English...what do we expect?

Exploring the mismatch between year 5 national tests in

English and competence aims at the end of year 4

Sarah Jeanne Grøthe

15.05.2017

Inland University of Applied Sciences, Lillehammer

Master's thesis of Education

Table of Contents

Forward

In October of 2013, I sent an email to the Norwegian Directorate for Education and Training voicing concerns about the difficulty on year 5 English national tests in correlation to my experiences as a teacher of English in the Norwegian primary classroom. I recently reread my email and the response from the Norwegian Directorate for Education and Training and have realized that, four years later, the same concerns have instigated this Master's thesis.

Throughout the writing process, I have struggled with how strongly I should make declarations, knowing that those who invest their careers in making national tests in English, may be offended. However, perhaps the research and ideas presented here will only make national tests in English better, or at least bring more understanding to the correlation (or lack of correlation) between year 5 English national tests, and competence aims in the subject at the end of year 4.

I would like to thank my husband Knut Grøthe, for many active discussions throughout the last year. Thank you to Yngve Nordkvelle, my advisor at Inland University of Applied Sciences, for insights and questions. Thank you to Kari Martens Meyer, a guardian angel who continually supports and believes in me and education. She asks questions which push me to reflect and she also knows when to offer encouragement. Lastly, thank you to colleagues at Sødorp Skole. Thank you to the school's administration team who has been positive and supportive. Thank you to my colleagues who, like teachers in Norway and across the world, continually strive to create a safe and enriching environment that allows for maximum learning.

Abstract

The main objective of this thesis is to investigate and attempt to explain the mismatch between year 5 national tests in English and competence aims for the English subject at the end of year 4 in hope to justify frustrations experienced in the classroom as an English teacher. In order to achieve this objective I lay a foundation consisting of three components. Firstly, I show how the English subject is organized in Norway at the national and community level, specifically in lower primary education. Secondly, I explain what second language acquisition means in a classroom context and factors contributing to learning another language in this setting. Thirdly, I describe some of the components that make up reading assessments.

Next, I embark on an exploratory study which analyzes versions of the year 5 English national tests from 2009-2016 through the method of document content analysis. By analyzing several tests, I observe how content in tests has changed over time and how exercises on tests correlate (or do not correlate) to competence aims being tested. Conclusions of the document content analysis have shown changes in some of the types of exercises, an increase in overall word-count for some of the exercises and an increase in reading comprehension complexity throughout the years. Significant changes were seen between 2012 and 2013.

The fact that recent year 5 national tests in English, which are second language tests for 10-year-olds, focus primarily on reading comprehension when language learning competence aims focus on all four areas of language development can be a reason for the mismatch. Further, it is clear that competence aims are open to interpretation and various interpretations can justify the mismatch between expectations on year 5 tests in English and competence aims for the subject at the end of year 4.

Chapter I: Introduction

Norwegian, mathematics and English as a foreign language are considered core subjects in the Norwegian primary school system. Correspondingly, reading, numeracy and parts of English are nationally assessed for the first time at the beginning of year 5. These national tests, which are meant to assess basic skills and competence aims from the previous four years, intend to provide schools with insight into learners' knowledge, skills and understanding of these subjects. The intent is that results from national tests will give insights which provide necessary information for schools to adapt their organization and teaching to provide for the needs of students by support or extension. This is stated clearly by the Directorate of Education and Training in the analysis of national tests when they say:

> *The goal with national testing is to evaluate and develop students' basic skills in literacy, numeracy and parts of the English subject. The results shall be used as a basis for quality development for all levels in the educational system.*
> (Utdanningsdirektoratet, 2016, translation mine).

Unfortunately, national tests have also been used as a measuring stick for individual schools and communities when media compares schools to each other and uses results from national tests to imply *better* schools. Completing simple searches on most online newspapers in Norway using *national tests* as key words proves this to be a hot topic. Articles range from praising schools, teachers and individual classes which score in the top percentage when it comes to national tests, to discussions of warning and worry about arranging schools from best to worst based on national test results. An academic, research report from the Nordic Institute for Studies of Innovation, Research and Education (NIFU) confirms teachers' and principals' frustrations surrounding publication of national test results and discussions in the media, especially because not all media personnel sit with necessary background knowledge and information about the tests. The report also states that over 80% of primary schools in Norway are so small that results are meaningless when used to compare with other schools (Seland, Vibe & Hovdhaugen, 2013).

Magnus Marsdal's book, *Kunnskapsbløffen* (the title's English translation is *The Knowledge Bluff*) has its main focus on the detriments caused by national tests (2011). He includes

interviews from principals and teachers describing how they prioritize performance outcomes by setting aside other subjects in order to drill and practice for national tests. He compares schools to businesses where results of national tests act as quality assurance of an overall product. The product in this case is individual learners' results on national tests. His book presents only a critical view of national tests, but perhaps some truth is hidden in there as well. The national tests have a potential of becoming steering documents for education instead of the national curriculum.

This is a side point, but perhaps also a starting point which fuels background for this thesis. Since 2008, I have taught English in Norwegian primary school and have consistently had year 5 learners take part in national testing. As intended, the national curriculum with its focus on the five basic skills and competence aims for the English subject have steered instruction and learning in the classroom. However, when it comes to the year 5 national test in parts of English, which assesses the same competence aims, I have seen frustration in learners who find the test overwhelming. At yearly parent meetings when presenting examples of previous English tests, parents are shocked by what is expected of their 10-year-old children. Thus, I am curious. Why is the year 5 national test in English difficult for so many learners? Are learners not prepared enough or does the test expect too much of them? Have the tests gotten more difficult over the years? Are the tests in line with the competence aims? There is a mismatch somewhere. As a teacher, I understand the need for national tests and I know they are here to stay but I also want learners to succeed and feel successful. Too often, I have seen learners' walk away from year 5 national tests in English with a look of defeat and I have overheard comments of their own failure.

In 2013 NIFU presented a detailed report about how national tests in general, and the system surrounding national tests as a whole, are used at the school, community and national levels. The research was based on national tests from 2012 and included surveys and interviews at most levels of the school system, including school owners, principals, teachers and parents. Although there was a lot of positive feedback about the *structure* of national tests, the report did not focus on *content* of any of the national tests (Seland, Vibe & Hovdhaugen, 2013). (Other articles and Masters theses relating to organizational structure of national tests include: Fikke & Helnes, 2012, Olaug-Berger, 2007; Nes, 2013; Isaksen &

Hjelm Solli, 2014). In contrast to focusing on the *structure* surrounding national tests as a whole, I choose to focus on *content* of national tests in English at the year 5 level and how this relates to the competence aims from a teacher's perspective. As a teacher who works directly with learners who take these tests, and the challenges that are involved, I find it vital and perhaps about time that content is addressed.

Therefore, the main aim of this thesis is an attempt to understand and explain the mismatch between English competence aims at the end of year 4 and English national tests at the beginning of year 5. In order to do this, I explain how the English subject in years 1-4 is organized in Norwegian schools at the national and community levels and what is known about second language learning and reading assessment. Then, I analyze year 5 national tests in English from 2009-2016 in an explorative study using the method of document content analysis. My intention is to bring an awareness to professionals who are interested in this topic in order to promote further research and (perhaps even) changes so that there are respectful, understandable correlations between English competence aims ending in year 4 and year 5 national tests in English that teachers, test-makers and politicians (can) find valuable for education.

Organization and viewpoint

After establishing a foundation for how the English subject is organized in Norwegian schools in chapter II, I look at important theoretical information surrounding second language acquisition in chapter III. Here, I explain the differences between first and second language acquisition and important factors in teaching English as a foreign language in beginning primary education. I conclude chapter III with information about assessment, focusing primarily on reading comprehension assessment. Then, in chapter IV I explain the method of document content analysis and the questions which have steered my investigation. In chapter V, I discuss findings and relate them back to information from chapters II and III. Chapter VI concludes with final remarks and ideas for further research.

I view the analysis as a teacher, not as a test-maker, second language acquisition researcher, linguist, politician, parent or administrator. Tone Kvernbekk has worked intently with studies in educational theorizing and part of her book from 2005 focuses on being an *insider* or *outsider* in regards to research and understanding a phenomenon. It is often difficult for

insiders to view a phenomenon as an outsider and vice versa. She uses the well-known Native American proverb "Don't judge a person until you walk two moons in his moccasins" to associate with what it means to be an insider. In essence Kvernbekk says it is difficult to truly understand something unless it is understood from the inside (2005). For example, a person doesn't really know what it is like to be a cancer survivor unless (s)he is or has been a cancer survivor. Similarly, a professional doesn't really know what it's like to teach children year after year in English and experience frustrations related to year 5 national tests in English, unless one has experienced it. As a teacher, I view myself as an insider of the specific activity of teaching and learning English in the Norwegian primary classroom. I have real-life, hands-on experiences in the classroom, every day, year after year with learners and perceive challenges and opportunities that are part of the changing classroom.

As a teacher, being an insider has influenced the topic of this thesis. It has also influenced themes focused on in the theory chapter, choosing to use the method of document content analysis and the view in which I make qualitative conclusions about findings in the content analysis.

It is important to obtain information, research and articles from insiders and outsiders of a phenomenon which is why I find it necessary to write this thesis. This insider teacher view, this asset, is what I bring to the discussion. In the introduction of Marysia Johnson's (2003) book, *Philosophy of Second Language Acquisition,* the author goes to great extent to explain the lack of connection between theorists and practitioners. Perhaps, like her, I am also trying to bridge this gap.

Chapter II: Organization of the English subject today

The Directorate of Education and Training (Utdanningsdirektoratet) under advisement from The Norwegian Ministry of Education and Research (Kunnskapsdepartementet) is responsible at the national level for education in Norway. The Directorate of Education and Training will henceforth be referred to as UDIR since the audience of this paper is mostly Norwegian and UDIR is a familiar abbreviation in the Norwegian educational system. Together these entities set up a framework and guidelines for education in Norway. The latest reform known as Kunnskapsløftet 06 (LK06) is the major steering document for education which includes basic skills, the core curriculum as well as competence aims for each subject from the beginning of primary school until the end of secondary school. This reform is originally from 2006. Some updates have been made since then (UDIR, 2006, 2015).

The competence aims currently fall under one of the following four main subject areas for primary and secondary English education:

- *language learning*
- *oral communication*
- *written communication*
- *culture, society & literature*

At the primary level, specific competence aims serve as reference points, key stages or benchmarks for what should be achieved after year 2, year 4 and year 7 (UDIR, 2015). National tests, course books, as well as what and how teachers teach should all be based on the reform of LK06, its updates and the corresponding competence aims for English. The major change in the English curriculum happened in 2013 when the original subject area of *communication* was split into two subject areas of *oral communication* and *written communication* (UDIR 2006, 2015).

In coordination with LK06, UDIR has created a *Framework for Basic Skills* document which recognizes five basic skills that are necessary for successful educational experiences and a foundation for life in both work and social situations. UDIR sees the basic skills as

fundamental for learning and they are therefore integrated in all subjects throughout the curriculum (2012). These skills are defined as: *oral, reading, writing, digital,* and *numeracy.*

In the *Framework for Basic Skills* document, UDIR goes into further detail about each basic skill mentioned above and includes rubrics which define a progression of levels within each area. The rubrics are general and should be used to relate skill development to subject content (UDIR, 2012). Further, competence aims should be made with these rubrics in mind. It is important to understand both competence aims and basic skills because in addition to creating guidelines for what should be taught and learned, they create structural, organized, evaluation criteria for national tests in Norwegian, numeracy and parts of English.

An example of a rubric from the *reading* portion of *The Framework for Basic Skills* document is included in appendix 1 on page 66. UDIR says that national tests in English focus primarily on the competence aims since English is not one of the five basic skills (2017). However, I find it important to include the reading rubric because in the instructional guide for teachers (2016), UDIR explains how questions on the national test measure specific levels of reading comprehension and these correlate with basic skills of reading.

National tests

National tests started in 2004, in the aftermath of the Programme for International Student Assessment (PISA) exams in 2001 which showed Norwegian learners were under average in reading and numeracy skills. Norwegian authorities made changes in the educational system and LK06 was created as well as national tests in reading, numeracy and part of English (Seland, Vibe & Hovdhaugen, 2013). UDIR is responsible for making and assessing national tests and providing feedback to teachers and schools about results. English national tests are constantly being developed by departments at the University of Bergen (Kunnskapsdepartementet, 2016). It is a requirement for individual exercises to be tried out at least once on a population of learners (UDIR, 2017). Learners currently take all national tests digitally. Since 2012, an instructional guide has been available for teachers of English at the year 5 level. This provides general information about the tests, suggestions about how to prepare learners and follow up with learners afterwards. Results from UDIR show that the average year 5 learner answers between 50 and 60 percent of the questions correctly on English tests (UDIR, 2014).

According to UDIR, the national test in parts of English assesses competence aims in *written communication* (UDIR, 2014). Although it can be argued that competence aims can indeed overlap with each other, the 2016 instructional guide for teachers states that the test is based on testing four of these aims which are included below. (The English competence aims in their entirety for the end of year 4 are included in appendix 2 on page 67.)

- use simple reading (and writing) strategies
- read, understand (and write) English words and expressions related to one's needs and feelings, daily life, leisure time and own interests
- understand the main content of simple texts about familiar topics
- use some common, short words and simple spelling and sentence patterns (UDIR, 2016)

Time allocation

The Norwegian Ministry of Education and Research and UDIR also dictate the amount of time allocated each subject (see appendix 3 on page 68). Generally, students begin learning English in year 1. However, it is up to individual communities and schools to divide the time among different year groups. Time is allocated in hours. Some schools use 45 minute *lessons*. In these schools one *60-minute-hour* is equivalent 1.25 *45-minute-lessons*.

In years 1-7, the following amounts of hours are allocated to the three core subjects: Norwegian—1372 hours, mathematics—888 hours and English—366 hours (UDIR, 2015). UDIR states that in order to meet all aims in the English subject, it is imperative to develop skills across other subject areas as well (2016). When comparing time allocation to Norway's neighboring countries, learners in Denmark complete 450 hours by the end of year 7 and learners in Sweden complete 480 hours. (Ministry for Children Education and Gender Equality, 2016; Skoleverket, 2016).

Norway uses about 100 hours less than its neighboring countries when it comes to the English subject in primary schools. It is interesting to mention how Norway compares to its neighboring countries in this area. It must also be said, however, that more time does not necessarily relate directly to greater learning in a classroom situation. There are other

factors that play a role in the quality of learning of course. Class size and learning environment are two examples (out of several other factors) that affect learning as well.

What does 366 hours of English mean in a normal school setting? Let me give an example from the school where I work. Here, there are 45-minute lessons. The following explains how many 45-minute lessons there are of English on a weekly basis. If there is half of a lesson, this means 22-23 minutes.

Year group	Number of 45-minute-lessons of English per week
1	0.5
2	1
3	1.5
4	2
5	2.5
6	2.5
7	3

How schools delegate and divide English lessons vary. This is part of the autonomy given individual schools throughout Norway. Some schools might decide to have one 45-minute lesson a week during the first three years for example. Another school in the community where I work tried out a system where they had zero lessons of English during year 1 and 2.5 lessons in year 4. Teachers, however, were strongly encouraged to include English in other lessons or at other times of the day during the first year of school, for example during the start of the day or in music lessons.

It is important to mention the 38 flexible hours that are allocated to schools as seen on the chart from UDIR (appendix 3 on page 68). Individual schools allocate 38 flexible hours to subjects they deem necessary. Some schools may decide to use a portion of 38 hours towards the English subject. It is also worth mentioning that time allocated the English subject has increased since the latest reform of LK06. In 2006, English was allocated a total of 328 hours for primary school in comparison to 366 hours it is allocated today (UDIR 2006, 2016).

Education of teachers

Central to the quality of teaching are teachers' deep understanding of what they need to teach and the pedagogical practices that can be used to represent such understanding to students (D.L. Ball 2000 in Wang et. al, 2010, p. 395).

Another part of organization at both national and community levels includes which teachers teach English. Ideally, teachers should be experts in the subjects they teach and have relevant education supporting that expertise. This is a goal of the Norwegian Ministry of Education and Research as presented in the Teacher's Lift (Lærerløft) in the autumn of 2014 (Kunnskapsdepartementet). The Teacher's Lift, with a primary focus on increasing education for teachers, was presented as a way to create better schools and raise the education level in Norway. It explains that in 2017 (now postponed to 2018) teacher education in Norway should be a 5-year program in which all teachers must obtain Master degrees.

That being said, there is a vast amount of teachers who teach English at the primary level who do not have the desired education for the subject. At the time the Teacher's Lift was written, 65% of teachers in years 1-4 and 50% of teachers in years 5-7 did not have the desired amount of education to teach English (Kunnskapsdepartementet, 2014; Utdanningsforbundet, 2016).

There is a debate about if experience can give the same qualification of years of education. There is no guarantee that more education automatically produces *good* teachers. Further, some student teachers have implied their years in the classroom at a university or teacher-college have done little to prepare them for the realities as a teacher in the classroom (Østrem, 2009). This is a debate I am not going to take up here. The important thing to remember, however, is summed up in the quote above. Teachers should be experts in their subjects and have pedagogical and didactical experiences as well to create the best environment to allow for maximum learning. Therefore when it comes to English, teachers should know their learners as well as the subject and approaches to teaching and learning a second language.

Course books

Course books chosen in conjunction with the English subject affects how English is organized at the community or school level. Course books are meant to be based on competent aims and be a resourceful tool for teachers to use in conjunction with particular subjects. In years 1-4, there are several course books that schools or communities can choose from in order to aid in teaching and learning English. Some common course books include *Quest* from Aschehoug, *Stairs* from Cappelen Damm AS, *Explore* from Gyldendal, *Junior Scoop* and *First Choice* from Fagbokforlaget. Course books in years 1-4 often include classroom CDs, textbooks and workbooks for learners, a teacher's manual with pedagogical instructional and activity ideas as well as information about how competence aims and basic skills are met. There are interactive textbooks and website links relating to learner textbooks and workbooks. Some course books include flashcards or differentiated readers that learners can read as well. Throughout this thesis, I bring in examples from the *Stairs* course book since it is the English course book I am most familiar with.

Moving on

It has been established that the Norwegian Ministry of Education and Training and UDIR have overall responsibility for providing the educational framework in Norway. More specifically, the framework creates expectations of what should be taught, learned and assessed in English primary and secondary education by way of basic skills and competence aims. Further, the framework provides information about time allocation. The framework also creates boundaries so that those who work within education such as teachers, principals, test-makers and authors of course books can create situations for teaching, learning and assessment that are in line with the English subject's intentions. This basic understanding of English structure is necessary because, as a teacher, information presented in chapter III has to then conform within the framework of the English subject and its realities in the Norwegian classroom.

Chapter III: Theory

In order to analyze year 5 national tests in English from a teaching perspective, it is necessary to have a theoretical basic understanding related to language acquisition, factors affecting second language acquisition in the daily classroom context, and assessment in order to create a platform for the analysis. I believe if teachers and test makers do not have an understanding of how a second language is learned and factors contributing to this learning in the modern day early primary classroom, teaching and assessing is far less effective and relevant for learning. I begin chapter III by explaining how first language is acquired. Then, I clarify differences between learning a second language in simultaneous bilingual contexts and sequentially in naturalistic and instructional contexts. Looking more closely at instructional contexts, I look at factors that affect teaching and learning in second language acquisition. I conclude chapter III by taking a closer look at reading assessment, specifically focusing on reading comprehension assessment and views presented by Melissa Lee Farrall (2012). This chapter intends to provide the reader with necessary information to create the types of eyes or glasses, if you will, in which to view and understand how I have both analyzed year 5 national tests in English and interpreted findings from the analysis.

The following abbreviations are used during the remainder of the thesis. **SLA** is the common abbreviation for **second language acquisition**. This abbreviation and term is used for people who might be acquiring a third, fourth or fifth language as well, but is an accepted term relating to learning *another* language after the first language (Hummel, 2014). **L1** is an abbreviation for the first language, whereas **L2** is an abbreviation for the second or *other* language.

First language acquisition in general

Knowing a language…. Is a slow, organic process that requires exposure to massive amounts of input (Ellis, 2015, p.16).

There are over 7 billion people in the world using nearly 7 100 languages (Ethnologue, 2016). Language is part of being human and all people with normal cognitive and physical development, who are exposed to language, learn to speak and understanding their L1 without difficulty. Often, it is with little effort that people become experts in their first

language, speaking with perfect pronunciation and using grammar skills correctly. Newborns respond to both voices and languages heard in the womb (Bialystok, 2013; Hummel, 2014). After birth, studies show similar stages in L1 acquisition even though languages differ (Tomasello, 2010; Hummel, 2014; Mellanby & Theobald, 2014) and by the age of five or six, young children understand and speak their L1 with correct syntax, grammar and pronunciation without cognitively reflecting on their own development due to the omission of metalinguistic awareness at a young age. Throughout this learning process, there is an abundance of repetition of words, phrases and correct grammar. Children are enveloped in a context where there is motivation to understand and communicate. In L1 acquisition, oral language is always learned before written language.

Second language acquisition

Approximately two-thirds of children in the world use more than one language (Bhatia & Ritchie, 2013; Mellanby & Theobald, 2014). Even though speaking more than one language is normal and not abnormal in today's globalized society, L2 acquisition can be more difficult than L1 acquisition. The study of SLA had its onsets with Noam Chomsky (1928-) in the 1960s (Ellis, 2015; Hummel, 2014). Chomsky with his cognitive ideas of Universal Grammar and the Sensitive period Hypothesis and Lev Vygotsky (1896-1934) and his ideas of learning through social interaction are two theorists who have shaped ongoing questions and research influencing SLA today both in Norway and elsewhere (Ellis, 2015; Mellanby & Theobald, 2014; Farrall, 2012; Johnson, 2003). Since SLA is a relatively new area of study, knowledge in this field is continually changing and adapting.

Types of SLA

SLA refers to different types of learners. A L2 can be acquired through simultaneous bilingual contexts, or sequentially through naturalistic or instructional contexts. All three types of language learners are in the Norwegian classroom today but it is important to distinguish differences here so there are appropriate expectations for learning in SLA. Let me explain what it means to learn a L2 simultaneously and sequentially in both naturalistic and instructional contexts and then I will explain why these differentiations are important to remember when making national competence aims and national tests.

Simultaneous bilingualism and SLA

When children acquire two languages at the same time, they learn languages simultaneously and are referred to as simultaneous bilinguals (Barron-Hauwaert, 2004; Mellanby & Theobald, 2014). A typical example is if two parents consistently speak a different language to a child from birth and the child grows up acquiring two languages at the same time. In general it is accepted that simultaneous bilingualism occurs if children begin acquiring two languages by the age of three years (Baker, 2000). Simultaneous bilinguals are indeed a part of SLA but learn more than one language in a way that is different to what is focused on in this thesis. These children do not use a great amount of energy in acquiring a L2. Two languages come effortlessly and naturally, just as the L1 comes effortlessly and naturally in L1 acquisition. It is a virtually unconscious process (Ellis, 2011).

Naturalistic contexts and SLA

When an additional language is learned after the age of three years old, children (and adults) learn additional language(s) sequentially. It is a complex process and in regards to SLA, there are two distinctions to make here. Firstly, SLA can occur through naturalistic contexts. This occurs when the L2 is acquired in a setting where language is heard and spoken often and acquiring language is meaningful (Hummel, 2014; Ellis, 2015; Mellanby & Theobald, 2014). An example of learning a L2 in a naturalistic context occurs when people move to another part of the world where a new language is spoken (Hummel, 2014). For example, when people from other countries move to Norway, they learn Norwegian as a L2 through naturalistic contexts. An important distinction between naturalistic contexts and instructional contexts as addressed below is that the L2 is also part of society surrounding learners. In a similar way to acquiring a L1, in a naturalistic SLA environment, learners have a great amount of exposure to the L2 and natural situations which offer continuous repetition and meaning for learning. Children speak the L2 not only during Norwegian 2 classes for example, but also throughout the rest of the school day with teachers and classmates as well as at after school activities, stores or the doctor's office. Learners acquire the L2 through immergence and in this process socialization is one large factor in both learning the language and being motivated to learn the language (Ellis, 2015).

Instructional contexts and SLA

In contrast, another type of sequential SLA, and the area of SLA that correlates with this thesis, is when the L2 is acquired in instructional or classroom contexts. Here, learners are exposed to the L2 for a few hours a week which leaves little room for constant repetition or practice. The process of acquiring a L2 is slower than in a naturalistic setting (Ellis, 2011). It may be difficult for learners to understand the importance of learning the L2 or find motivation or meaning when it is not used outside of the classroom. Having said this, I believe it is easier for learners to obtain a sense of meaning for understanding English in Norway compared to other L2 languages. It is important to note the difference between acquiring English as a L2 in an instructional setting and acquiring for example German, French or Mandarin in an instructional setting. In Norway, learners are exposed to English outside of the classroom. English is heard through music, on television and digital games. English words have crept into the Norwegian language and English is written on t-shirts and other clothing. Children who learn English only in the Norwegian classroom definitely fit into the category of instructional SLA, however with the examples of English in the surrounding environment as listed above, arguably, it is easier for learners to find meaning and/or motivation in learning English compared to other languages in Norway that have less influence on daily life such as German, French or Mandarin. I find this important to mention because perhaps this extra exposure to English contributes to expectations in regards to outcomes of learning English.

It has now been established that the type of SLA focused on in this paper is in instructional contexts. Due to limited amount of L2 exposure in this context, it is vital to not have the same expectations for learners in instructional contexts as learners in naturalistic or simultaneous bilingual contexts. Learners who become bilingual simultaneously or learn English as a L2 in naturalistic contexts have far more exposure to the L2 than learners in instructional contexts. They have a greater vocabulary and are able to focus on meaning rather than the form of language or individual vocabulary words in both written and spoken language (Ellis, 2011).

In fact, the reason I have chosen to differentiate between L2 learners in simultaneous bilingual, sequential naturalistic and sequential instructional contexts is because there are

learners in Norway who take year 5 national tests in English who come from simultaneous bilingual or sequential naturalistic contexts in regards to English language. In addition, learners who attend English-speaking schools such as Oslo International School in Bærum take national tests in English. These types of L2 English learners need to be considered when teaching and learning in the classroom. However, I believe they should not set standards for skills assessed or expectations for reading comprehension ability in year 5 national tests in English. Clearly the level of oral and written understanding and production is far greater for learners who have massive exposure to English.

Factors influencing SLA in instructional contexts

Within the discipline of SLA, there are a plethora of contemporary theories and approaches which serve as a foundation for different ways of teaching and learning a L2 in the classroom. Depending on what is accepted as theory or practice there are 40-60 to choose from (Long, 2014). Three examples are: The Direct Method which includes the teacher speaking only the L2 in instructional contexts, The Task-Based Approach which provides students with real-life communicative situations and The Grammar Translation Method which is phasing out of Norwegian schools today but includes translation of texts and drilling of grammatical rules. A compilation of many of these theories and approaches is found in Garcia & Martinez (2013). From their studies, it is easy to conclude that both past theorists and current research vary in how best to use these theories and approaches in instructional settings. I am not focusing on one particular L2 teaching method or approach because it seems this varies depending on trends and empirical research in L2 education. However a commonality is that SLA in instructional contexts happens either through implicit acquisition, explicit instruction and learning or a combination of both. Therefore, ideas relating to implicit acquisition and explicit instruction and learning are factors that affect SLA in instructional contexts and are discussed below. In addition, individual factors vary from learner to learner and affect how one learns a L2. In the classroom setting, these factors affect teaching, learning and assessing of English on a daily basis in the Norwegian classroom and are also discussed below.

Implicit acquisition and explicit instruction & learning

Implicit acquisition in regards to language acquisition refers to when language is acquired simply and without conscious operations (Ellis, 2011). In both L1 and simultaneous bilingual SLA, there is only implicit acquisition needed in order to become an expert in a language orally. The learner doesn't know why or how s/he knows something, s/he just knows it. Implicit acquisition takes place as a result of exposure and repetitive experiences in meaningful language settings (Ellis, 2011; Mellanby & Theobald, 2014).

In the instructional context, there can be implicit acquisition even though there are explicit learning outcomes. For example when learning about colours, a teacher in a year 1 or year 2 classroom in Norwegian schools might ask learners, *What colour is your house?* expecting learners to respond with *My house is _____.* If the teacher asks this to 18-20 learners, they not only learn about colours which they are aware of learning (explicit learning), but through repetition, they also learn about word order in both questions and statements, the use of *is* with a singular noun and the use of *my* as a pronoun referring to oneself. By focusing on meaning, learners acquire parts of the language without being aware of it which is implicit acquisition.

When reading, if learners understand main meaning in a text already, they can also acquire new language skills implicitly such as word order and new vocabulary without being aware of it. Steven Krashen (1941-), a known researcher and advocate of implicit SLA states that reading is an overwhelmingly proven way to learn another language (2008). However, it has also been said that when it comes to understanding new words in a text, readers need to know between 90 and 95% of the existing text! (Ellis, 2015; Hummel, 2014; Nagy & Scott in Farrall 2012). This means readers needs to understand most of the text if they are going to learn new, unknown words or understand main ideas of texts.

In contrast to implicit acquisition in the L2 context, explicit instruction and learning can be explained as "any attempts by either learners or teachers to manipulate learning from the outside" (Van Patten, 2011, p.10). Learning to read in any language is an obvious example of explicit learning. Learners need to connect sounds they already know to written symbols. This cannot be done without a conscious effort by either a learner or instructor. Beginning reading abilities can only occur through explicit instruction and learning.

When it comes to SLA, examples of explicit instruction and learning include translation of new vocabulary words or drilling verb conjugation. In addition, when a teacher explains differences of spelling rules of writing *toys* or *boys* compared to *babies* or *parties*, the teacher gives explicit instruction and learners need to make conscious efforts to learn the correct spelling rule. A further example of explicit instruction and learning in SLA is the idea of noticing (Long, 2014; Ellis, 2011). Often there is a grammatical rule that learners do not pick up on if they are in a context that focuses only on meaning or implicit language acquisition. When a teacher draws attention to a specific rule, the learner then notices or is made aware of this rule and can then accommodate new information with existing knowledge of the language. Explicit instruction and learning focuses on form of language rather than meaning of language as in implicit acquisition. After understanding and/or enough repetition of rules, learners do not need to think consciously about language rules anymore and at this point it becomes a part of learners' implicit acquisition. It becomes automatic (Long, 2014; Ellis, 2011; Van Patten, 2011).

It is a goal for the L2 to be automatic in the four areas of language learning (listening, reading, speaking and writing). In a perfect world, the goal is to learn and know a L2 through implicit means, but in a L2 instructional context with limited time and repetition, much of the learning remains explicit (Ellis, 2011). There are discussions amount how much explicit instruction and learning should be focused on in an instructional context. There are those who believe that teachers should create meaningful, natural environments in instructional settings so that learners focus on meaning and communication and acquire the L2 implicitly (Mellanby &Theobald, 2014). Stephen Krashen's Natural Approach is one example of this (Krashen & Terrell, 1995). The Natural Approach, as well as other approaches that are inspired by implicit learning, focuses on communicative, task-based instruction and learning that tries to create real-life meaningful situations using the language. Authors of *Stairs* course books seem to agree with this approach. Both editions of the teachers' manuals clearly state the importance of learners being active and finding meaning in learning English as the L2 (Håkenstad & Undheim Vestgård, 2007; Håkenstad, Morten & Undheim Vestgård, 2013). For children under the age of seven, it has been said that focusing on meaning should be a major focus in the classroom (Long, 2014). Part of the reason for this is because

younger children are not cognitively developed enough to understand processes of a L2 through explicit means, especially grammar skills (Long, 2014).

Others believe teachers should rely more heavily on explicit instruction and learning in the classroom (see examples in Farrall, 2012, p.57). When it comes to grammatical processes, reading, writing and understanding a L2 for children (and adults) ages 12 and up, they rely more heavily on cognitive abilities to learn a L2 (Long, 2014). For example, older learners compare languages with each other in vocabulary learning. In addition, when the L2 grammar system differs from the L1, explicit means of instruction and learning of grammar can be beneficial since learners are not familiar with the grammar system from before (Mellanby & Theobald, 2014; Long, 2014).

Based on experiences and theoretical information addressed above, my standpoint is that an instructional L2 context for English in beginning primary schools in Norway should provide room for both implicit acquisition and explicit instruction and learning. Focusing on meaning through communication is important. Focusing on form of language is also important. Both are necessary if teachers are to meet all of the national competence aims in English throughout primary school and provide learning environments that suit several types of learners at different levels in language learning within the existing allocated time given the subject. Acquiring English in the Norwegian primary classroom whether implicitly, explicitly or a combination of both is highly influenced by teachers. Let's take a look now at factors surrounding L2 learning that are influenced by learners.

Individual factors

From robust findings over four decades, it is clear that learners, not teachers, have most control over their language development, and they do not move from ignorance to native-like command of new items in one step (Long, 2014, p. 22).

Individual differences in a classroom are considerable factors which influence learning in SLA instructional contexts. Individual factors affecting how one learns can both be biological and environmental (Mellanby & Theobald, 2014; Long, 2014). For example, learners are biologically at different stages in their developmental processes which affect attention, understanding and production in instructional SLA contexts (Long, 2014). Environmentally,

learners grow up in different homes which affect learning development and attention or motivation at any given time for any given subject. The following are some detailed examples of both biological and environmental individual factors and how those factors affect acquisition of a L2.

Working memory is a biological cognition process that has been found to influence SLA and differs amongst learners in the classroom. There are several recent studies that look at how working memory is significant to L2 processing and it has been accepted that a large capacity of working memory aids in SLA (Hager, 2011; Alexiou, 2009; Ellis, 2015). To explain it simply, working memory affects how quickly one can take in new information and connect it to existing ideas before storing information in long term memory where it can be easily retrieved. In beginning sequential SLA, new information is unending. Learners connect new information for example to pictures (either tangible or abstract) or current understanding of words in their L1. Some learners need more exposure than others in order for sustained understanding to occur. In instructional SLA contexts in 2010, Ulf Andersson conducted research to see how processes in working memory correlated to second language processing of sentences and short stories in learners ages 9-12 whose L1 was Swedish and L2 was English. In presenting his findings which are also supported by previous studies, Andersson says that learners need to have a "capacious" working memory in order to adapt to a L2 (2010). When there is limited working memory capacity, it is more challenging for learners to acquire a L2 in an instructional context. Learners with a lower capacity of working memory need even more repetition and experiences using the L2 in order for new information to become automatic and this is especially true in implicit SLA settings (Ellis, 2015). Working memory is an example of one biological factor that contributes to how or why students process and produce a L2 at different rates in an instructional context. (Other studies about working memory include Alexiou, 2009; Hummel, 2014, chapter 8).

Although obvious to some, it is important to mention how cognitive skills in the L1 generally link directly with the L2. If learners have challenges decoding in their L1, they will also have problems decoding in their L2, which affects learners' overall reading abilities. Likewise, if learners are quick to make inferences in the L1 when reading, they will also be quick to make inferences in the L2, providing they understand the words being read.

Environmental factors connected to learners' L1 also affect the rate at which the L2 is learned. It has been accepted that language ability both in reading and oral communication in the L1 strengthens the L2 (Mellanby & Theobald, 2014; Andersson, 2009; Ellis, 2015; Farrall, 2012). The opposite is also true. In 1995 Hart and Risley documented language experiences in the home environment in children ages one to three years old and found that language experiences differed greatly in families. Language experiences varied from three to 11 million words in a year for children in their first three years (in Farrall, 2012). Hart and Risley compared these experiences to language acquisition and skills at age nine and found that those who were exposed to more words and had greater language experiences at home in their first three years had a more secure platform in which to acquire language and reading skills. This affects a learner's L1 and in turn, the L2. In addition, those who have a firm foundation in their L1 are able to use transfer to a greater degree in order to understand the L2 better. Transfer involves using what is known or has been learned in one language in order to learn or understand another (Hummel, 2014).

Besides cognitive processes such as working memory and environmental factors such as language experiences, there are other reasons why learners learn a L2 at different rates. Motivation and attitude for learning English are individual factors that vary. They are connected to the acquisition process because motivation and attitude affect the amount of attention given a L2. This is especially true in skill building because skill building requires more attention (Seland, Vibe & Hovdhaugen, 2013; Hummel, 2014). Robert Gardner (1985) is known for his work in motivation and attitude in SLA (in Hummel, 2014). He says that attitudes affect motivation which affects overall SLA. Krashen also agrees that motivation in learning a L2 is undeniably important for successful SLA (2008). Unlike some biological factors, attitudes and motivation are not static and can change. When working with young learners in an instruction SLA context, teachers can potentially greatly influence individual factors in this area. As the quote at the beginning of this section implies, if learners themselves are not actively engaged in learning processes in the English L2 classroom, it is difficult to acquire English at all. Therefore, it is important for teachers to create learning environments that motivate and engage learners in their own learning processes, at any age.

The above factors are not inclusive of all individual factors that contribute to SLA. Learning styles, social, emotional and learning challenges in individuals are among some of the other factors. However, the above does attempt to explain that individual factors affect the overall rate of learning and success of SLA in instructional contexts. I believe it is important to include information about individual differences because they not only explain the differences in individual learners' acquisition of L2, but also the reasoning behind why teachers' may choose one L2 method or approach over another. Although teachers do not have access to each learner's cognitive processes such as working memory nor do they have a clear picture of how much L1 has been used in the home, competent teachers can clearly see which learners grasp new words and phrases quickly and which learners need constant repetition. Similarly, they see which learners have difficulty putting sounds together in English to make words when reading and which students use decoding skills from their L1 to read English. Teachers get an overview of which learners focus on understanding each word in a text and which learners try to grasp overall content, using words they know to derive meaning. Teachers also see which learners are already motivated to learn English and which learners they need to motivate for the subject. Further, they (should) have pedagogical professional knowledge to understand that learners are motivated by different means. Therefore, in SLA instructional contexts, I believe individual differences in learners is one factor influencing the approach(es) used in instructional contexts (whether implicit or explicit or both).

I find it important to mention how SLA research in naturalistic contexts is sometimes used as a basis for generalizations supporting expectations for organizational choices in instructional contexts (Munoz, 2006; Mellanby & Theobald, 2014). As a final reflection in this section, I would like to resonate on a current situation where a colleague of mine, who is currently studying English as part of teacher training, has been told to only speak English to learners (The Direct Method). This is an example where research from naturalistic SLA contexts has been used to support one specific method of teaching and learning in an instructional context. In a naturalistic setting, for example where a learner from Afghanistan is learning Norwegian in a Norwegian school, The Direct Method of learning the L2 is obvious and unavoidable. However, by only speaking English in a Norwegian primary classroom with

limited exposure to the language may not take into account learners' individual needs. If learners are overwhelmed, not motivated or do not have the attention span to figure out what is being said in English during one lesson a week, they may switch off and learning will not be accomplished. My point here is that when teaching English as a L2 in the Norwegian primary classroom today, I believe it is imperative as it is with all subjects, that teachers know the subject and their learners when making decisions about how to teach in the classroom realizing that the method(s) chosen may differ from year to year and class to class. I am not against The Direct Method and I have indeed used this method at times as well. I do not, however, agree with promoting only one specific method of L2 in instructional contexts since, as we have seen above, there are several factors influencing the approaches used in SLA instructional settings.

Assessment

When we assess children…..we need to make our decisions based on careful reading of research, our knowledge of children as learners and our knowledge of the tools of assessment (Farrall, 2012, p. 26).

According to *The Glossary of Education Reform* (GER), in education "assessment refers to the wide variety of methods or tools that educators use to evaluate, measure and document the academic readiness, learning progress, skill acquisition or educational needs of students" (2015). In the L2 classroom, competent teachers informally and formally assess learners constantly through both oral and written means in individual and group settings. In a Norwegian primary classroom, this assessment covers the four skills of language learning, the basic skills and the competence aims.

In this section of the thesis, I often cite Melissa Lee Farrall, PhD and her book from 2012 entitled *Reading Assessment: Linking Language, Literacy and Cognition*. I find her information about reading and assessment to be clear and easy to understand for a professional teacher like myself, who is not an expert on reading comprehension assessment and testing.

National tests as assessments

National tests in English are, of course, one type of assessment. There are many decisions to make before the question of content is even considered in such a test. For example, should all four basic skills of language be tested? Should learners have the same test or should the tests vary according to learners' individual needs? During which year of schooling should national tests be taken? Should the test be criterion-referenced or norm-referenced? (Farrall, 2012).

Norway differs from its neighboring countries in these decisions. In Denmark, for example the national test in English is taken in year 7. It is a 45-minute computerized adaptive test which focuses on reading, vocabulary & language and language use (Ministry for Children, Education and Gender Equality, 2016). A computerized adaptive test means that the computer adapts or adjusts the questions throughout the test based on learners' responses to previous exercises, making each test suitable to the individual learner (GER, 2015). Both Norwegian and Danish national tests in English use primarily multiple choice and fill-in-the-blank questions. Both tests are scored by the computer. Similar to the English test in Denmark, Norway's 60-minute computerized national test assesses one of the four areas of language: reading comprehension, which includes vocabulary and language use (Ministry for Children, Education and Gender Equality, 2016; UDIR, 2016). In contrast, the national test in English in Sweden tests all four areas of language: listening, speaking, reading and writing. It is taken in year 6. Teachers, who teach the learners, correct the tests with help of an instructional guide (Skoleverket, 2015). Since it is scored by teachers, results in Sweden are more subjective when compared to national tests in English in Norway and Denmark. Subjective results make it challenging to compare students/schools/communities to each other but easier to use test results for further learning in the classroom (Farrall, 2012).

In Norway, the national test in English is a standardized test. This means the test is made for a large population and scored in a standardized or consistent way so that results can be compared with other learners or populations of learners (GER, 2015). The test is norm-referenced, which means the results reference how learners perform *in relationship to their peers*. The advantage of norm-referenced tests is that it is easy to compare learners' skills with each other and use the results to adjust national educational plans (Farrall, 2012). In

contrast, two major disadvantages with norm-referenced tests are that they do not necessarily assess what learners know and they do not allow for optimal further learning (Farrall, 2012).

Reading Comprehension

As stated previously, national tests in English focus on reading comprehension and vocabulary so let's look at how reading comprehension and vocabulary are tested.

First of all, for reading comprehension to occur, the reader has to above all make some type of connection to the text. This connection is generally based on learners' background knowledge (Farrall, 2012; Alderson, 2005). Some of learners' background knowledge is related to previous experiences. Going to birthday parties, travelling on a bus, train or plane, experiencing foggy weather and experiencing sadness all help 10-year-old readers relate to, understand and make inferences in texts. Exposure and experiences with literature, hearing and retelling stories also are included in background knowledge which helps learners relate to texts. In addition, learners' history, culture and knowledge of factual information create a data bank that can be drawn upon when relating to texts as well. It is nearly impossible to test an individual learner's background knowledge and how s/he relates that to a specific text, however test-makers try to include a variety of topics of both familiar and unfamiliar topics so that all learners can (hopefully) use background knowledge to relate to some texts. Background knowledge is necessary to recognize when testing reading comprehension because background knowledge helps in understanding texts. If the reader cannot make any connection to a text, it makes it difficult to understand, in both a L1 and L2.

When testing reading comprehension, it is accepted that three broad areas are referred to:

1. understanding the main purpose/main idea of a text
2. finding specific information/details in a text
3. making inferences (which requires going beyond the literal meaning of a text) (Alderson, 2005; Farrall, 2012; UDIR, 2016).

Information from the 2016 instructional guide for teachers showing how test exercises relate to these three areas is include in appendix 4 on page 69. These three areas are also reflected in the reading portion of the *Framework for Basic Skills* document (2012). Apart from the

reflect & assess category on the *Framework for Basic Skills* document, the year 5 national test in English assesses learners varying levels of reading in these areas.

When it comes to vocabulary, UDIR says that understanding of words individually is not enough for testing reading comprehension and that mastery of vocabulary is realized by understanding words and phrases within a text (2016). Although this is true, it must also be iterated that if learners do not understand individual words in front of them (especially in their L2), it is impossible to decipher main ideas, find details or make inferences and connections within a text, let alone find motivation to continue reading. The authors of the first edition of *Stairs* say in the teaching manual that reading comprehension is difficult if learners do not work with and understand vocabulary first (Håkenstad & Undheim Vestgård, 2007). Farrall says that in order to truly understand a new word, learners need exposure to that word through different contexts up to 12 times (2012, p.241). Here she is referring to the L1 but constant exposure and use of vocabulary words are extremely important in the L2 as well when it comes to vocabulary development. Remember that readers need to understand about 90% of the words in order to make sense of the text.

Understanding vocabulary plays one part of reading comprehension. Farrall says that complexity and length of sentences play a role in reading comprehension as well (2012). In general, the longer the sentences are, the more difficult they are to comprehend. Additionally, she makes a distinct difference between simple sentences, compound sentences and complex sentences. Simple sentences with few words are often the easiest to understand. Learners need less time to derive meaning from these types of sentences. Texts build in their complexity when they move to compound and complex sentences or when there is a combination of these types of sentences (See Farrall, 2012, p.151 for further clarification).

Recognizing how vocabulary, as well as complexity and length of sentences directly influence reading comprehension, has led to decisions I've made in part of my investigation of year 5 national tests in English. Keeping in mind the above information about reading assessment, as well as how a L2 is acquired in an instructional context and some factors influencing methods used in the classroom, let's now move on to chapter IV. This will explain details of my investigation and the method of document content analysis.

Chapter IV: Method

In the introduction of this thesis, I mentioned Tone Kvernbekk's theoretical ideas relating to being an insider or an outsider when investigating a phenomenon. I made the connection that I, as a teacher, have an insider role when analyzing content of year 5 national tests in English. According to Kvernbekk, by being an integral participant in an activity, insiders use all sides of themselves including knowledge, perceptions, experiences and general discoveries and this often creates shared values and/or understandings amongst insiders (2005). I would like to think there are other teachers in Norway who share similar perceptions and experiences when it comes to a mismatch between English competence aims at the end of year 4 and expectations on year 5 national tests. Therefore, I hope that my viewpoint as an insider reflects views from other insiders as well.

The perception of an insider is both a strength and a weakness. One strength is that insiders sit with privileged knowledge and/or understanding of a phenomenon in a unique way to that of outsiders. Outsiders can look at the same phenomenon and take it as they see it, whereas insiders sit with knowledge or perceptions about characteristics that are not directly observable by outsiders (Kvernbekk, 2005). School administrators, test-makers and politicians for example, can clearly observe outcomes of national tests to make educational reforms or compare students and schools with each other. However, they are not likely to see individual learners in test situations.

In contrast, a weakness to being an insider is that insiders are invested in a specific activity. Insiders must be cautious to not let feelings get in the way of objective perceptions (Kvernbekk, 2005). Because feelings, perceptions, knowledge, experiences, decisions, attitudes and values are delicately intertwined, this objectivity is not always easy for insiders.

Throughout this thesis, the overall goal is to understand the mismatch between year 5 national tests in English and the competence aims in English at the end of year 4. Not knowing the answers to these questions, I find the investigation to be exploratory. Research questions I asked have led me to the approach of document content analysis of several of the year 5 national tests in English. By comparing content in several tests to each other and to the competence aims, I hope this leads to possible conclusions about why the tests are

difficult for many learners. Due to my role as an insider, it is necessary that I use an approach to the analysis that is as objective as possible and also quantifies meaningful results. In order to minimalize subjectivity and maximize objectivity in the following analysis of national tests, I have therefore used the approach of document content analysis.

Content analysis of documents

Content analysis has been an accepted research technique since the early 1900s (Holsti, 1969). Together with Philip J. Stone, Ole Holsti defined content analysis as "any technique for making inferences by objectively and systematically identifying specified characteristics of messages" (1969, p. 14). Instead of using the word *messages*, Alan Bryman refers to "documents and texts" (2008, p. 274) in his definition of content analysis. Therefore, in my investigation, the *messages* that Holsti refers to are the documents which make up year 5 national tests in English from 2009-2016.

To achieve objectivity in document content analysis, it is necessary to use both a coding schedule and a coding manual. Categories in the coding schedule and coding manual should be steered by research questions (Bryman, 2008). A coding schedule is a chart which shows results of categories being analyzed in each document. Each result is given a code and by comparing codes as opposed to words, it makes it easier to structurally organize and analyze content in documents.

The Coding Schedule

Below is an example from my investigation in order to give a visual idea of what a coding schedule looks like.

	A	B	C	D	E	F	G	H	I
1	Coding Schedule for year 5 English national test content analysis								
2									
3					(18 choices)	(15 choices)	(3 choices)	(3 choices)	(6 choices)
4	test year	exercise #		# of words	topic	exercise type	sentence structure	sentence length	verb use
5	2010	1		16	4	1	1	1,2	1
6		2		14	2,4,11	1	1	1	1,2
7		3		7	18	1	1	1	1
8		4		33	2,11	1	1,3	1,2	1
9		5		41	2,10,17	4	1,3	1,2	1,2,4
10		6		9	4,12	4	1	1	1
11		7		51	2,4,12	4	1,3	1,2	1,6

The columns referring to *test year* and *exercise number* simply help with organization so that it is easy to know which test year and exercise number the analyzed categories refer to. Categories being analyzed are: *number of words, topic, exercise type*, *sentence structure, sentence length* and *verb use*. The number of choices, as seen in row three, refers to how many code options there are for any particular category. For example, there are 18 different code options which relate to the category of *topics*. Each number below the categories, then, refers to a specific code. For example, under *topics*, exercise 1, has been assigned code 4. Likewise, under *topics* exercise 2 has been assigned code numbers 2, 4, and 11. These codes, however, mean nothing without a coding manual. The coding manual is found in appendix 5 on page 70. It serves as an instructional guide and explains what the codes, or in this case numbers, refer to.

By using the coding manual, it is easy to understand that in the previous example under *topics,* exercise 1 which has code 4, actually means the topic of exercise 1 is animals. Anytime code 4 is used under *topics*, it is referring to the topic of animals.

The content analysis is both quantitative and qualitative. A quantitative analysis is an analysis which interprets results based on numbers (Hosti, 1969). Some of the results reflect this as I make interpretations based on frequency (or lack of frequency) of a code on the coding schedule. By making inferences based on details of the analysis and not merely frequency results of the coding schedule, makes it also qualitative. Holsti says that qualitative and quantitative content analysis supplement each other and that "by moving

back and forth between these approaches, the investigator is most likely to gain insight into the meaning of data" (1969, p. 11). Thus by using both qualitative and quantitative content analysis of year 5 national tests in English, I hope the findings to be valuable and meaningful.

In large content analysis projects where there are several researchers, there needs to be inter-coder reliability (Bryman, 2008). This is a set of rules that enable several researchers to apply the same code in each category and ensures validity and reliability. In the investigation I have carried out where I am the only researcher, rules still need to be made. This is not for inter-coder reliability purposes but rather for intra-coder reliability purposes (Bryman, 2008). Basically, there needs to be rules or attributes assigned options in each category so that I continually agree with myself when assigning codes. Detailed attributes for categories on the coding schedule and coding manual are explained further in this chapter under the title *Explanations of the coding schedule*.

The fact that results of the same codes should (hopefully) be found by others, makes results repeatable and thus reliable (Bryman, 2008). Holsti says that if attributes have been formulated well in a content analysis, the inferences made afterwards are meaningful (1969). That being said, I am aware that interpretations of results in a content analysis can be debated depending on the one who interprets them. Thus, it is necessary to mention the role hermeneutics play in interpretation and analysis of results in the coding schedule. According to Hans-Georg Gadamer (1900-2002) and his accepted contemporary views on hermeneutics, a person's past, values, traditions and previous knowledge are some factors which influence one's pre-understandings in any particular situation (2004). The interpretations I make are influenced by the fact that I am a teacher and not a test-maker for example. Another factor influencing my interpretations is that I speak English as my L1 and teach it to learners in their L2. Who I am influences the interpretations I have made from findings in the investigation. I recognize that other professionals may have different interpretations from the same results.

Let me explain the process I have gone through for this investigation of year 5 national tests in English. I will explain categories I have analyzed in the coding schedule and attributes given options in each category, when necessary.

Explaining the process

I chose to analyze some of the content in national tests in English from 2009 until 2016. Although national tests began in 2004, in 2008-2009, there was a shift in the type of software platform used in taking national tests (Fikke & Helness, 2012). Therefore, UDIR does not have digital copies of tests prior to 2009 which is why I have not analyzed tests previous to 2009. Each year, more than one version of the test is given, but I have only looked at one version of each test. This version is the one I received from UDIR when I contacted them in September 2016. UDIR sent me one version of each test from 2009-2013 and I used the existing versions of the 2014, 2015 and 2016 tests that were found on UDIR's website, which was also suggested to me via email through UDIR (the 2014 version is no longer available on the website.) I chose to look at several tests instead of analyzing just one test in hopes of finding some correlations between tests or evidence of change in analysis of the tests.

Research on the tests was completed in autumn and winter of 2016. I read and reread the tests several times in order to find answers to the following three questions.

1. **What types of exercises are included in national tests and how many words are included in each exercise?**

2. **How complex are the exercises and how has this changed over time?** I look at sentence structure, sentence length and verb use. The results here are purely quantitative as results are based on measurable interpretations found in the coding schedule.

3. **How do exercises relate to the four competence aims that the test refers to?** I look at topics, and interpretations of three specific words that are repetitively used in the competence aims that UDIR says the tests are based upon. Initial results are quantitative based on results in the coding schedule, but qualitative content analysis is also used to interpret findings in detail.

I ask the first question in order to get a general overview of tests and their content. The second question evolved out of ideas from Melisssa Lee Farrall and understanding how sentence complexity and sentence length affects reading comprehension in general, whether in a L1 or a L2. Learning a L2 has several components but I wonder if there are

implications that exercises generally test reading comprehension and if so, I wonder if this can explain a mismatch. I believe the third questions is the heart of the analysis because it is here, that I critically analyze tests in comparison to the four competence aims UDIR currently says the tests are based upon.

Information on the coding schedule relates to one of these three questions. So, it is the questions themselves, which have led to inclusion of categories in the coding schedule. Below is information about categories used in the coding schedule and attributes that aided in analyzing the categories consistently.

Explanations of the Coding Schedule

Categories included in the coding schedule are:

Test Year: The test year ranges from 2009-2016.

Exercise number: The exercise number relates directly to the exercise number used on each test. This ranges from 1 to 49.

Number of words: This helps to answer the first question. When counting number of words, I do not include the initial instructional text which is *italicized* in each exercise. For example "Read the text. Click on the correct item." is not included. I do not include initial instructional texts in overall word-count because this short text is included in all exercises. However, if there is further instruction in an exercise after the initial instructional text, I include these words in the overall word-count. For example, here is the text after the instructional text in exercise 3 from 2012: "There are two rugs in the room. One is under the footstool Roger is sitting on. The other is hanging above the sofa. Click on the one on the wall." Here, the words "click on the one on the wall" are included because this is part of the exercise text and not merely the instructional text. I am able to observe if word-count in different exercise types changed over the years and to see which types of exercises used most words.

Topic of exercise: This helps to answer the third question. By looking at the topic of each exercise, I specifically observe how topics relate to two of the four competence aims referred to in the national tests which state that learners can

- read, understand (and write) English words and expressions related to one's needs and feelings, daily life, leisure time and own interests.
- understand the main content of simple texts about familiar topics.

I use 17 general topics to see how they fit into feelings, daily life, leisure time and interests. Out of all the exercises in eight tests analyzed, there are five exercises that do not fit into any of the general topics and are assigned code 18 which stands for *other*. Therefore, there are actually 18 codes in this category. Sometimes, an exercise includes more than one topic. This is reflected in the coding schedule by including more than one code here as well. When referring to topics included in the tests, other investigators may have chosen to include more (detailed) topics to or less (broader) topics. Keep in mind, the goal of analyzing topics is to see if/how they align with the competence aims mentioned above, therefore, minor discrepancies in details about including more or less topics should not influence general observations of topics within the texts when referring to the two competence aims above.

Exercise type: This helps to answer the first question. I identify 15 different types of exercises. (Examples and descriptions of the most common exercise types are included in chapter V.) I observe which types of exercises are consistently part of the English national tests and which exercises have developed and then not been used again. I also observe the frequency of each exercise type on the tests.

Sentence structure: This helps to answer the second question. Using ideas from Farrall, I observe the use of simple sentences, compound sentences and/or complex sentences in each exercise. I have not gone into further depth by looking at if sentences have compound subjects or predicates or the type of subordinators used in complex sentences. Although delving deeper in these areas would contribute to the discussion about the complexity of sentences themselves and exercises as a whole, I believe the general sentence type also shows differences in complexity in sentences and exercises. I see if sentence complexity has changed throughout the years. Some exercises include more than one sentence type. This is reflected on the coding schedule by including more than one code here as well.

Sentence length: This helps to answer the second question. Using ideas from Farrall about how sentence length relates to sentence complexity and therefore overall complexity within reading comprehension, I set three codes according to sentence length. I observe if sentence length has changed throughout the years and if a correlation between exercise types and sentence length exists. Some exercises include sentences of varying lengths. This is reflected on the coding schedule by including more than one code here as well.

Verb use: This helps to answer the second and third questions. This category developed as I was reading through tests and after I started initial research. I became aware of different verb tenses used in texts. I made a connection that present tense verbs are the verb format learners have most exposure to during beginning primary education in English classrooms in Norway. They are the first types of verbs that are explicitly taught in the primary classroom as seen in course books such as *Stairs* (Håkenstad, Undheim Vestgård, 2007; Håkenstad, Morten & Undheim Vestgård, 2014). Also, in my experiences as a L2 learner of three languages, I know that present tense is often explicitly taught before other tenses. Knowing this, I became interested in finding out which verb tenses are used, if this has changed throughout the years and if there is a correlation between exercise type and verb use. Some exercises include more than one type of verb and this is reflected in the coding schedule by using more than one code here as well.

Simple/common/familiar

I have now explained information found on the coding schedule and attributes used to ensure intra-coder reliability. Lastly, I need to explain choices made surrounding analyzing vocabulary words and expressions in national tests. This helps to answer the third question. Three of four competence aims that the national tests are currently based upon specify that learners should understand "*simple* texts about *familiar* topics", use "*common short* words and *simple* spelling-and sentence structures" and use "*simple* reading strategies" (UDIR, 2016, italics mine). What do the words *simple*, *common* and *familiar* mean? Perhaps my interpretation of simple or common words for 10-year-olds after working with them for several years in the classroom is different than test-makers or linguists who study L2. Perhaps differences in how the words *simple*, *common* and *familiar* are interpreted can help

explain the mismatch between content in national tests and competence aims. However, is it possible to objectively, structurally find out what is *simple*, *common* or *familiar* when it comes to vocabulary used on national tests? These thoughts and questions have steered the final portion of my investigation but cannot be contained to the method of document content analysis, coding schedule and coding manual which is why I find it deserving of its own explanation.

I believe it is possible to come to a conclusion about what 10-year-olds, teachers, linguists who specialize in L2 and test-makers view as *simple*, *common* or *familiar* words on year 5 national tests in English. However, this would require another research investigation of its own. Whether by way of surveys given to a sample of the above mentioned groups or by carrying out investigations through discussion or interview settings about words and expressions used on national tests, conclusions could be drawn. However, there needs to be a starting point and this is where my investigation can be beneficial. I have made charts from each test about words and expressions that I question as *simple*, *common* or *familiar* to 10-year-olds. (The charts are included in appendix 11 on page 77.) The vocabulary chosen is based on background knowledge presented in chapter II that 90-95% of words need to be understood in order to comprehend texts, the understanding of the types of texts used in curriculum in years 1-4 and tacit knowledge that comes from my experience in the English L2 classroom in Norway. These charts can be debated and I hope they will be. I argue, however, that some conclusions can be drawn based on vocabulary in the exercises. These conclusions are explained in the next chapter. In addition, the charts can serve as a starting point for other research investigations in the future.

Chapter V: Results

By researching several tests at a time, it is possible to find correlations between tests and see if or how some of the content has changed throughout the years from 2009 to 2016. By interpreting results from document content analysis and reflecting upon what is known about SLA in early primary education, conclusions can be drawn that attempt to explain the mismatch between the competence aims at the end of year 4 and the year 5 national tests in English which is the intention of this chapter. Chapter V is divided into three parts, relating to the three questions that have steered the investigation of document content analysis. There are several examples and I also discuss conclusions or implications based on the findings.

Question # 1: What types of exercises are included in national tests and how many words are included in each exercise?

The instructional manual for teachers from 2016 includes a list of nine exercise types that are used in the 2016 national test (UDIR), but since I have looked at tests since 2009, I have described them slightly differently. There are seven exercise types that have been used consistently in all tests from 2009-2016. They are:

- *Read text, find correct item and put in correct place*

- *Read text and click on correct picture*

- *Read several texts and match with one picture*

- *Fill-in-the-blank*

- *Read and answer a question about a text*

- *Read and answer questions about the same text*

- *Who-could-say*

In addition, ***click on the correct item*** is an exercise used on all tests except for 2015 and ***set the time*** exercises have been used since 2011.

Below are examples of the exercise types mentioned above. Representations from all of the tests from 2009 to 2016 are included in the examples. Sometimes two examples are included to show differences in difficulty or length in the same exercise type. I find it

relevant to include examples because this investigation is based on content and it is necessary to provide content examples instead of only writing about them.

Click on the correct item

This type of exercise has been used on all tests except for 2015. This type of exercise comes first on the tests and includes understanding a short text. These exercises range from five to 40 words and there are between two and five exercises of this type on each test. The below examples are from 2016.

Oppgave 1

Oppgave 2

Read text, find correct item and put in correct place

This type of exercise includes finding important information from a text, understanding words and understanding prepositions. There are between two and four exercises of this type on each test ranging from 20-70 words. The following example is from 2014.

Read the text. Click and drag.

My sister runs every afternoon and she has just bought a new pair of colourful trainers. She has put them on top of the white chest of drawers. Put the item in the correct place.

Read text and click on the correct picture

This type of exercise includes understanding important information in a text. There are between four and eight exercises of this type on each test ranging from eight to 73 words. The following examples are from 2010.

Oppgave 12

Read the text. Click on the correct picture.

Every summer we see a lot of goats up in the mountains.

Oppgave 13

Read the text. Click on the correct picture.

The captain of our soccer team works in the local factory. They produce newspaper and it's her job to run one of the machines. The local council is concerned about all the smoke that is released from the two chimneys.

Read several texts and match with one picture

This type of exercise includes understanding important information in a text. There are between two and five exercises of this type on each test ranging from 50-165 words. The first example is from 2009 and the second is from 2015.

Oppgave 17

Look at the picture. Click on the correct text.

My father woke up with a stomachache and didn't get up all day yesterday.	I am going for a long walk this afternoon. I know what to wear on my feet.	It's getting colder and I need to find something warm to put around my neck.	After the basketball match, I was so thirsty. I went to the fridge and drank some milk.

Look at the picture. Click on the correct text.

Carol is visiting her grandparents and she likes to do odd jobs to help them out. Yesterday she offered to make cupcakes, but her grandpa asked her to help him saw up a tree he had just cut down in the garden instead. Carol really enjoyed it!

Carol's parents have just had a tree in their garden cut down. She has promised to help them saw it up when she finishes her science project. She is not going to rush with her project because she enjoys looking at small organisms through the lens.

Carol's parents asked her if she could make some cupcakes for her younger brother's birthday. Carol would love to, but she doesn't have time. She has to fix her bicycle in order to get to school tomorrow. The chain broke on her way home today.

Carol wants to go over to her grandparents' place after school and help them with some firewood. Her mother tells her that that has to wait because she needs Carol to make her special cupcakes. It's her little brother's birthday tomorrow.

Fill-in-the-blank

This type of exercise includes understanding the use of common words within a text, generally relating to grammar and sentence structure. There are between two and six exercises of this type which include one to four blanks in each question. The number of words used in each exercise ranges from 15-100. The following example is from 2013. I have chosen to include this particular example because it shows if learners understand pronouns, the verbs *to be* and *to have*, and use of the definite or indefinite article, all of which relate to grammar skills.

Oppgave 27

Read the text. Click on the correct word.
My best friend Janet has a beautiful little pony called Sunshine. Janet lives on a farm. I get to ride Sunshine when I visit [she, he, him, her] . Janet has two rabbits and a hamster, too. It is so much fun to [is, are, be, am] on the farm and to help her take the animals out and feed them. I live in [the, that, a, an] small flat with my parents and sister, Lea Ann. She is allergic to fur so we can't [have, has, had, having] a pet. After playing with Janet and her pets, I have to change clothes and wash my hands before going home.

Read and answer a question about a text

This type of exercise includes understanding the main idea and/or details of a text. As in the following example, this exercise also tests learners' ability to make inferences. The question comes after the text as opposed to before the text as in most exercises. There are between one and five exercises of this type on each test ranging from 37 to 140 words. The following examples are from 2012 and 2016.

Oppgave 35

Read the text. Click on the correct answer.
People are rushing to the zoo to see Sara the walrus. She can hold a saxophone between her flippers and play a note. The children love the show.

What can Sara the walrus do?
- ○ Play with children
- ○ Play with a ball
- ○ Play some games
- ○ Play some music

Oppgave 26

Read the text. Click on the correct answer.

Dear Sarah

I'm having a great time here at Grandma's! I have met a nice girl who lives next door. We are going to the cinema tonight. Yesterday, Grandma and I went swimming at the local pool. Tomorrow, we are going to an adventure park. I think my parents will pick me up next week. John is coming along for the ride. I miss you! See you in a week!

Hugs from Diana

Where is Diana staying?
- ○ With John
- ○ With Sarah
- ○ With her parents
- ○ With her grandmother

Read and answer different questions about a text

This type of exercise includes understanding the main idea and details of a text and sometimes, learners need to make inferences. The same text is the basis for three to five exercises ranging from 135-230 words. The following example is from 2011.

Oppgave 28

THE PARTY - QUESTIONS 28-31

Read the text. Click on the correct answer.
Noah's class was having an end-of-year party. They had invited all the parents. Their teacher helped them to plan it. They had baked cakes and made popcorn. The whole class was also going to sing and dance on stage.

Noah was not looking forward to it. It was the singing and dancing that was the problem. He hated to be on stage! Just the thought of everyone staring at him made him sweat. It didn't help that all his classmates were going to be on stage with him.

When it was time to go on stage, Noah tried to hide behind his friends. He was shaking and his feet could hardly move. The music started, and they all began to sing. Noah looked down at the audience, but because of the bright light in his eyes, he could hardly see anyone. Then the strangest thing happened. He wasn't afraid anymore! He sang along, and his feet started to dance, just like when they had practised. He had a great time! Nothing mattered as long as he couldn't see anyone!

Who had prepared the party?
○ The class and their teacher
○ The class and their parents
○ The teacher and the parents
○ The class, teacher and parents

Who-could-say

This type of exercise includes understanding the main idea of a text and making inferences. The same text is the basis for five to six exercises ranging from 130-220 words. These exercises have made up the last exercises on the test, except for in 2016 where there were two fill-in-the-blank exercises at the end of the test. The following examples are from 2010 and 2015.

Oppgave 39

GUINEA PIG - QUESTIONS 34-39

Read the text. Click on the name of the person who could be speaking.
Who could say: "The boys will have to watch their pets"?

Mathew has a pet guinea pig called Spot. It lives in a cage in his bedroom. One day Mathew's friend Tommy came over to play. Tommy brought his new puppy, Max. Mathew's mother, Mrs Heather, told them to keep an eye on the puppy. She was worried that he would frighten the guinea pig. Max sniffed at the guinea pig's cage. He looked like a giant to Spot. She hid in her nest. Max climbed up on top of the cage. Then the boys saw what was happening. Tommy picked up Max and put him out in the garden. At last Spot could come out. Mathew gave her a treat.

Read the text. Click on the name of the person who could be speaking.

Who could say: "This was something I had dreamed of doing"?

Cheryl and her family went to South Africa for Christmas. They visited one of the large game parks and saw a lot of elephants, leopards, giraffes and impala. It was very different from seeing the animals on the screen! She took lots of pictures to show her friends.

Zack and his father visited Arenal Volcano National Park in Costa Rica last summer. He had always wanted to see a live volcano. Few plants grew on the lava fields, and hot lava and glowing rocks were shooting into the air. They didn't get close to the peak.

Lynn and her older sister went to Croatia last Easter on a kayaking holiday. They spent the first day learning what to do if the kayak tipped. The rest of the holiday was spent paddling along the coast, exploring caves and swimming on remote beaches. They even saw some dolphins!

Christopher visited his American online friend last summer. They went to see the Grand Canyon. It took them a long time to hike down in the heat and he was glad they didn't have to climb up again. Instead, they took a raft down the river.
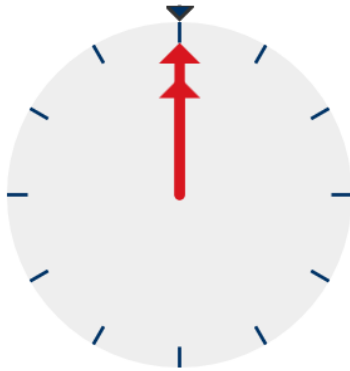
Set the time

This type of exercise includes the ability to tell time. Since 2011, there have been between two and four exercises on tests using about 10 words. The following example is from 2014.

There are certain types of exercises that have come and gone throughout the years. They include exercises about: colour, the calendar, charts, reading notices and read several texts and answer one question. These exercise types are not included in the above examples.

General discussion and conclusions about question #1:

When it comes to overall amount of words in tests, there are necessary comments that need to be made. First of all, in certain exercise types, the word-count is roughly similar in the years from 2009 to 2016. However, there are two exercise types that have significantly increased in word-count from 2009-2016. The *read and answer questions about the same text* exercises have expanded their word-count by almost 100 words since 2009. Similarly the *who-could-say* exercises have seen a gradual increase in word-count as well, with a steady increase by 80 words or more.

Secondly, questions relating to colour were tested prior to and including 2012. From 2013 and onwards, questions do not relate directly to colour. Similarly, tests including and preceding 2012 had questions that simply tested certain vocabulary words. For example, the text from 2011, exercise 1 says "click on the shoes". Exercises do not exist in this simple

form after 2012. Instead, learners need to understand several words in order to derive meaning from the text and connect it to a picture.

Furthermore, tests including and preceding 2012 had exercises relating to placement that directly tested skills in using prepositions and understanding specific vocabulary. For example the text from 2012, exercise 12 says "put the chair next to the sink." From 2013 onwards, there are still questions that focus on prepositions. However, texts surrounding these questions are longer, implying these exercises are not only testing learners' skills relating to prepositions and individual vocabulary words, but also reading comprehension in general.

These findings point towards an obvious change between 2012 and 2013 such that simple vocabulary in and of itself was and is no longer tested, but rather understanding vocabulary in its context is tested. The ability to understand vocabulary in a context, instead of only understanding a word or expression, makes reading comprehension more complicated and more demanding in a L2.

Question #2: How complex are the exercises and how has this changed over time?

Keeping this in mind, I also found results showing a change in reading comprehension complexity, especially between 2012 and 2013. Below, I resonate about how sentence type, length of sentences and use of verbs correlate with complexity of sentence structure and therefore overall reading comprehension complexity in texts.

Types of sentences, sentence length and verb use

Looking at types of sentences first, all tests from 2009-2016 include simple, compound and complex sentences in the texts. I looked at the amount of exercises in each test that included one or more complex sentences and which exercises included only simple sentences. In general, results from bar chart 1 (appendix 6 on page 72) indicate that it has always been necessary for learners to show understanding of texts that include some compound and complex sentences. However, bar chart 2 (appendix 7 on page 73) clearly indicates that the amount of exercises using only simple sentences has decreased since 2009. Simple sentences are easier to comprehend. Therefore a decrease in using only simple sentences points towards an increase in exercises containing compound and complex sentences which is an increase in reading comprehension complexity.

When it comes to sentence length, all tests have some exercises with sentences containing more than 20 words and all tests have exercises with sentences containing 10 words or less. When comparing the amount of exercises that use only sentences with 10 words or less, I found a significant difference between 2012 and 2013. (See the bar chart 3, appendix 8 on page 74 for further details.) Tests after and including 2013 include more exercises with longer sentences (by at least 10%).

When it comes to verb use, I looked at the percentage of exercises on tests that include texts only in the present tense. By looking at bar chart 4 (appendix 9 on page 75) results show that between 55% and 70% of exercises on tests, include other verb forms than present tense verbs. This means learners have to be able to derive meaning from texts that have many verb forms, including those which have not necessarily been explicitly taught. Relating this back to theoretical information on implicit and explicit learning situations, these results imply that it is necessary for learners to be able to understand meaning in texts based on implicit learning since several verb forms are not explicitly taught by the end of year 4.

General conclusions about question #2
Analyzing sentence structure, sentence length and verb use gives insight into reading comprehension complexity of texts on year 5 national tests in English. I believe this understanding offers a perception of what is expected from learners and also helps explain the mismatch that this thesis is trying to understand.

As previously discussed, Farrall says that both length of sentences and complexity within sentences play a significant role in reading comprehension complexity (2012). Based on this investigation through document content analysis, it is clear that the year 5 national test in English has increased in reading comprehension complexity. First of all, both the percentage of exercises containing sentences with more than 10 words and the percentage of exercises using compound and/or complex sentences has increased. Additionally, there is a large percentage of exercises which use other verb forms than present tense. Lastly, as mentioned in discussing the first question, there has also been a removal of exercises testing specific vocabulary.

Therefore, the increase in complexity over the years, specifically between 2012 and 2013, leads to the conclusion that expectations of reading comprehension in a L2 have increased. The specific change as seen in results could be credited to change in competence aims in 2013. As mentioned in chapter II, the category of *communication* was split into two categories, recognizing *oral* communication (focusing on listening and speaking) and *written* communication (focusing on reading and writing). Even more recently, a change in the instructional guide for teachers between 2015 and 2016 shows this as well (UDIR, 2015; UDIR, 2016). In 2015, the instructional manual from UDIR states that year 5 English test content includes reading comprehension, vocabulary and grammar. In 2016, only reading comprehension is mentioned. Currently, the year 5 national test in English is not based on testing skills in listening, speaking or writing in English which is contrary to what UDIR writes in the forward of the English curriculum which says "The subject shall help build up general language proficiency through listening, speaking, reading and writing" (UDIR, 2015, p.55). Reading comprehension relates primarily to only one of four major skills in language learning. I believe by focusing on only one of these areas helps explain the mismatch because teachers who adhere to the competence aims ending in year 4 focus on the other language skills as well. Of course, the four basic language skills are intertwined with each other, but that does not neglect the fact that the national tests seem to focus primarily on reading comprehension.

Relating to information discussed in chapter II, I wonder if the primary focus on reading comprehension on national tests has influenced course books used in the classroom. Let me explain by using *Stairs* as an example. Heidi Håkenstad and Marianne Undheim Vestgård are authors of the first edition, published in 2007. The same authors, as well as Kristin Morten are authors of the second edition, published in 2014 after changes in the 2013 competence aims. In year 4, the second edition offers additional opportunities for reading comprehension to occur. Simply looking at the amount of pages shows this. The first edition of the year 4 learner textbook has 93 pages and the second edition has 180. Similarly the first edition of the year 4 learner workbook has 80 pages and the second edition has 108.

Interestingly, the second edition has added lengthy factual texts as well. The introduction of the teacher's manual, second edition says there is increased focus on reading and reading

comprehension. According to the authors, this is intended to better prepare students for years 5-7 (Håkenstad, Morten & Undheim Vestgård, 2014). I wonder if the greater attention given reading comprehension in the second edition is also (or perhaps because of) the increased attention of reading comprehension on year 5 national tests in English. The second edition of the year 4 learner workbook of *Stairs* contains exercises that are exactly the same as exercises on the national test (see workbook pages 19 and 76). This is pertinent to mention because the ideal is for subjects to be based on competence aims. In the *Stairs* course books, this is undeniably true, but there are indeed implications that the year 5 national test in English has also influenced the second edition of the course book. I am not saying that the increase in focus of reading comprehension on the national tests is the only factor that has influenced changes in the *Stairs* course books, but I am wondering if the focus on reading comprehension in national tests is one of the factors that has influenced this change.

Question #3: How do exercises relate to the four competence aims the tests refer to?

The instructional guide for teachers from 2016 says that the test is based on four specific competence aims (UDIR, 2016). In general, I believe results from this investigation confirm that different interpretations of competence aims can explain a mismatch between the national tests and aims. Let me clarify.

First of all, the *fill-in-the blank* exercises test skills referring to learners' use of pronouns, conjunctions, verbs and/or prepositions. *Fill-in-the-blank* is the only exercise type that intentionally assesses the competence aim where learners "use some common, short words and simple spelling and sentence patterns" (UDIR, 2016). In other exercises, a basic understanding of spelling, words and sentence structure is needed to derive meaning within sentences and texts, but only here is it directly tested. The competence aim is vague and does not specifically refer to grammar relating to pronouns, conjunctions or verbs. In fact, none of the competence aims refer specifically to an understanding of grammar skills at all, which leaves the above mentioned competence aim open to interpretation for test-makers, teachers, book-makers, politicians, etc. Also, as mentioned earlier in the chapter, there are

and have always been exercises about placement of objects, but there are no competence aims that relate specifically to understanding prepositions either.

Topics

To see if exercises correlate with the competence aim referring to "familiar topics" and to see how topics related to "one's needs and feelings, daily life, leisure time and own interests" (UDIR, 2016), I sorted content of exercises into 18 general topics. After completing the coding schedule, I analyzed topics in further detail by seeing how frequently each topic was mentioned throughout the tests. (A spreadsheet of these results is found in appendix 10 on page 75.)

These results clearly show that *family/friends, sports/hobbies* and *animals* are the three topics that are generally mentioned most throughout tests from 2009-2016. These topics relate to the competence aim which refers to reading and understanding texts about "one's needs, feelings, daily life, leisure time and own interests" (UDIR, 2016). In addition, generally these topics also fit with the competence aim relating to understanding main texts about "familiar topics". Therefore, results from the coding schedule and coding manual show that year 5 national tests in English generally refer to topics reflected in the competence aims.

Knowing this, I find it necessary to include two interpretations of my observations as well. First of all, some of the topics included in the coding schedule are recorded even if the topic is not important for understanding a text. For example, exercise 9 from the 2012 test begins with "Ben's uncle Ken is very interested in history..." Here, family is mentioned (and has been recorded as one of the topics for the exercise) but the text is really about a hobby related to history. Knowing that Ken is Ben's uncle does not contribute to understanding the main idea of the text. Furthermore, although feelings are mentioned specifically in one of the four competence aims, there are no exercises that are *only* about feelings. For example a scary story from the 2016 test (exercises 29-33) and an exciting text about a dream from 2012 (exercise 38) definitely include feelings as a part of other topics in these texts, but feelings themselves are not the primary topic. Therefore, even though some of the texts relate to topics that have been recorded on the coding schedule, the topics don't always

necessarily contribute to the main topic which is needed for understanding a text or answering an exercise question.

Secondly, keeping in mind that this investigation looked at content of topics to see how they relate to specific competence aims, a critical finding needs to be discussed. When analyzing topics of specific exercises, there are certain topics found throughout the national tests that I argue have no place in being part of topics represented in a L2 language test for 10-year-olds. (The following are only some examples and are not all inclusive.) Topics related to jobs and businesses as seen in exercise 10 from 2013 about an English map maker is not a part of 10-year-olds' daily lives (other examples are in tests 2010 exercise 18, 2013 exercise 1-2, 2015 exercises 20, 21-22). Jobs and descriptions of businesses are familiar topics about adults' daily lives and are likely familiar topics for L2 reading comprehension tests for adults but not for 10-year-old learners. Topics relating to the body or clothes that have to do with hippies (2010 exercise 17) or hairstyles from the 60s (2014 exercise 22) are not familiar topics to 10-year-olds neither in their L2 nor in their L1. School topics relating to 7th grade (2011 exercise 9), home-economics (2011 exercise 8), chemistry sets (2013 exercise 15), examining organisms through a lens (2015 exercise 21) or studying abroad (2016 exercise 6) are appropriate topics for older learners. Hobbies relating to WWII (2012 exercise 7) and Greek gods (2012 exercise 10) can also be debated if these are hobbies that are familiar to 10-year-olds or part of their leisure time activities. Further, in the topic of sports/hobbies, toys and movies that adults find in popular culture such as he-man figures (2010 exercise 4) or the movie "Titanic" (2014 exercise 33) are not modern popular culture for 10-year-olds anymore. Again, these topics are more appropriate for adult learners who can connect their background knowledge to such texts.

In this criticism of appropriate topics, I see that a teacher's perspective (and a 10-year-old's perspective) is important because not all test-makers are aware of topics that fall under the category of *familiar* topics to 10-year-olds. Not only do these examples show a lack of consistency with the two competence aims being referred to, they also make it difficult for learners to use background knowledge to relate to texts, therefore making reading comprehension difficult before readers even begin reading. Of course the test needs to have difficult questions to test learners at all levels and there is something to be said about using

topics that are not familiar, but the questions on the test should still be compatible with the few competence aims that it proclaims it is based upon. As stated at the beginning of my analysis of topics, there are several questions where specific topics are definitely in line with competence aims relating to daily lives and familiar topics. However, I have chosen examples to show that there are exercises each year that do not adhere to the same competence aims. This criticism is necessary, if nothing else, than to remind test-makers to be aware of 10-year-olds' daily lives, not 13-year olds' nor adults' daily lives. These findings indicate that the exercise texts do not always relate to the competence aims that are analyzed here and can indeed explain the mismatch between national tests and competence aims.

Simple/common/familiar

The tables in appendix 11 starting on page 77 show vocabulary words and/or expressions that I question on year 5 national tests in English. I question if these words and expressions fit into the category of being simple, common or familiar. I argue they are not. Understanding this is only my argument, I included the tables in their entirety which can serve as discussion for other professionals or a starting point for further research. Although the tables show my arguments, there are important conclusions to be made here.

First of all, in some exercises there are only a few words and expressions in question. In other exercises, there are many. Although it can be argued that a few unfamiliar words do not necessarily deter learners from comprehending or drawing conclusions from a text, when there are several unfamiliar words or expressions, comprehension is difficult. Therefore, this should be considered when determining the amount of words that are not simple, common or familiar in each exercise. I believe this is especially important to remember since the tests are meant to compliment the competence aims.

Secondly, some words could be replaced with words that are in fact simpler or more familiar words to 10-year-olds in the L2. For example, the verb *notices* (2009 exercise 13) as in *Mr. Wilson **notices***, could be replaced with *sees* so that the beginning of the sentence could read, *Mr Wilson **sees***. Similarly *wade through* (2012 exercise 6) could be replaced with *walk through, contains* (2013 exercise 10) could be replaced with *has*, *prefers* (2014 exercise 22) could be replaced with *likes*, *chops* (2015 exercise 14) could be replaced with *cuts* and *resemble* (2016 exercise 9) could be replaced with *looks like*. The same can be said about

other word forms as well. Here are a few examples with nouns. *Shelter* (2010 exercise 29, 2016 exercises 29-33) could be replaced with *building* or *house*. *Rodent* (2012 exercise 10) could be replaced with *animal*. *Lyrics* (2015 exercise 30) could be replaced with *music* or *words*. Here are a few examples with adjectives. *Hopeless* (2012 exercise 36) could be replaced with *sad*. Second-hand (2013 exercise 13) could be replaced with old. *Rather distinct* (2014 exercise 10) could be replaced with *very special*. *Exquisite* (2010 exercise 18) could be replaced with *beautiful* or *lovely*. Even using the word *walrus* in the second exercise of the most recent test (2016) is a vocabulary word that is not common to most 10-year-olds. I believe learners know a lot of words relating to animals, but I argue that walrus is not a common animal word that they learn in beginning primary school in their L2. The spelling is different enough between English and Norwegian that it can't necessarily be expected for learners to use their L1 knowledge or transfer skills to understand the word in English.

The fact that these examples (and many others) have words that can be replaced by easier, more common words, show that several exercises within the tests use words that are not in line with (my interpretation of) the competence aims that refer to using *simple*, *common* and *familiar* words and expressions. I believe different interpretations of these words contribute to the mismatch between English competence aims at the end of year 4 and year 5 national tests in English.

As a final reflection about both topics that are inappropriate for 10-year-olds and vocabulary that expects a lot of 10-year-old learners in their L2, I find it necessary to refer to the Common European Framework for Reference of Languages, known as the CEFR. The CEFR is a recognized European document that refers to "what language learners have to learn in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively" (Byram & Parmenter, 2012, p. 2). This framework aids in making policy documents and assessments for L2 learners in Europe in a way that is similar between the countries that use it. I mention the CEFR because the national competence aims and tests are implicitly linked to the CEFR (Fikke & Helness, 2012). Although it can be seen as positive to have an international reference point when it comes to SLA and testing, I must voice one concern. According to the Cambridge University Press (2013) in its *Introductory Guide to the Common European Framework for Reference of*

Languages (CEFR) for Teachers, the CEFR is intended as a tool for language learners who are teenagers or adults. I have not researched the degree to which the CEFR influences the year 5 national test in English, but some of the topics and words used in the test, relate to learners who are older than 10-year-olds and I wonder if there is a correlation here.

I am aware of a project led by Angela Hasselgreen which started in 2007 called *Assessment of young learner literacy linked to the Common European Framework of Reference for Language (*AYLLIT*).* This project has looked at how the CEFR can be used for younger learners in reading and writing. (My understanding of this project is that it has looked at the CEFR goals and *informal* assessment.) My point, however is that both information from the Cambridge University Press and the AYLLIT project shows that the CEFR is not originally intended for 10-year-olds. Without jumping to inappropriate conclusions but at the same trying to explore options around the high expectations in vocabulary use and topics, I wonder to what degree outside international influences have influenced the national tests.

Chapter VI: Conclusion

People working with test development and research most often have positive experiences with school and learning, in this case languages. Obviously, this has a number of good effects, but it does not optimize their ability to understand the affective reactions of students who are not quite as positive (Erickson & Åberg-Bengtsson, 2012, p. 97).

In chapter V, conclusions have been made and discussed about the mismatch between year 5 national tests in English and competence aims ending in year 4. Chapter VI provides final reflections based on those conclusions and some ideas for further research in the area.

Various interpretations

In addition to five basic skills, teachers in Norway are expected to use all 25 competence aims (to be met by the end of year 4) to steer instruction and learning in early primary L2 English education. The year 5 national test in English is currently based on four of these aims and I believe the *interpretations* of these aims can greatly vary. I have tried to show that varying interpretations of these aims help explain the mismatch.

An article about national tests in literacy (as opposed to English or numeracy) explains how a major goal for national tests in general is to see how learners perform in regards to basic skills and competence aims (Frønes, Roe & Vagle, 2012). As stated in the introduction, UDIR portrays this as a primary goal of national tests and I believe educators want to believe this as well. However, after completing my investigation, I have difficulties recognizing that a main goal of year 5 national tests in English is to see how learners perform in regards to basic skills and competence aims. I have only viewed four competence aims, but these are the same aims that serve as a basis for the tests. If there was a common interpretation of the competence aims, perhaps this idea would be easier to digest, but I have shown that interpretations differ. Specific differences in interpretations relate to using *simple*, *common* and *familiar* words and expressions and appropriate topics surrounding daily life for 10-year-olds.

Focus on reading comprehension

In general, year 5 national tests in English are primarily reading comprehension tests which findings from my investigations reveal is even more evident after 2012. Reading is one of four large components of L2 learning, but it is only one component. The fact that instruction

and learning English in Norwegian early primary L2 classrooms is expected to focus on all four areas of language learning (listening, speaking, reading and writing) and the test focuses heavily on one area can also explain the mismatch.

There is something called a washback effect which is the effect tests have on influencing teaching and learning in the classroom (Carlsen, 2007). If concepts or skills are thought to be important by educators because they are assessed on prioritized tests and if educators focus on those concepts or skills in classrooms contexts because they are on the tests, the tests have a washback effect. A washback effect can be positive, negative or a combination of both. If a teacher focuses on a skill that is pertinent in learning that he/she wouldn't otherwise have focused had it not been on tests, the washback effect is positive. However, if a teacher drills on areas because they are in tests and this creates lack of motivation for the subject area, this can have a negative effect on teaching and learning in the classroom (Carlsen, 2007). The common term of "teaching to the test" is an example of a washback effect but generally has a negative connotation. The 2016 instructional guide for teachers for the year 5 national test in English provides several examples about how to use results of the year 5 national test in English to further learning in the classroom (2016) which is also an example of a washback effect.

In the past school year, I have attended meetings with educators from several communities where discussions about national tests furthering learning in the classroom have been the primary focus. This is another concrete example of the washback effect. *What do we know from results of the national test*, *how can our school get better at them* and *how can we make this happen* are questions that have steered these meetings. This thinking concerns me for two reasons. First of all, the national test is a standardized test and learners are assessed based on how they perform in relation to their peers. One of the drawbacks of standardized, norm-referenced tests is that they do not further learning in the classroom (Farrall, 2012). So, by using year 5 national tests in English to further learning in the classroom, it is serving a function that a standardized norm-referenced test is not meant to serve.

Secondly, because year 5 national tests in English focus primarily on reading comprehension, I am concerned that the washback effect could create classrooms that focus primarily on

reading comprehension in early primary L2 education. (Perhaps this is already seen in the course book of *Stairs*.) These are of course important skills to develop, but if schools begin to focus primarily on reading comprehension in early primary L2 classrooms where time allocation in the L2 is limited, schools are not effectively using what is known about L2 learning in early primary education to benefit learning and motivate learners as explained in chapter III. In early L2 learning with a limited amount of time, it is important for massive amounts of language exposure so that understanding is maintained. Comprehending an L2 takes time and repetition. It is also important to create meaningful situations that inspire and motivate early learners in their L2. With young learners this includes a lot of implicit acquisition of language but explicit instruction and learning is also important in order to meet the needs of several types of learners.

If the year 5 national test in English is already showing these signs of washback effects, then I must ask, what is steering the English L2 educational curriculum in early primary schools in Norway today? Is it the national curriculum with its competence aims and basic skills or is it the year 5 national test in English?

Ideas for further research

The above question leads to opportunities for further research in the field. I am not the first person to ask this question of course. For example, Astrid Birgitte Eggen interviewed several school leaders in regards to their perceptions of PISA questionnaires, national curriculum and assessment. She wondered if professionals should be guided by curriculum content or test content (2010). In the March, 2017 edition of *Bedre Skole*, the leader of Union of Education Norway strongly suggests that Norway considers dropping out of PISA because Norwegian education has become too preoccupied with test results (Brøyn, 2017). Several written resources exist which warn against educational systems ruled by results. Further research, then, could be done to inquire about outside international influences such as PISA or the CEFR and how it has trickled down to influence year 5 national tests in English. On a more detailed level, perhaps this type of research could attempt to also answer the question of why the tests have changed throughout the years, especially between 2012 and 2013. My investigation clearly shows that sentence complexity and the amount of words used in the tests have increased, especially after 2012. However, I have not attempted to seek out

possibilities for why this has happened. Knowing the competence aims have only changed slightly between 2012 and 2013, it could be interesting to find out if this alone is adequate reasoning for changes in the test or hypothesize about other factors.

Looking at year 5 national tests in English from a completely different viewpoint, research could be done to explain the learners' viewpoints of the year 5 national test in English. Here, interviews could be done on a sample of learners to find out what they know, which exercises they guess on, etc. Further research could also be done to find out the English vocabulary level of average 10-year-olds from various perspectives to have a greater understanding of what words are simple, common and familiar to 10-year-olds.

A final reflection

The Norwegian government is in the process of renewing the national educational plan with its competence aims (Meld. St. 28 (2015-2016)). Perhaps this thesis comes at the right time in regards to renewing the English subject of the national plan. A plethora of research exists about how to acquire a L2 in addition to knowledgeable information about different approaches that can be used in instructional contexts. I believe this research and knowledge needs to intertwine with organization of the Norwegian system today including time allocated the subject, course books, competence aims and national tests. This is pertinent to keep in mind when making educational reforms and policy so that L2 educational policy does not become steered by tests. The quote at the beginning of this chapter is a vital reminder that learners are a various group of individuals. Those of us who work in any and all parts of education, need to remember this.

References

Andersson, U. (2009). The contribution of working memory capacity to foreign language comprehension in children. *Memory,* 18(4), 458-472

Alexiou, T. (2009). Young Learners' Cognitive Skills and Their Role in Foreign Language Vocabulary Learning in Nikolov, M. (ed): *Early Learning of Modern Foreign Languages: Processes and Outcomes.* Bristol: Multilingual Matters

AYLLIT (2016, January 5). Assessment of young learner literacy linked to the Common European Framework of Reference for Languages. Retrieved from ayllit.ecml.at

Baker, C. (2000). *The Care and Education of Young Bilinguals: An Introduction for Professionals*. Clevedon: Multilingual Matters Ltd.

Barron-Hauwaert, S. (2004). *Language Strategies for Bilingual Families: The One-Parent-One-Language Approach.* Clevedon: Multilingual Matters Ltd.

Bialystok, E. (2013). The Impact of Bilingualism on Language and Literacy Development. In T. Bhatia and W. Ritchie (Ed.), *The Handbook of Bilingualism and Multilingualism Second Edition* (624-648). West Sussex: Blackwell Publishing, Ltd.

Bhatia, T.K. & Ritchie, W.C. (2013). *The Handbook of Bilingualism and Multilingualism Second Edition.* West Sussex: Blackwell Publishing, Ltd.

Bryman, A. (2008). *Social Research Methods third edition.* Oxford: Oxford University Press.

Byram, M. & Parmenter, L. (2012). *Common European Framework of Reference : The Globalisation of Language Education Policy* (1). Clevedon, GB: Multilingual Matters. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Brøym, T. (2017). Meld Norge ut av PISA! *Bedre Skole,* 1*,* 84-85

Cambridge University Press (2013). *Introductory Guide to the Common European Framework of References for Languages (CEFR) for English Language Teachers.* Cambridge

Carlsen, C. (2007). Language testing—a matter of ethics in Carlsen, C & Moe, E. (Eds), *A Human Touch to Language Testing.* Oslo*:* Novus Press.

Eggen, A. (2010). PISA's gyldighet blant skoleledere og lærere in Elstad E., & Sivesind, K. (Eds.) *PISA-sannheten om skolen?* Oslo: Universitetsforlaget.

Ellis, R. (2015). *Understanding Second Language Acquisition second edition.* Oxford: Oxford University Press.

Ellis, N (2011). Implicit and Explicit SLA and their Interface in Sanz, C. & Leow R.P. (ed): *Implicit and Explicit Language Learning: Conditions, Processes, and Knowledge in SLA and Bilingualism.* Washington, DC, USA: Georgetown University Press. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Ethnologue (2016, June 4). Retrieved from www.ethnologue.com/world.

Erickson, G. & Åberg-Begtsson, L. (2012). A Collaborative Approach to National Test Development in Tsagari, D. & Ildiko, C. (Eds.) *Language Testing and Evaluation: Collaboration in Language Testing and Assessment*. Vol. 26. Frankfurt am Main: Peter Lang. Retrieved from http://www.ebrary.com.ezproxy.no

Farrall, M. (2012). *Reading Assessment: Linking Language, Literacy and Cognition.* Sommerset: Wiley.

Fikke, A.J. & Helness, H.L. (2012). Synergies and Tensions in Computerized Language Testing in Tsagari, D. & Ildiko, C. (Eds.) *Language Testing and Evaluation: Collaboration in Language Testing and Assessment.* Vol. 26 (159-170). Frankfurt am Main: Peter Lang. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Frønes, T., Roe, A., & Vagle, W. (2012). Nasjonale prøver I lesing-utvikling, resultater og bruk in Hopfenbeck, T. Kjærnsli, M. & Olsen R. (Eds.). *Kvalitet i Norsk Skole: Internasjonale og nasjonale undersøkelser av læringsutbytte og undervisning.* Oslo: Universitetsforlaget

Gadamer, H-G. (2004). *Truth and Method* (second edition)*.* New York: The Continuum International Publishing Group Inc.

Ganem-Gutierrez, G. (2013). Sociocultural Theory and second language development: Theoretical foundations and insights from research in García, M. & Martínez, A. (Eds.). *Contemporary Approaches to Second Language Acquisition*. Amsterdam: John Benjamins Publishing Company. Retrieved from http://www.ebrary.com.ezproxy.hil.no

García, M. M. D. P., & Martínez, A. M. (2013). Contemporary Approaches to Second Language Acquisition. Amsterdam, NLD: John Benjamins Publishing Company. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Glossary of Education Reform (2015, November 10). Assessment Retrieved from http://edglossary.org/assessment

Hager, M. (2011). *Intercultural Studies and Foreign Language Learning : Culture, Psychology, and Language Learning.* Bruxelles, GB: Peter Lang AG, Internationaler Verlag der Wissenschaften. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Holsti, O. (1969). *Content Analysis for the Social Sciences and Humanities.* Reading, Massachusetts: Addison-Wesley Publishing Company.

Hummel, K. (2014). *Introducing Second Language Acquisition Perspectives and Practices.* West Sussex: John Wiley & Sons Inc.

Håkenstad, H. & Undheim Vestgård, M. (2007). *Stairs My Textbook 4.* Oslo: J.W. Cappelens Forlag AS.

Håkenstad, H. & Undheim Vestgård, M. (2007). *Stairs My Workbook 4.* Oslo: J.W. Cappelens Forlag AS.

Håkenstad, H. & Undheim Vestgård, M. (2007). *Stairs Teacher's Book 4.* Oslo: J.W. Cappelens Forlag AS.

Håkenstad, H., Morten, K., & Undheim Vestgård, M. (2014). *Stairs My Textbook 4.* Oslo: Cappelen Damm AS.

Håkenstad, H., Morten, K., & Undheim Vestgård, M. (2014). *Stairs My Workbook 4.* Oslo: Cappelen Damm AS.

Håkenstad, H., Morten, K., & Undheim Vestgård, M. (2014). *Stairs Teacher's Book 4.* Oslo: Cappelen Damm AS.

Isaksen, T.H., Hjelm Solli, A-M. (2014). *Resultaten er kommet! Hva nå?: Nasjonale prøver— fra focus på resultat til utvikling av kvalitet.* Master thesis, Norges Artiske Universitet, Tromsø.

Long, M. (2014). *Second Language Acquisition and Task-Based Language Teaching.* Somerset: Wiley-Blackwell. Retrieved from ProQuest ebrary.

Johnson, M. (2003). *Philosophy of Second Language Acquisition.* New Haven: Yale University Press. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Krashen, S. (2008). *Language Education: Past. Present and Future.* In RELC Journal. Vol 39, Issue 2. retrieved from http://journals.sagepub.com.ezproxy.hil.no

Krashen, S. & Terrell, T. (1995). *The Natural Approach: Language Acquisition in the Classroom.* Hertfordshire: Prentice Hall Europe.

Kunnskapsdepartementet (2014). Læreløftet: På lag for kunnskapsskolen. Aurskog: Kunnskapsdepartement.

Kvernbekk, T. (2005). *Pedagogisk Teoridannelse, insidere, teoriformer og praksis.* Bergen: Fagbokforlaget.

Marsdal, M. (2011). *Kunnskapsbløffen: skoler som jukser, barn som gruer seg.* Oslo: Manifest.

Meld. St. 28 (2015-2016). *Fag-Fordypning-Forståelse en fornyelse av Kunnskapsløftet.* Oslo: Det Kongelige Kunnskapsdepartement

Mellanby, J. & Theobald, K. (2014). *Education and Learning : An Evidence-based Approach.* Somerset: Wiley. Retrieved from ProQuest ebrary.

Ministry for Children Education and Gender Equality (2016, June 14). Subjects & Curriculum. Retrieved from http://eng.uvm.dk/education

Munoz, C. (2006). *Age and the Rate of Foreign Language Learning.* Clevedon: Multilingual Matters LTD.

Nes, H. (2013). *Nasjonale prøver på mellomtrinnet: Lærernes elevforberedelser i et accountability-perspektiv.* Master thesis, Universitetet i Oslo, Oslo.

Olaug Berger, M. (2007). *Nasjonale prøver 2007: Arbeidet skolene etter intensjonen? Hvordan takler skolene preset?* Master thesis, Universitetet i Oslo, Oslo.

Reber, A. (2011). An Epitaph for Grammar: An Abridged History in Sanz, C. & Leow R.P. (ed): *Implicit and Explicit Language Learning : Conditions, Processes, and Knowledge in SLA and Bilingualism.* Washington, DC: Georgetown University Press. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Seland, I., Vibe, B. & Hovdhaugen, E. (2013). *Evaluering av nasjonale prøver som system* (NIFU rapport 4/2013) Oslo: Utdanningsdirektoratet

Skoleverket. (2016, June 15). Timeplan for Grundskolan. Retrieved from http://www.skoleverket.se

Skoleverket. (2015, September 16). Prøv & bedømning. Retrieved from http://www.skoleverket.se

Tomasello, M. (2010). Language Development in Goshwami, U. (Ed).*Wiley Blackwell Handbooks of Developmental Psychology: Wiley-Blackwell Handbook of Childhood Cognitive Development (2)*. Hoboken: Wiley-Blackwell

Utdanningsdirektoratet. (2016, March 15). *Analyse av nasjonal prøver i engelsk, lesing or regning på 5. trinn 2015*. Retrieved from http://www.udir.no/Vurdering/Nasjonale-prover

Utdaningsdirektoratet (2015). *Læreplanverket for Kunnskapsløftet Grunnskolen.* Oslo: Pedlex

Utdaninngsdirektoratet (2006). *Midlertidig utgave av læreplanverket for Kunnskapsløftet.* Oslo: Interface Media AS

Utdanningsdirektoratet (2009). *Nasjonale prøver i engelsk 5. trinn 2009.*

Utdanningsdirektoratet (2010). *Nasjonale prøver i engeslk 5. trinn 2010.*

Utdanningsdirektoratet (2011). *Nasjonale prøver i engelsk 5. trinn 2011.*

Utdanningsdirektoratet (2012). *Nasjonale prøver i engelsk 5. trinn 2012.*

Utdanningsdirektoratet (2013). *Nasjonale prøver i engelsk 5. trinn 2013.*

Utdanningsdirektoratet (2014). *Nasjonale prøver i engelsk 5. trinn 2014.*

Utdanningsdirektoratet (2015). *Nasjonale prøver i engelsk 5. trinn 2015.*

Utdanningsdirektoratet (2016). *Nasjonale prøver i engelsk 5. trinn 2016.*

Utdanningsforbundet (2016). Nøkkeltall for grunnskole t.o.m. skoleåret 2015-2016: Faktaark 2016:1. Retrieved from www.utdanningsforbundet.no on 15.01.16

Utdanningsdirektoratet (2012). *Framework for Basic Skills.* Retrieved from www.udir.no on 10.10.16

Utdanningsdirektoratet (2016, April 19). Metodegrunnlag for Nasjonale prøver. Utdanningsdirektoratet. Retrieved from www.udir.no

Utdanningsdirektoratet. (2017, March 1). Rammeverk for nasjonale prøver. Retrieved from https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-nasjonale-prover

Utdanningsdirektoratet. (2015). Rettleiing til lærarar: Oppfølging av vidare arbeid med prøve på 5. trinn: Nasjonal prøve i engelsk.

Utdanningsdirektoratet. (2016). Rettleiing til lærarar: Oppfølging av vidare arbeid med prøve på 5. trinn: Nasjonal prøve i engelsk. Retrieved from www.udir.no on 1.12.16

Utddanningsdirektoratet. (2014, July 17). Vurdering av Nasjonale prøver engelsk. Retrieved from http://www.udir.no/Vurdering/Nasjonale-prover/Engelsk/Engelsk/

Van Patten, B. (2011). Stubborn Syntax: How it Resists Explicit Teaching and Learning in Sanz, C. & Leow R.P. (ed): *Implicit and Explicit Language Learning: Conditions, Processes, and Knowledge in SLA and Bilingualism.* Washington, DC: Georgetown University Press. Retrieved from http://www.ebrary.com.ezproxy.hil.no

Wang, J., Odell, S.J., Klecka, C.L., Spalding, E., & Lin, E. (2010). Understanding Teacher Education Reform. *Journal of Teacher Education 61(5)395-402.*

Østrem, S. (2009). The Contribution of teacher education in teachers' knowledge development. *Nordic Studies in Education*, 61, 369-383.

Appendix 1: The Framework for Basic Skills-reading rubric

The following is the rubric for reading skills as defined in *The Framework for basic skills* document from 2012.

The reading rubric incorporates reading levels throughout the entire primary and secondary educational system.

Sub-category	Level 1	Level 2	Level 3	Level 4	Level 5
Understand	Can use previous knowledge to get an overview of the text prior to reading. Can read simple texts on paper and screen and process the text with some help.	Can read simple texts with fluency and perseverance. Can find and read texts on the internet. Can ask questions and talk about content.	Can vary the reading text type and purpose. Can read and navigate effectively on the internet. Can use different methods to structure content.	Can apply subject-related and general knowledge actively. Can read screen and paper texts critically. Can interpret texts in an independent manner.	Can choose and use reading strategies relevant to a wide variety of text types and purpose. Can assess one's own reading and reflect on the strategies applied.
Find	Can recognize explicitly expressed information centrally placed in simple texts.	Can identify central subject-related texts. Can identify explicitly expressed information in texts with competing information. Can refer to sources.	Can obtain and combine information in texts with competing information. Can assess source credibility.	Can identify implicit information in complex texts. Can use sources critically and refer to them in a systematic manner.	Can obtain detailed and implicit information in texts without prior knowledge of text type and content.
Interpret	Can draw simple conclusions based on information in texts and can use one's own words to express them.	Can identify the main topic and understand clear connections explicitly expressed.	Can infer and understand information implicitly expressed.	Can deal with ambiguity. Can identify contradictory information and deal with information that does not correspond to expectations.	Can show holistic as well as detailed comprehension of complex texts. Can systematize and draw conclusions based on implicit information.
Reflect and assess	Can comment on content and meaning in simple texts.	Can assess the content of subject-related texts.	Can give substantiated assessment of form and content in subject-related texts.	Can crucially analyze and assess form and content. Can compare and systematize information in different subject-related texts.	Can assess complex texts about unfamiliar topics in a critical manner and incorporate subject-related general perspectives.

Appendix 2: The Competence Aims after Year 4

The following is a complete list of competence aims after year 4, as described by UDIR in their English

version of the curriculum. (UDIR, 2015)

Language Learning: The aims of the studies are to enable learners to

- identify situations where it might be useful or necessary to have English language skills
- converse about one's own work in learning English
- find similarities between words and expressions in English and his/her own native language
- use digital resources and other aids in exploring the language

Oral Communication: The aims of the studies are to enable learners to

- use simple listening and speaking strategies
- listen to and understand the meaning of words and expressions based on the context they are used in
- understand and use English words, expressions and sentence patterns related to one's needs and feelings, daily life, leisure time and own interests
- understand the main content of nursery rhymes, word games, songs, fairy tales and stories
- use some polite expressions and simple phrases to obtain help in understanding and being understood
- participate in everyday conversations and simple phrases related to local surroundings and own experiences
- be able to repeat the English alphabet and spell names and home town
- understand and use English words and expressions related to prices, quantities, shape and size when communicating about one's daily life, leisure time and own interests

Written Communication: The aims of the studies are to enable learners to

- use simple reading and writing strategies
- understand the relation between English phonemes and letters and put sounds together to form words
- understand the meaning of words and expressions based on the context they are used in
- read, understand and write English words and expressions related to one's needs and feelings, daily life, leisure time and own interests
- understand the main content of simple texts and familiar topics
- write short texts that express opinions and interests, and that describe, narrate and enquire
- use some common short words and simple spelling and sentence patterns
- use digital tools to retrieve information and experiment in creating texts

Culture, Society and Literature: The aims of the studies are to enable learners to

- give some examples of English-speaking countries and famous people from these countries
- converse about some aspects of different ways of living, traditions and customs in English-speaking countries and in Norway
- participate in presenting nursery rhymes, word games, songs, short plays and stories in English
- express own thoughts and opinions in the encounter with English-language literature and child culture
- create own texts inspired by English-language literature and child culture

Appendix 3: Time allocation

The diagram below is taken from UDIR's website and shows the allocation across subjects and year groups for the 2015-2016 school year. Reference: www.udir.no. (This can also be found in the newest version of the KL06 (UDIR, 2015).

2.2 Ordinær fag- og timefordeling

Fag- og timefordelingen i tabell 1 gjelder for alle som ikke skal ha spesiell fag- og timefordeling (jf. punkt 2.3 og 2.4).

Tabell 1 Ordinær fag- og timefordeling for elever på 1.-10.trinn skoleåret 2015-2016

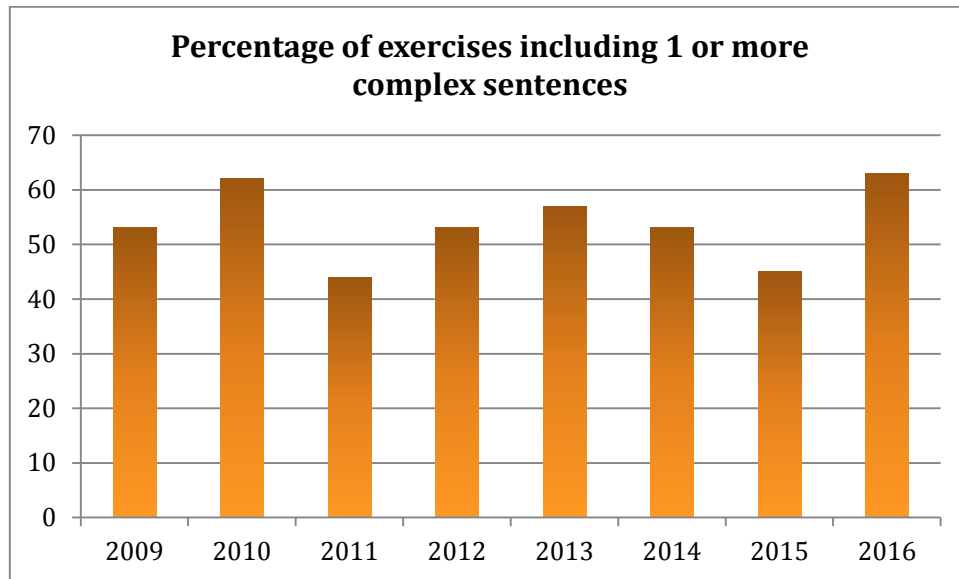| Fag/trinn	1.-4. trinn	5.-7. trinn	Sum	8. -10. trinn	Sum grunnskole
KRLE	-	-	427	153	580
Norsk	931	441	1372	398	1770
Matematikk	560	328	888	313	1201
Naturfag	-	-	328	249	577
Engelsk	138	228	366	222	588
Fremmedspråk / fordypning / arbeidslivsfag	-	-	0	222	222
Samfunnsfag	-	-	385	249	634
Kunst og håndverk	-	-	477	146	623
Musikk	-	-	285	83	368
Mat og helse	-	-	114	83	197
Kroppsøving	-	-	478	223	701
Valgfag	-	-	0	171	171
Utdanningsvalg	-	-	0	110	110
Elevrådsarbeid	-	-		-	-
Fleksibel time	-	-	38	0	38
Fysisk aktivitet	0	76	76	0	76
Samlet minstetimetall	5 234			2 622	7 856

Appendix 4: Levels of reading comprehension

The following is from the 2016 teacher's guide about the year 5 national test in English. It identifies which of the three areas of reading comprehension that each exercise tests. (UDIR, 2016).

Oversikt over oppgåvene til nasjonal prøve 2016

Opp. nr.	Innhald	Teksttype	Formål med oppgåva:	Nivå av leseforståing:	Format
1	Fritid og interesser: museum	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå setning(ar)	Click item
2	Fritid og interesser: museum	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå vokabular	Click item
3	Daglegliv: veret, vekedagar	Forteljing/bilete av værmelding	Finne informasjon/ forstå detaljar	Kople setningar/ delar av tekstar	Click item
4	Fritid og interesser: avis	Forteljing/bilete av avisforside	Finne informasjon/ forstå detaljar	Kople setningar/ delar av tekstar	Click item
5	Daglegliv: klede	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå vokabular	Click picture
6	Fritid og interesser: reise	Forteljing/ bilete	Finne informasjon/ forstå detaljar	Forstå vokabular	Click picture
7	Fritid og interesser: reise	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå vokabular	Click picture
8	Fritid og interesser: reise	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå vokabular	Click picture
9	Fritid og interesser: fuglar	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå setning(ar)	Click picture
10	Høflegheits-fraser	Teikneserie	Finne informasjon/ forstå detaljar	Forstå vokabular	Click and drag
11	Daglegliv: bustad	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forståsetning(ar)	Click and drag
12	Daglegliv: bustad	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forståsetning(ar)	Click and drag
13	Fritid og interesser: dyr på garden	Forteljing/bilete	Finne informasjon/ forstå detaljar	Forstå setning(ar)	Click and drag

Appendix 5: Coding Manual

Exercise type	Topic of exercise
1. Click on the correct item 2. Placement questions (relating to prepositions) 3. Read text, find correct item, put in correct place 4. Colour 5. Set the time 6. Calendar 7. Find the correct notice 8. Read text and click on the correct picture 9. Read several texts and match with 1 picture 10. Read several texts and answer 1 question (question comes before the text) 11. Fill-in-the-blank (followed by number of blanks) 12. Read and answer a question about a text (question comes after the text) 13. Read and answer different questions about a text (same text for several questions) 14. Who-could-say 15. Read a chart/table	1. business 2. family/friends 3. school 4. animals 5. time 6. clothes 7. jobs 8. vehicles/transportation 9. body 10. food 11. sports/hobbies 12. house/buildings/rooms 13. weather 14. calendar 15. celebrations 16. countries/culture of a country 17. feelings 18. other
Sentence structure:	**Sentence length:**
1. Simple sentence 2. Compound sentence 3. Complex sentence	1. 10 words or less 2. 11-20 words 3. More than 20 words
Verb Use:	
1. Present tense (including simple present, present continuous, and commands) 2. Past tense 3. Present perfect 4. Past perfect 5. Conditional (should, could, would) 6. Future	

Appendix 6: Bar chart 1

Percentage of exercises including 1 or more complex sentences

Year	Percentage
2009	53
2010	62
2011	44
2012	53
2013	57
2014	53
2015	45
2016	63

The diagram shows the percentages of exercises that include one or more complex sentence, which according to Farrall (2012) are the most difficult sentence structure to comprehend out of simple, compound and complex sentences. (Note that the vertical axis ranges from 0-70%.)

The amount of exercises using one or more complex sentence varies from 45-63%, showing that about half of the exercises from 2009-2016 consistently include at least one complex sentence.

Appendix 7: Bar chart 2
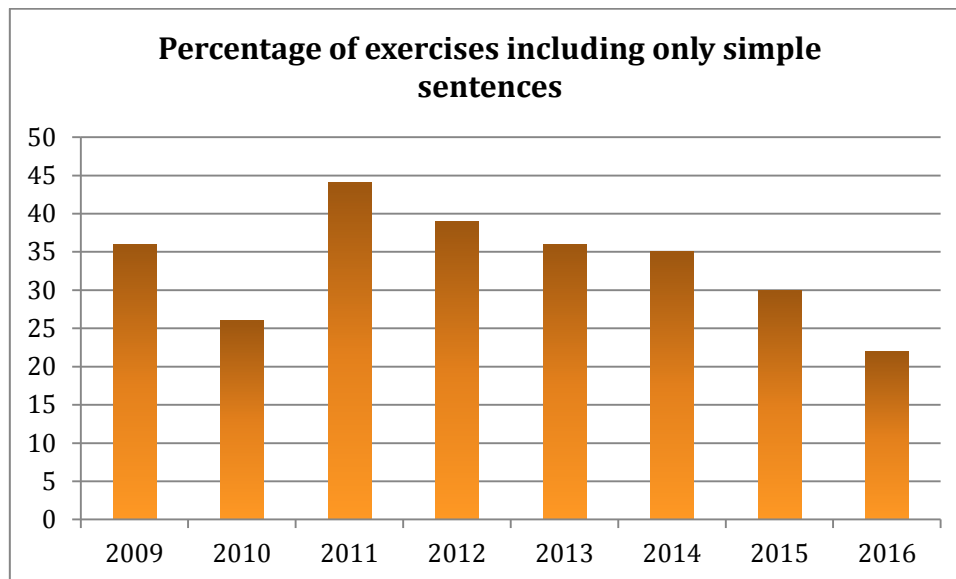
The diagram shows the percentage of exercises in each test that include only simple sentences. (Note that the vertical axis ranges from 0-50%.)

Results show that from 2009-2016, a range from 22% to 44% of the exercises on each test include only simple sentences. There are three exercise types that never use only simple sentences in any of the tests. These exercise types are: *read and answer a question about a text, read and answer different questions about a text* and *who-could-say* (exercise types 12, 13 and 14 on the coding manual). These three exercise types consistently use compound and/or complex sentences as well. Another observation relating to the exercise type is that *set the time exercises* account for 2-4 exercises in each test from and including 2011 and always include only simple sentences. *Set the time* exercises were not included on the tests in 2009 or 2010. This is interesting for interpreting details in the results. For example in 2016, 22% of the exercises included only simple sentences, where 5% of these exercises were *set the time* exercises. Regardless of keeping this in mind or not, the diagram shows evidence that the amount of exercises using only simple sentences has been decreasing.

Appendix 8: Bar chart 3

Percentage of exercises containing sentences with only 10 words or less
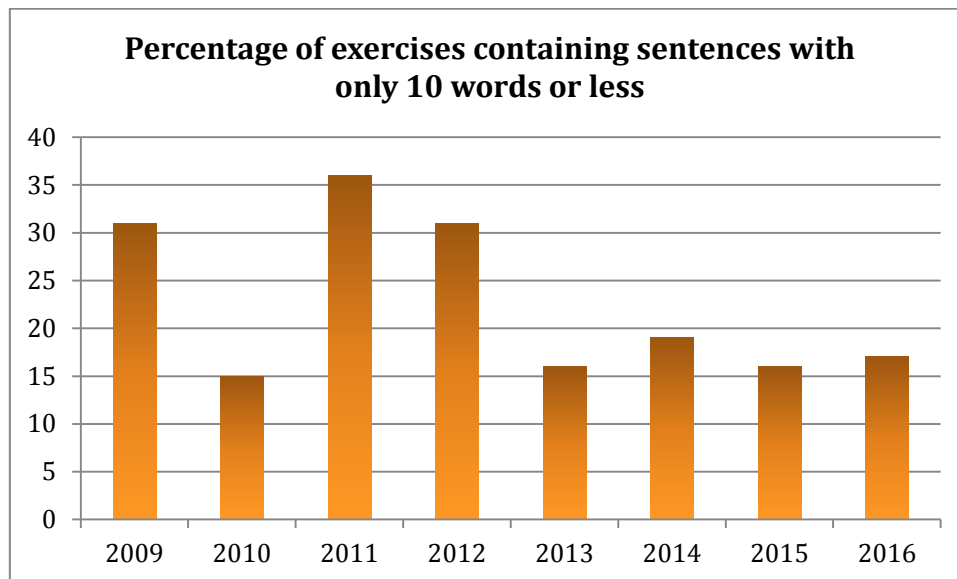
This diagram shows the percentage of exercises containing sentences with 10 words or less, and contributes to conclusions made about sentence length. (Note that the vertical axis ranges from 0-40%.)

Results show that from 2009-2016, a range from 15% to 36% of the exercises in each test include sentences that only use 10 words or less. With the exception of 2010, the tests preceding 2013 had a greater amount of exercises that only included short sentences. From 2013 and onwards, this percentage dropped significantly, by more than 10%. The most common exercise types including sentence length of 20 words or more include: *fill-in-the-blank, read and answer a question about a text, read and answer different questions about a text*, and *who-could-say* (11, 12, 13 and 14 on the coding manual).

Appendix 9: Bar chart 4

Percentage of exercises with only present tense verbs

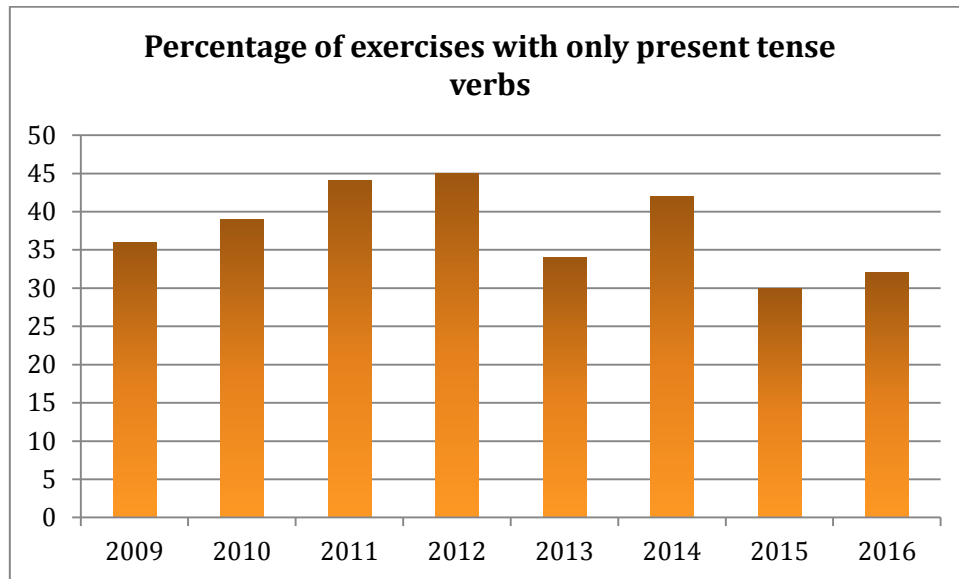This diagram shows the percentage of exercises including only present tense verbs. (Note that the vertical axis ranges from 0-50%).

Results show that from 2009-2016, a range from 30% to 45% of the exercises in each test include only present tense verbs in the texts. Tests from 2009-2012 averaged around 40% and after 2013, this average is about 30%. This percentage dropped significantly, by about 10%. With the exception of 2014, tests from 2013 onwards include fewer exercises containing verbs only in the present tense.

Appendix 10: Topics

Topics: sorted by topics used most to topics used least

		2009 (29)	2010 (31)	2011 (27)	2012 (40)	2013 (33)	2014 (32)	2015 (32)	2016 (30)		total
family/friends	2	12	13	7	8	12	7	13	9		81
sports/hobbies	11	11	11	4	11	9	4	11	10		71
Animals	4	9	11	10	8	4	5	3	8		58
house/buildings/rooms	12	5	4	2	7	5	7	5	10		45
Clothes	6	2	4	6	4	4	7	2	3		32
School	3	4	1	2	3	5	2	3	6		26
Food	10	4	2	4	4	7	3	2	0		26
countries/cultures	16	4	3	0	3	4	5	3	4		26
Weather	13	3	5	1	7	2	2	1	3		24
Time	5	1	0	3	4	4	4	4	2		22
Feelings	17	3	3	5	2	3	1	3	1		21
Celebrations	15	0	1	2	2	3	2	3	0		13
jobs/business	7, 1	0	1	0	0	5	1	4	2		13
vehicles/transportation	8	1	1	3	2	2	1	1	1		12
Body	9	0	2	1	0	0	3	1	2		9
Other	18	0	2	0	1	0	0	2	0		5
Calendar	14	0	0	0	0	0	1	2	0		3

The first two rows are the categories of topics chosen and the corresponding number on the coding schedule and coding manual. The table is organized by topics mentioned most to topics mentioned least in the test, using the total as a filter. The first column corresponds to particular tests where the number in parentheses represents the number of actual texts. (Remember that there are more exercises than texts). The numbers under each test column reflects the amount of times that a topic is mentioned. I used different colours to easily differentiate between the most mentioned topics in each test. The topic mentioned the most is in blue. The second most mentioned topic is in red, followed by green and purple. If topics are mentioned equally, they are coloured the same colour. For example, in 2014, the top three topics are all coloured blue because these topics are mentioned the same amount of times throughout the test. In addition, topics that are not a part of a particular test are coloured in pink. Some exercises include more than one topic.

The topic of *sports/hobbies* is very general. Hobbies include music, travel, books, map-making, museums, interest in history, etc. This particular category could be divided into

several sub-categories to be analyzed in further detail about which hobbies relate to 10-year-olds. The topic of *house/buildings/rooms* is also general. This refers to explaining both the inside and outside of a house or building. Bedrooms, kitchens, living rooms and bathrooms are all included in this topic, as well as objects that might be found in these rooms. In contrast, businesses and jobs are given separate attention. Originally, I differentiated between exercises mentioning a type of business and exercises about occupations. Upon reflection, I believe these should have been combined into one general topic. They should have been combined just like the two topics mentioned above are combined. This would have taken a lot of work to reorganize the coding schedule which is why I have not made these changes in the coding schedule. However, when analyzing the results, I did put them together. Lastly, there were several sentences throughout the test which said that something was liked by a person. Because this was prevalent in a plethora of texts, I did not include liking something in the category of *feelings*. I did, however include hunger, being scared and showing excitement. I thought *feelings* was an important category to include because one of the competence aims directly refers to feelings in its description.

Appendix 11: Simple/Common/Familiar

2009

Exercise #	Exercise type	Vocabulary words & expressions in question
8	3	Seldom
13	8	Notices (Mr. Wilson never notices)
19	10	Environment, gossiping, cafeteria,
20	10	Reward
21	9	Comfortable, collection of
27	12	Since then, experts
32-36	14	Branch, middle of the night, came rushing to the rescue

2010

Exercise #	Exercise type	Vocabulary words & expressions in question
1	1	Wooly
4	1	Nobody else is allowed to touch
5	4	Luckily, plenty of, join me
7	4	to gather
8	3	Covered with, newly fallen, impossible, fetching, grabbed
9	3	Searching
10	8	Species, shoulder height
13	8	Local factory, produce, run the machines, local council, concerned, released
14	8	Especially
16	8	Knotting my tie loosely
17	9	Old-fashioned, particularly, dyed, string tie, begged, hippie, track suit
18	10	Exquisite, front locker, public transportation, violent winds, sinuses were clogged, disgusting, traffic, environmentally friendly vehicle, contributing, climate change is problematic, man-made, considering, triathlon winner, in my thirties, admit, travel at the pace, impressive, smog, coughing fit
19	9	Exchanged
21	7	Forbidden, unsafe, not allowed
23	9	Coming up with, unexpected, impressed, managed to keep his balance, seaside resort, wind and wave conditions, absolutely, luckily, situated on, equipment, satisfied, glides through, gathered, all kinds of, competitors
24	11	Nearby, hesitates, honest, reward
27	12	Attacked by, has had to manage, seldom, for instance, owner, balances nicely, riding along
29	12	Pond, dried up, find themselves, well (deep well), since, shelter, such depth
30-33	13	Emptied the contents, hardly wait, expression, poison, dawned on them
34-39	14	Keep an eye on, treat

2011

Exercise #	Exercise type	Vocabulary words & expressions in question
8	8	Home economics, vitamins, decided
9	8	Invited, join them, activity
12	9	Car crash, though
13	9	Rounded, double doors, beige, straw roof, window frames, surrounding the building, tower, shutters
15	4	Envy
16	4	Other end of the row
19	3	Decides to distract him
23	11	Envy
24	11	Secret admirer
25	12	Pasture
26-27	12	Wool, chewed, to pieces
28-31	13	hardly move, audience, nothing mattered
32-36	13	Accidents, wreck, twisted ankle, warned, full speed, blocked, complaining, weren't serious

2012

Exercise #	Exercise type	Vocabulary words & expressions in question
3	1	Footstool
4	1	Especially, comfortable
5	1	Afloat
6	8	Decided to, white blanket (referring to snow), wade through
7	8	Particularly, local film club, historical films
8	8	Most easily recognizable, travelling in a herd, never been tamed
9	8	Collection, design
10	8	Ancient Greek gods, caused plagues, rodent, companion, symbolize, laurel leaf crown
13	3	Seldom seen
15	3	Wondered how
24	4	Totally lost in it (referring to a book), supposed to be
27	9	Far end of, noticed, flock of swallows, heading south, desperate, just a matter of, rushed into, amazing sight, multi-coloured arch, recreate on canvas, welcomed the new day, branches
28	9	Comfortable chair, collection of
29	9	Rounded, double doors, beige, straw roof, window frames, surrounding the building, tower, shutters
33	11	Redecorating
34	12	Experts
35	12	Rushing to
36	12	Vineyard, hanging, stretched, tasty, hopeless, no matter, wore himself out, surely, served on a silver platter
37	12	Including, transferred, bendy
38	12	Took off, crawled back under
39-43	13	Might even, hedgehogs, pollute, suggest

2013

Exercise #	Exercise type	Vocabulary words & expressions in question
1	1	Customers, expects lots of business
2	1	Recently
8	8	Primary school librarian, poppy (flower)
9	8	In search of food
10	8	Map maker, generally, contains, complete
11	2	Chemistry set
13	3	Second-hand tractor, unfortunately, arrived
15	3	Home economics, responsible, peeling off
21-22	9	Unusually, hind (legs), divided into several parts, feeds on nectar, powerful flier, around the edges, lives off, rotting fruits, cadavers and dung, trim
23-24	9	Serves him, serving, roast, carving knife
25	10	Advertised, cross between, expected, great demand, town landmark, damaged, pulled down, site, announced, ice slope, documentary, claimed, produce electric energy, swim upstream, motion, generate enough electricity
26	11	covered in, latest forecast predicts
27	11	Allergic to
28	11	Racing down the court
29	12	Sink (verb), for up to
30	12	Decided to, teamed up, protect themselves from attacks, rushing out
31	12	Collected, discovered, tame, set up
32	12	Eight-film series, showed an interest in, role on screen, novel, stunts, curious
33	12	Logs (journal), cartoonist, struggles, series
34-38	13	Pranks, realized, switched, decided, snuck, notice, quite a sight
39-45	14	Street music, odd jobs, managed, donate

2014

Exercise #	Exercise type	Vocabulary words & expressions in question
4	1	Realistic
6	6	Keep up to date, saving up
8	8	Sweet tooth
9	8	All-time favourite
10	8	Rather distinct look, specially trimmed
11	8	Has a thing about, beret
16	3	Suspects
21	9	Rough weather, increase, towards, reach a high, western regions, experience, temperatures will drop
22	9	Pattern, prefers, swept upwards
23-24	9	Doesn't mind, pocket money
25	10	Gnawing animal, colonies, mate for life, prepare, crawl into position to nurse, social groups, consist of, various sounds, gestures, mammals, threatens, fluid
26	10	Centuries, in fact, prefer, require, average, water quality, properly cared for, regularly clean, depending on
27	11	Imitating, in addition to
28	11	Celebrating, normally
29	11	Male, female, average, cubs, normally, prey

33	12	set sail, currency, constructed, unsinkable, costing roughly
34-37	13	deadly, creatures, tentacles, toxic, logs, reptile, inland, venom, mentioned
38-43	14	Hold world records, hardly, previous, parachute jump trainers, competed, participated, competitors, managed

2015

Exercise #	Exercise type	Vocabulary words & expressions in question
1	8	Judges, track
2	8	Equipment, eyesight
3	8	Performance, audience, feather
7	8	Sloping
8	8	Decided to, packed up
12	3	Decorated
13	3	Annoyed
14	3	Chop the ingredients
19	10	Surface, wind pipe, circulates, rapidly, speed at which something travels
20	9	Teaching staff, in charge of, pastries, customers, tidy, suggests
21	9	Not going to rush, small organisms, lens, in order to, chain, firewood
22-23	9	Patients, recovering from surgery, loaves of bread, set the dough, doesn't mind, famer's union, harbor, sets his nets, good catch
24	11	Among the branches, wonders, decides to
25	11	Fictional superhero, according to, survive, destruction, adopted
30	12	Viral, ordinary, lyrics, insists
31-32	12	Get in touch, statement, contacted
33	12	Burglar, unlock, rushes, keep you company
34-38	13	Comes alive, gets light, Christmas classic
39-44	14	Game parks, lava fields, glowing rocks, peak, tipped, remote beaches, raft

2016

Exercise #	Exercise type	Vocabulary words & expressions in question
2	1	Walrus
4	8	Definitely, take advantage of, offer (sale offer) offers (tilbyr)
8	8	Structure, consists, passenger cabin
9	8	Resemble
13	3	Nearby
17	9	Redecorating, luckily, good mood, cuddling
18	9	Front crawl, improves, racket sports
19	9	Waterfront, skyscrapers, entire, cascading
20	9	Architect, designed, beside, canals, noticed, particularly, unusual
22	11	Exhausted, rescued
27	12	Flea market, spots (sees)
29-33	13	Cliffs, bucketing down, shelter, deserted, ran for cover, raced inside, enormous, creaked, wimp, whistling, whirling, rolled down
34-39	14	Ensemble, keen, considering
40	11	Supposed to, notice
41	11	Partly, gliding across