not simple for all participants. However, and consistent with the use of tier 2 words, for the older Ice2 group initial vocabulary measure (Orðskilningur) explained the most of the variance in rubric scores, with 12% explanation; whereas for the older Ice1 group, initial vocabulary measure (also Orðskilningur) explained the least of the variance in rubric scores, or 7%. This supports our suggestion that the older Ice2 youngsters were putting more effort into completing the writing task than the other groups. These Ice2 participants were significantly older when they arrived in Iceland than the younger Ice2 group (see section 3.4). Some of these children had possibly been trained in comparable writing skills in the country of origin before they moved to Iceland. Another possibility is that older arriving Ice2 learners were performing in the writing task as, in general, in all areas of schoolwork, using their full potential as constantly
faced with the double load, to keep up with their Ice1 peers in all subject areas while they are at the same time learning Icelandic.

However, unstable relations between earlier word skills and later advanced free productive word choice and writing skills indicate that, at least those Ice2 and Ice1 children with larger initial vocabularies, would have been capable of better performance if only they had acquired the opportunity to practice those skills, that they were not using their full potential.

5.1.5 Summary

The purpose of this research was to explore the development of vocabulary and reading comprehension among Ice2 middle school learners, and the predictive power of vocabulary for their reading comprehension growth. Moreover, we sought to examine the extent to which theoretically important linguistic and contextual factors might influence the emergence of these skills. The research builds on prior studies with second language learners of English, Dutch, and Norwegian, revealing that vocabulary deficiency is a major cause for their frequent lag in reading comprehension and academic development. Our findings demonstrated that the Ice2 children had lower scores on vocabulary measurements in grade four and the gap between the two groups widened during the study period. However, while both the Ice2 and the Ice1 children made an unstable improvement in reading comprehension, the gap between the two groups remained the same. Moreover, the findings revealed that Icelandic vocabulary skills in fourth grade positively predicted the rate of growth in reading comprehension, for both the Ice2 and the Ice1 learners. This means that those Ice2 and Ice1 learners with larger initial vocabularies made a faster improvement in reading comprehension than those with smaller initial vocabularies. Reading fluency positively influenced initial reading comprehension scores for both the Ice2 and the Ice1 learners. On the other hand, for the Ice2 children, those fourth graders with poor reading fluency managed to catch up with the fast readers in their reading comprehension scores, on the multiple-choice L.N. reading comprehension test. On same test, the reading comprehension gap between slow and fast Ice1 readers, however, remained the same throughout the time of the study. This also applied to the Ice2 children on the NARA reading comprehension measurement, administered orally, as the slow initial readers did not reach the fast readers over these years. Moreover, those Ice2 learners who arrived older in Iceland increased their
vocabulary and reading comprehension scores at a faster rate than those who arrived at a younger age. A difference between Ice2 learners with European first languages and Ice2 peers with non-European first languages emerged throughout these years in their vocabulary and reading comprehension performance, always in favor of the European group. On the writing test, the Ice2 learners lagged behind the Ice1 group in grade four and six, in the use of academic tier 2 words and in the ability to express their ideas in an organized and advanced way. Our findings strongly indicate payoffs for all Ice2 children of focused and rich vocabulary instruction, and that even those who have the longest residence in the country need support, and in particular Ice2 children whose first language is not European.

6 Conclusion
This first longitudinal study of vocabulary and reading comprehension development of Icelandic compulsory school students whose first language is other than Icelandic has demonstrated that this group of learners makes slower progress in the acquisition of Icelandic vocabulary than their first language age peers, from grade four and up to grade eight. Furthermore, we found that those Ice2 children who started with richer Icelandic word skills made a slower progress than those with poorer initial word skills, meaning that even the best performers remained on a depressed trajectory. These findings are consistent with a previous cross-sectional MA study with Ice2 learners in primary school (S. Ólafsdóttir & Ragnarsdóttir, 2010), and a three-year longitudinal study with Ice2 children in grades one to ten (Thordardottir & Juliusdottir, 2012), both indicating little and no increase between grades in Icelandic word acquisition. The results are also in line with previous studies with second language learners of Norwegian, Dutch, and English (e.g. Lervåg & Aukrust, 2010; Droop & Verhoeven, 2003; Lesaux & colleagues, 2010; Verhoeven & Vermeer, 2006; Roessingh, 2008; 2009; 2003; 2005), that have demonstrated a widening vocabulary gap between L1 and L2 children with each school year. Our research addresses limitations and extends the findings of the above quoted studies, as norm-referenced and standardized Icelandic vocabulary measurements were used and a comparison group of Ice1 children was tested at all time points. Data was collected from grade four and all the way up to grade eight,
presenting development over five years, and the shape and the rate of vocabulary growth was explored.

Moreover, our results indicate that the literacy instruction the Icelandic school environment is providing Ice2 children, is not doing more for their reading comprehension skills than helping them to continue along the same pathway throughout the middle school years as they started in grade four, behind their Ice1 peers. Furthermore there seemed to be a barrier there that even the best initial performers could not exceed, represented by an unstable progress between grades and a trend towards the mean, and this applied to both the Ice2 and the Ice1 children. Our findings indicate that Ice2 and Ice1 middle school learners need more effective instructional support to speed up their reading comprehension development, a key to good academic progress. However, the slowing growth curve is of most concern for those children with poor reading comprehension skills, as they could not accelerate their growth when their peers slowed down.

Furthermore, we extended the two prior studies with Ice2 learners by demonstrating the consequences of a poor progress in Icelandic vocabulary acquisition of this population, as our findings make it clear that Icelandic word skills are a strong predicting factor on their development in Icelandic reading comprehension. Indeed, a widening reading comprehension gap emerged throughout the middle school years between high and low vocabulary scorers in fourth grade, for both the Ice2 and the Ice1 children. Moreover, we found that the variance in basic Icelandic vocabulary of Ice2 children, and also in higher order academic word skills, measured in the first year of middle school, predicted the rate of growth in reading comprehension along the continuum, from story reading and all the way to an advanced text comprehension associated with academic work.

Our findings underline the importance that educational policy makers in Iceland take into account positive outcomes that have resulted from instructional, vocabulary focused programs in The Netherlands, Canada, and in the US (Appel & Vermeer, 1997; Roessingh, 2008; Snow & Lawrence, 2011) in bridging the reading comprehension gap between learners. Indeed, three and a half decades have passed since vocabulary deficit was first highlighted as the main reason for academic failure among disadvantaged students (cf. Becker, 1977). In the Icelandic curriculum (Aðalnámskrá grunnskóla: almennar hluti 2011 og greinasvið 2013), vocabulary as fundamental skills for reading comprehension development (Blachowicz, et al., 2006; Davis, 1944, 1972; Snow, et al., 1994) is still
overlooked. Also decades after researchers in neighboring countries have demonstrated with their findings that vocabulary deficiency of second language learners is a key variable in explaining their frequent lag in academic achievement compared to first language age peers (Appel & Vermeer, 1997; Roessingh & Kover, 2002; Verhoeven, 1990), vocabulary is ignored in this respect in the Icelandic curriculum.

Taking into account the trivial role vocabulary plays in the Icelandic school system (see section 2.4), it can be suggested that the reading comprehension gap between the Ice1 and Ice2 groups will not remain the same from grade eight and on, but rather widen, in line with increasing demands on vocabulary skills (see section 2.6.1.1).

The responsibility of Icelandic schools is great, as for our Ice2 participants, with both their parents with other first languages than Icelandic, all their Icelandic quality language input occurs within the school settings. It is indeed quality language input, including low-frequency academic words, that is requisite for the development of academic literacy (Roessingh & Douglas, 2013). These fundamental skills are not acquired through daily language experience; these pivotal skills for academic development can be obtained from quality conversations and, importantly, from reading.

An important aspect of this research was that a measurement of the learners’ use of academic vocabulary in a writing task was included, which has been used in very few literacy studies. Taking into account the importance of quality language skills for academic procedure it was disappointing that the eighth grade Ice2 and Ice1 participants did not use a higher number of academic tier 2 words in their writings than the sixth graders. Nonetheless, the number of words in this category used by the writers went up in line with higher productive language proficiency, based on their ability to express their ideas in a convincing and an organized way. The Ice2 children lagged behind the Ice1 peers in both skills, in the use of tier 2 words and in writing skills. Importantly, applying to both the Ice2 and the Ice1 children, the relation between earlier vocabulary skills and productive language proficiency was not simple, indicating that, at least those with the largest vocabularies were not stretching their full range of lexical resources. The middle school years are crucial in this respect, as this is the age at which learners absorb increasingly new knowledge, match with prior knowledge, constantly encountering and learning new concepts and lexical items important for higher order thinking. For intellectual
development sophisticated receptive vocabulary is far from being sufficient. Using academic words freely in speech and writing is the optimal word knowledge, a key to being able to use the language to explore ideas and thoughts. Equipped with these skills students are not only consumers of academic texts but also producers.

The finding that maternal education influenced word skills of the Ice1 learners is in line with studies indicating that quality conversations at home occur in general to a greater extent in families of educated parents (Hart & Risley, 1995). However, this background variable did not have any impact on vocabulary and reading comprehension skills of the Ice2 group, and the latter also applied to the Ice1 children. This indicates that quality language input in one language influences vocabulary acquisition in that language, and that more than domestic quality language experience is needed for the complex reading comprehension proficiency to grow. These findings underscore the responsibility of the school system, in particular as regards Icelandic vocabulary acquisition of Ice2 children and Ice1 peers with small vocabularies, and reading comprehension development of all Ice2 and Ice1 learners.

It was a remarkable finding that reading fluency skills predicted the growth in reading comprehension differently for the Ice2 group than for the Ice1 children. While slow initial Ice2 readers managed to catch up with the fast Ice2 readers on the multiple choice reading comprehension test the reading comprehension gap between the different initial reading fluency performers remained the same for the Ice1 group. Our finding, regarding Ice1 children, extend those of Birgisdóttir (2010) showing that poor decoders were still struggling in grade three, as our results indicate that poor Ice1 decoders struggle with reading comprehension throughout the middle school years. The same applied to the Ice2 children on an oral reading comprehension test, as those who demonstrated poor reading fluency skills in grade four had lower scores in same grade and continued to lag behind the fast readers up to grade eight. These results give a reason to suggest that it is only when Ice2 learners are required to read aloud and answer comprehension questions orally that reading fluency continually hampers their reading comprehension performance. These findings merit further investigation with both types of reading comprehension tests.

Our finding that the older arriving Ice2 learners enriched their Icelandic word and reading comprehension skills at a faster rate than the younger arriving peers is in line with the inter-dependence theory (Cummins, 1979),
revealing a positive transfer between the first and the second language for those who start to learn a second language at an older age, when they have acquired more first language skills to build on (see section 2.3.5.1 and 2.7.2). Our finding supports a number of studies that have demonstrated that bilingual children who have a long residence in the host country, and are even second or third generation immigrants, need support from the very start (Roessingh & Elgie, 2009). Thus, our findings do not support the belief that ‘the younger the better’, but on the contrary confirm that ‘the older learn faster’. The two groups are, however, both faced with double the load, to acquire proficiency in Icelandic and develop academic skills across the curriculum.

We could, nonetheless, not demonstrate an influence of the Ice2 children’s first language proficiency or literacy, as estimated by their parents, and even not of first language instruction, on their Icelandic vocabulary and reading comprehension skills. In fact, studies on first language impact on second language acquisition have given conflicting outcomes. Positive influence has predominantly and most clearly become evident in the age at which children start to learn a second language, as an advantage emerging and increasing with age. As regards to first language instruction as a contributor to second language acquisition, findings are also conflicting and indicate strongly that it is important to take into account the quality of instruction, at least in the second language when that is under study (see section 2.7.2) (e.g. Reese, et al., 2005). In future studies it is important to use experimental groups in bilingual and monolingual educational programs, with pre- and post-tests, include measurements in both languages, and track development at least throughout the middle school years, when learners no longer are learning to read and have started the longstanding process of reading to learn. In spite of our inability to demonstrate an influence of first language skills and instruction on Icelandic second language acquisition it is important to underscore what a treasure it is to have good skills in the mother tongue, irrespective of its impact on second language learning. As Icelandic second language acquisition was the focus of the present study, only Icelandic measurements were included and for that reason the research gives only partial information about vocabulary and reading comprehension skills of the participants.

The fact that a difference emerged between Ice2 learners with European and non-European first languages in each grade on vocabulary and reading comprehension tests, always in favor of the Europeans, indicates that Ice2
children with non-European first languages need even more assistance to boost their Icelandic literacy acquisition throughout the compulsory school years. The results of the present study extend prior cross-sectional study with Ice2 primary school children (S. Ólafsdóttir & Ragnarsdóttir, 2010) in which a difference in vocabulary skills emerged between the two first language groups only among those with the longest residence in Iceland, the European group surpassing the other. Further research with Ice2 learners is needed to detect if there are specific aspects in the Icelandic language that give European first language groups a prevailing advantage in the acquisition of literacy, over those Ice2 children whose first language is not European.

Taken together, the implications of the findings are clear for Ice2 learners: solid Icelandic word skills are, first and foremost, fundamental for their academic progress in Icelandic speaking schools. If the education of Ice2 children, from the earliest grades to the latest, is not organized around bolstering their Icelandic vocabulary skills, their reading comprehension growth in the Icelandic language will be hampered. The results should provide the impetus for launching an effective campaign in advancing Ice2 children, and Ice1 as well, in their Icelandic word acquisition.

First, it is about time that a focus on Icelandic vocabulary as fundamental for reading comprehension is inscribed in the Icelandic curriculum. Good work of individual teachers in some schools is far from being sufficient. What is needed is a comprehensive school wide approach, that vocabulary instruction is the core focus of the entire educational society, for the benefit of all children in all grades and across the curriculum. An emphasis on vocabulary acquisition should be introduced to a large part of teaching material intended for both Ice2 and Ice1 children, providing multifarious vocabulary activities, in all grades, and in all subject area. Ice2 and Ice1 children need to acquire the specific vocabulary that is particular in each area of schoolwork, but also they need the vocabulary that plays a key role for the ability to use language for learning, the academic language.

Second, the education of teachers should provide them with skills that will lead to rich vocabulary and language activities in their future classes. Vocabulary evolves with increased experience, and knowledge. Teachers should foster the learners’ language skills through rich discussions and, importantly, writing opportunities to express their ideas and thoughts related to new learning, and with the use of a rich variety of words. Opportunities to
develop academic word skills can go hand in hand with modern technology, opening up new resources, challenging reading comprehension and writing skills more than ever.

Third, measurements are needed that have been pre-tested and standardized for Ice2 children specifically, in order to track their Icelandic word acquisition and reading comprehension development from the very start and throughout the whole educational system. There is also a need for research in which effective approaches in the teaching of important Icelandic lexicon are experimented with pre- and post-tests and these then used as a foundation for teaching practices and material. Moreover, further studies on the frequency of Icelandic words are needed, and indeed with the aim to detect the Icelandic words fundamental for academic success in Icelandic schools. In light of the complexity of the Icelandic language, due to the inflection of verbs, nouns, adjectives, numbers, and pronouns, studies are needed of effective ways in meeting the needs of Ice2 learners in this respect for both receptive and productive use.

The findings of the research make it clear that Icelandic vocabulary has to become a ‘hot’ literacy research topic (cf. Cassidy & Cassidy, 2004/2005; Cassidy & Ortlieb, 2013) in the next few years. Only then can we expect to be able to offer Icelandic citizens with other first languages than Icelandic the same opportunities in Iceland as those with Icelandic as their first language.
Appendix 1 – Letter of approval for school principals

Reykjavík, 27. febrúar 2012.

Ágæti skólastjóri

Ég undirrituð er doktornemi á Menntavísindasviði Háskóla Íslands og er að hefja rannsókn sem ber heitið próun orðaförða og lesskilnings íslenskra grunnskólabarna sem hafa annad móðurmál en íslensku. Leiðbeinendur minir eru dr. Freyja Birgisdóttir lektor og dr. Hrafnhildur Ragnarsdóttir professor á Menntavísindasviði Háskóla Íslands.


Ég óska góðislegslega eftir leyfi þínu til að prófa isl2 börn í skólanum þínun í 4. og 6. bekk, fyrrum þátttakendur í MA rannsókninni ásamt óðrum isl2 börnum.

Meginmarkmið rannsóknarinnar er að auka þekkingu á þróun íslensks orðaförða og lesskilnings hjá börnum sem eiga annad móðurmál en íslensku og bera saman við börn sem hafa íslensku sem móðurmál. Síðar upplýsingar er síðan hægt að nota til að gera kennslaðferðir markvissari og skilvirkari.


Foreldrar þeirra barna sem valin verða til þátttöku fá bréf með upplýsingum um rannsókna og þurfa þeir að skila inn skriflegu samþykki fyrir þátttöku barna sinna. Einnig biðjum við þá um að svara bakgrunnspurningum sem nauðsynlegar eru við úrinnslu niðurstöðna.

Ég vona að verkefníð þyki áhugavert og vel verði tekið í þessa málaðeitan.
Sigríður Ólafsdóttir
doktornesi á Menntavisindasviði Háskóla Íslands
sími 8495411
sio24@hi.is
Meðfylgjandi er leyfisbréf, sem sent verður til foreldra nemendanna á viðkomandi tungumálum.
Appendix 2 – Letter of approval for parents of Ice2 participants

Reykjavík, February 2012

Dear parents/guardians,

I am a doctoral student in the Educational Department at the University of Iceland. I am conducting a research project, *The development of vocabulary and reading comprehension among Icelandic second language learners*. I request permission for your child to participate.

The main aim of the research is to advance our understanding of the process of Icelandic vocabulary development among children with Icelandic as a second language and to explore its possible influence on reading comprehension development, and compare this with children with Icelandic as a first language. Such knowledge can contribute to more effective teaching methods and materials.

I will visit the children three times during school hours, one at a time and together in a group, which will take about 40 minutes each time. The tasks are in the form of word plays and reading texts and I will try to make them as pleasurable for the children as possible. The information obtained from this research will be kept strictly confidential and anonymous. The children will be tested in spring 2012, 2013, and 2014. Your child will only participate if you grant your permission. To increase the value of the research it is important that the most children will take part in this doctoral research.

In the space below, please indicate whether you *do* or *do not* want your child to participate in this project, and if so, kindly answer the enclosed questions; and return this to your child’s teacher **within 2 days**.

Sincerely,

Sigríður Ólafsdóttir

I do give permission ...

I do not give permission...

...for my child ____________________________

name of child

to participate in the doctoral research project of Sigríður Ólafsdóttir
Parent/guardian’s signature:

__________________________________________

If you give permission for your child to participate please answer the following questions:

• What education has the mother’s child completed?
  Not compulsory  Compulsory  College  University  Post
  graduate University
• Does your child know how to read in his/her first language?
  If yes, how well can he/she read in his/her first language
  compared to his native speaking age peers?
  Very well  Well  Neither well nor poorly  Rather poorly
  Poorly
• How well does your child speak his/her first language
  compared to his native speaking age peers?
  Very well  Well  Neither well nor poorly  Rather poorly
  Poorly
• Does your child get any lessons in his/her first
  language?
  If yes, how many lessons a week on average?
Appendix 3 – Letter of approval for parents of prior MA Ice2 participants

Reykjavík, February 2012

Dear parents/guardians,

Your child participated in my MA research in 2009, *Icelandic Vocabulary of children with Icelandic as a second language*. I now ask you to permit your child to participate in my doctoral research project, *The development of vocabulary and reading comprehension among Icelandic second language learners*, which is an extension of the prior research.

The research is conducted from The Educational Department at the University of Iceland. The main aim of the research is to advance our understanding of the process of Icelandic vocabulary development among children with Icelandic as a second language and to explore its possible influence on reading comprehension development, and compare this with children with Icelandic as a first language. Such knowledge can contribute to more effective teaching methods and materials.

I will visit the children three times during school hours, one at a time and together in a group, which will take about 40 minutes each time. The tasks are in the form of word plays and reading texts and I will try to make them as pleasurable for the children as possible. The information obtained from this research will be kept strictly confidential and anonymous. The children will be tested in spring 2012, 2013 and 2014. Your child will only participate if you grant your permission. To increase the value of the research it is important that the most children who participated 2009 will take part in this doctoral research.

In the space below, please indicate whether you *do or do not* want your child to participate in this project, and if so, kindly answer the enclosed questions; and return this to your child’s teacher **within 2 days**.

Sincerely,

Sigríður Ólafsdóttir

I do give permission ...  
I do not give permission...

...for my child ________________________________.

name of child
to participate in the doctoral research project of Sigríður Ólafsdóttir

Parent/guardian’s signature:

__________________________________________

If you give permission for your child to participate please answer the following questions:

• What education has the mother’s child completed?
  Not compulsory  Compulsory  College  University  Post graduate University

• Does your child know how to read in his/her first language?
  If yes, how well can he/she read in his/her first language, compared to his native speaking age peers?
  Very well  Well  Neither well nor poorly  Rather poorly  Poorly

• How well does your child speak his/her first language compared to his native speaking age peers?
  Very well  Well  Neither well nor poorly  Rather poorly  Poorly

• Does your child get any lessons in his/her first language?
  If yes, how many lessons a week?

  ______________________
Appendix 4 – Letter of approval for parents of Ice1 participants

Reykjavík, mars 2012
Ágetu foreldrar

Ég er doktornemí á Menntavísindasviði Háskóla Íslands og er að hofja rannsóknina Próún orðaförða og lesskilnings íslenskra grunnskólabarna sem hafa annað möðurmál en íslensku. Leiðbeinendur mínir eru dr. Freyja Birgisdóttir og dr. Hrafnhildur Ragnarsdóttir.

Meginmarkmið rannsóknarinnar er að að auka þekkingu á þróun íslensks orðaförða, hvernig orðaförði tengist lesskilningi og bera saman börn sem hafa íslensku sem möðurmál og börn sem hafa íslensku sem annað tungumál. Slikar upplýsingar er síðan hægt aðnota til að gera kennslaðferðir markvissari og skilvirktir.


Skilyrði fyrir þáttnökku barns er að forráðamaður veiti upplýst samþykki. Til að auka gildi rannsóknarinnar er mjög mikilvægt að sem flæst börn taki þátt í rannsókninni.

Til að einfalda skipulag, er þess farið á leit að meðfylgjandi eyðublaði verði skilad til kennara barnsins eða skólastjóra innan 2 daga, hvort sem veitt er samþykki fyrir þátttökun barnsins í rannsókninni eða ekki.

Með bestu kevðjum og fyrirfram þókk
Sigríður Ólafsdóttir

Vinsamlega merkið í viðeigandi reit:
Ég undirrituð/aður veiti samþykki mitt...
Ég undirrituð/aður veiti ekki samþykki mitt...
...fyrir því að barnið mitt ___________________________.
nafn barns

Taki þátt í doktorsránsókn Sigriðar Ólafsdóttur

Undirskrift foreldris/ fórðamanns:

________________________________________________________________________

Ef þið gefið samþykki fyrir ránnsókninni vinsamlegast svarið
eftirfarandi spurningum:

Hvaða menntun hefur módir barnsins lokið?
Ekki grunnskólaprófi grunnskólaprófi framhaldsskólaprófi
háskólaprófi framhaldsnámi í háskóla
Appendix 5 – The writing prompt: The empty Space behind the Schoolyard.

Name:
Grade:
School:

WRITING TASK
All students are asked to do some writing. Read the information below and think about how you will do the writing assignment.

TO THINK ABOUT BEFORE WRITING
Imagine that there is a large undeveloped space in your schoolyard. Every student in the school has been asked for ideas about what to put there. A committee of teachers and parents will choose the best suggestion.

IN YOUR WRITING
Write a proposal for the committee to read. Describe what you would put in the space. Then, convince the committee that your idea is the best way to use the space.

DIRECTIONS FOR WRITING
You will have up to 45 minutes to plan and write, so budget your time carefully.
Use the PLANNING page to plan your writing. You many brainstorm, web, draw, or list ideas. Think of details that will be interesting and entertaining.
Use the WRITING pages to write a first draft. You may show changes and corrections on your first draft. Do not write a ‘good copy’. If you need more space to write, use the back of the writing pages. Please number your extra pages.
Your work will be evaluated on WHAT you write and HOW WELL you write. Remember to:
• CONSIDER your audience
• PRESENT your ideas in prose
• ORGANIZE your writing as required by the task
• FOCUS on the purpose of your writing
Appendix 6 - Rubric for holistic scoring of the writing prompt (Edmonton Public Schools, 2008)

<table>
<thead>
<tr>
<th>PERFORMANCE CRITERIA-GRID FORMAT</th>
<th>EXCELLENT</th>
<th>PROFICIENT</th>
<th>SATISFACTORY</th>
<th>LIMITED</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Era</strong></td>
<td>The writer addresses the task and offers clear, logical, and original insights.</td>
<td>The writer addresses the task, offers some logical insights, but may not have a clear focus.</td>
<td>The writer addresses the task, but has difficulty in providing logical insights.</td>
<td>The writer does not address the task, or addresses it in a confusing or illogical manner.</td>
</tr>
<tr>
<td><strong>Audience</strong></td>
<td>The writer clearly identifies the audience and tailors the writing to the audience's needs.</td>
<td>The writer identifies the audience, but the tailoring is not always effective.</td>
<td>The writer fails to identify or understand the audience.</td>
<td>The writer does not address the audience or fail to engage the audience.</td>
</tr>
<tr>
<td><strong>Audience engagement</strong></td>
<td>The writer's writing effectively engages the audience and maintains reader interest throughout.</td>
<td>The writer's writing is engaging, but may lose the audience's attention in some parts.</td>
<td>The writer's writing is often off-topic and lacks interest.</td>
<td>The writer's writing fails to engage the audience.</td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td>The writer uses an extensive and varied vocabulary in a meaningful way.</td>
<td>The writer's vocabulary is adequate, but may lack variety.</td>
<td>The writer's vocabulary is limited and repetitive.</td>
<td>The writer's vocabulary is basic and repetitive.</td>
</tr>
<tr>
<td><strong>Usage</strong></td>
<td>The writer's usage is accurate, clear, and logically consistent.</td>
<td>The writer's usage is generally accurate, but may contain occasional errors.</td>
<td>The writer's usage contains frequent errors and confusion.</td>
<td>The writer's usage is incorrect and confusing.</td>
</tr>
<tr>
<td><strong>Mechanics</strong></td>
<td>The writer's mechanics are flawless, and the writing is well-organized and easy to follow.</td>
<td>The writer's mechanics are generally sound, but may contain minor errors.</td>
<td>The writer's mechanics are flawed and the writing is difficult to follow.</td>
<td>The writer's mechanics are extremely flawed and the writing is unorganized.</td>
</tr>
</tbody>
</table>

**NOTES:** For HLAT writing, performances are judged on GRADE AND PERFORMANCE. A student's writing is judged to be no particular grade level.
Appendix 7 - Tier 2 words used by participants

**Frequency number on the MÍM list:**
(Stofnun Árna Magnússonar í íslenskum fræðum, n.d.)
(see section 3.5.3)

<table>
<thead>
<tr>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upplagt - upplagður</td>
<td>7590</td>
</tr>
<tr>
<td>Halli</td>
<td>3316</td>
</tr>
<tr>
<td>Illilega</td>
<td></td>
</tr>
<tr>
<td>Lélegur</td>
<td>2151</td>
</tr>
<tr>
<td>Reitur</td>
<td>4427</td>
</tr>
<tr>
<td>Eins konar</td>
<td></td>
</tr>
<tr>
<td>Fjarlægja</td>
<td>2309</td>
</tr>
<tr>
<td>Ganga upp (= virka)</td>
<td></td>
</tr>
<tr>
<td>Meiðast - meiða</td>
<td>5230</td>
</tr>
<tr>
<td>Hifa</td>
<td>10799</td>
</tr>
<tr>
<td>Eins og sjálft nafnið segir</td>
<td>5717</td>
</tr>
<tr>
<td>Öðruvisi</td>
<td></td>
</tr>
<tr>
<td>Troða</td>
<td>3274</td>
</tr>
<tr>
<td>Aðstaða</td>
<td>1251</td>
</tr>
<tr>
<td>Vinsæll</td>
<td>1653</td>
</tr>
<tr>
<td>Koma með skemmtunina til þeirra</td>
<td></td>
</tr>
<tr>
<td>Plássfrekt</td>
<td></td>
</tr>
<tr>
<td>Skipta út</td>
<td></td>
</tr>
<tr>
<td>Blettur</td>
<td>3913</td>
</tr>
<tr>
<td>Félagsskapur</td>
<td>3210</td>
</tr>
<tr>
<td>Innilokaður</td>
<td></td>
</tr>
<tr>
<td>Hlekkaður - hlekkur</td>
<td>6167</td>
</tr>
<tr>
<td>Frjósa</td>
<td>8566</td>
</tr>
<tr>
<td>Snyrtilegur</td>
<td>5264</td>
</tr>
<tr>
<td>Spjalra</td>
<td>2399</td>
</tr>
<tr>
<td>Endurnýja</td>
<td>2437</td>
</tr>
<tr>
<td>Í fyrsta lagi</td>
<td></td>
</tr>
<tr>
<td>Litrikur</td>
<td>4716</td>
</tr>
<tr>
<td>Liflegur</td>
<td>5046</td>
</tr>
<tr>
<td>Ræktunarstaður</td>
<td></td>
</tr>
<tr>
<td>Rækta</td>
<td>2312</td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Slaka á</td>
<td>3946</td>
</tr>
<tr>
<td>Risaleikstaður</td>
<td></td>
</tr>
<tr>
<td>Brú</td>
<td>2426</td>
</tr>
<tr>
<td>Kíkir – kíkja</td>
<td>2771</td>
</tr>
<tr>
<td>Frumskógur</td>
<td></td>
</tr>
<tr>
<td>Ljósastangir</td>
<td></td>
</tr>
<tr>
<td>Áhugamaður</td>
<td>3891</td>
</tr>
<tr>
<td>Spara</td>
<td>2517</td>
</tr>
<tr>
<td>Betrumbæta</td>
<td></td>
</tr>
<tr>
<td>Tenging</td>
<td>2317</td>
</tr>
<tr>
<td>Hanki</td>
<td></td>
</tr>
<tr>
<td>Taka til umhugsunar - umhugsun</td>
<td>9694</td>
</tr>
<tr>
<td>Náttúrulegur</td>
<td>5201</td>
</tr>
<tr>
<td>Útlit</td>
<td>1611</td>
</tr>
<tr>
<td>Huggulegt</td>
<td>9078</td>
</tr>
<tr>
<td>Útsýnispallur – útsýni</td>
<td>2440</td>
</tr>
<tr>
<td>Slökkviliðssúla – súla</td>
<td>4683</td>
</tr>
<tr>
<td>Marglítaður</td>
<td></td>
</tr>
<tr>
<td>Óveður</td>
<td>7834</td>
</tr>
<tr>
<td>Taka til máls</td>
<td></td>
</tr>
<tr>
<td>Örvæntingafullur</td>
<td>9965</td>
</tr>
<tr>
<td>Huggulega</td>
<td>9078</td>
</tr>
<tr>
<td>Braut</td>
<td>1562</td>
</tr>
<tr>
<td>Skreyta</td>
<td>6178</td>
</tr>
<tr>
<td>Staðsetning</td>
<td>2030</td>
</tr>
<tr>
<td>Handföng</td>
<td></td>
</tr>
<tr>
<td>Virka</td>
<td>2105</td>
</tr>
<tr>
<td>Sóa</td>
<td>9180</td>
</tr>
</tbody>
</table>
Appendix 8 – An excellent writing sample

Kæri skólastjóri

Ég hafð trúða hugmynd um það hvað hægt er að gera við auða svæðið á skólalóðinni.

Mér finnst að það ætti að vera náttúrulegur kastali. Það sem ég er að meina eru tveir steinar eða plasthlutir sem eru hlíð við hlíð. Á einum er stigi upp sem kemur á aðra hæð og þar er bekkur sem fólk getur sest á til að spjalla eða kannski bara til að hvíla sig. Þaðan er rennibraut níður því það er alltaf gaman að renna sér níður rennibraut. Á annarri hæðinni er kíkir bara til gamans fyrir kannski einhverja leiki eins og þykjustunni leiki og þar er líka útsýnispallur. Til að hafa það skemmtilegra væri sniðugt að hafa slökkviliðssúlu níður, þannig að maður verði fljótari að fara níður.

Til þess að tengja þessa tvo steina/plast/grjót hluti þá væri skemmtilegt að hafa brú á milli og þar kemst maður níður stiga með því að vera inni í kastalanum. Það er hægt að hafa líka líttinn bekk þar inni því það getur verið frekar huggulegt. Svo það verði ekki ógæðelega dimmt þá gæti verið gluggi á miðju kastalans. Það er líka sniðugt að hafa svona innistiga, þegar það er kalt eða fólk er að fela sig, þá getur það farið þangað inn.

Um útlitið var ég að pæla að hafa þetta smálitrikt eins og rauðar tröppur og bekkrinnir margliðaðir. Það væri líka fæld ef það væri blómapottur sem krakkarnir gætu ræktað blóm eða plöntur í. Því þetta er uppi í lofti þá er sniðugt að hafa gríðingu í kring.

Mér finnst að þú ættrir að gera þetta því að það væri gaman að hanga í þessu leikteki og væri flott að horfa á ef þetta myndi heppnast vel.
References


Compulsory School Act, no. 91 (2008).


A. Nizegorodcew (Eds.), EUROSLA Yearbook 1 (pp. 7–28). Amsterdam, Netherlands: Benjamins.


Reese, L., Garnier, H., Gallimore, R., & Goldenberg, C. (2000). Longitudinal analysis of the antecedents of emergent Spanish literacy and middle school English reading achievement of


212


Sigurðsson, Á. (1940). Algengustu orðmyndir málsins og stafsetningarkennslan [The most frequent word forms and the spelling instruction]. *Menntamál*, 13, 8–42.


