Development of morphological awareness

A longitudinal study of the development of morphological awareness across morphological domains in Norwegian children from kindergarten to first grade

Ane Roalkvam Skåra

Master Thesis in Special Needs Education

Institute of Special Need Education

Faculty of Educational Sciences

UNIVERSITY OF OSLO

Date: 12.06.2019

Development of morphological awareness

A longitudinal study of the development of morphological awareness across morphological domains in Norwegian children from kindergarten to first grade.

2019

Development of morphological awareness: A longitudinal study of the development of morphological awareness across morphological domains in Norwegian children from kindergarten to first grade

Ane Roalkvam Skåra

http://www.duo.uio.no

Abstract

Background and rationale

Previous research indicates that there is a difference in when children acquire awareness of the different morphological domains. For instance, studies across different languages have found that the awareness of inflectional morphology precedes that of derivational morphology (Kuo & Anderson, 2006). It has also been hypothesised that linguistic awareness is a skill that develops from a more implicit to an explicit level. The level preceding an explicit, or "meta", awareness, is termed "epilinguistic control" by Gombert (1992). On this level the child is thought to be able to judge correct vs. incorrect language use. On the meta-level, on the other hand, the child is expected to be able to manipulate linguistic structures in an unfamiliar context. It is further believed that these levels exist for morphological awareness and may be measured by targeting the skills related to the different levels. Such a study has not yet been performed in Norwegian, and this study aims to give insight in how the domains of morphology and levels of awareness interact with age. In this longitudinal study, different domains and levels of morphological awareness have been measured in kindergarten and first grade. The research topic for this thesis is:

A longitudinal study of the development of morphological awareness across morphological domains in Norwegian children from kindergarten to first grade

As morphological awareness cannot be observed directly, the measurement of this skill may be affected by other skills. In this study the tests that have been used are oral and may thus put demands on working memory. The level of vocabulary may affect the measurement of morphological awareness because in one of the tasks the children are expected to manipulate real words, and familiarity with the words may facilitate this task (Kuo & Anderson, 2006). Moreover, a large vocabulary may indicate generally higher linguistic abilities, which in turn may affect the development of morphological awareness, as this too is a linguistic skill. On the basis of these skills possibly affecting the general results on the tests of morphological awareness and the development of morphological awareness, they have been included as covariates in this study. The following research question will therefore also be addressed:

Do working memory and vocabulary skills affect the development of morphological awareness differently across morphological domain and linguistic level?

Method

This study is based on data from the longitudinal study "Development of Numeracy and Literacy in children" (NumLit). Since this research project has gone through two measurement periods, I have had access to data from both kindergarten and first grade. The independent variables in this study consist of three tests of morphological awareness, developed for the purpose of being used in NumLit by Vassiliki Diamanti and Germán García Grande (as reported in the master thesis of García Grande (2018)) by adapting tests previously developed for the Greek language (Diamanti, Benaki, et al., 2017; Diamanti, Mouzaki, et al., 2017). Measures for vocabulary and working memory are used as covariates. Vocabulary was measured with British Picture Vocabulary Scale II adapted for Norwegian. Working memory was measured with Backward Digit Span.

Analyses

Data were analysed with a repeated measures ANCOVA conducted using IBM SPSS version 25.

Results

The results indicate a difference in how much the mean score for each test improved from kindergarten to first grade. Vocabulary and working memory did not affect this interaction between grade and test, but high scores on the covariates were related to higher scores on the overall mean of the morphological awareness tests.

Preface

First, I would like to thank the research group behind NumLit for giving me the opportunity to be a part of this interesting research project. It has been a great experience to participate in the data collection consisting of such a wide range of tests. I am grateful for the access to the data material from two measurements of so many children.

I would also like to thank my supervisor Athanassios Protopapas for all the thorough feedback and for helping me understand issues that seemed incomprehensible. I am truly grateful for having had a supervisor who has shared so much of his knowledge and who has encouraged me to push the boundaries of what I thought I could accomplish. I also want to thank my other supervisor Vasiliki Diamanti, for introducing me to field of morphological awareness and inspiring me to choose this topic for my thesis.

I cannot imagine how this process would have been without my classmates. Thank you for being such great company in the many and long breaks. I would also like to thank all my other friends who have supported me in this process. A special thank you goes to Anette for reminding me these past months that there is a life outside of Helga Engs.

Finally, I want to thank my family who has always supported and believed in me. I want to thank my mother for inspiring me to study special needs education and sharing with me her experiences from the field, my father for taking such interest in what I am doing and for proof reading my thesis. Last, but not least, I want to thank my brother for being the best brother and friend I could have wished for.

To the readers of this thesis, I hope you enjoy it.

Table of Contents

List of figures

List of tables

1 Introduction

1.1 Background and rationale

Morphology is the study of the structure and content of morphemes, which are the smallest, meaningful grammatical units in language, and the process of word-formation (Nagy, Carlisle, & Goodwin, 2014). Morphological awareness refers to the ability to analyse these units in a non-communicative context (Kuo & Anderson, 2006). Morphological awareness has been an emerging topic of research in recent decades. In the course of time, research increasingly suggests that morphological awareness plays a role in the development of literacy skills such as word reading, reading comprehension, spelling, and vocabulary (Carlisle & Goodwin, 2014; Kuo & Anderson, 2006; Nagy et al., 2014). Further, findings indicate that morphological instruction benefit both students with and without reading difficulties (Bowers, Kirby, & Deacon, 2010a). Morphological instruction may thus have educational consequences for all students and may ultimately contribute to reducing the gaps between the average and the weaker readers, in turn, making this an interesting field for both general and special needs educators. Even though most of the research on morphological awareness has been done on English speakers, research on the development of morphological awareness has indicated a similar development of morphological awareness across languages (Kuo & Anderson, 2006). In this study data from two measurements of morphological awareness are analysed to study whether the development of morphological awareness to follows a similar course in Norwegian monolingual children. There are some challenges related to the measurement of constructs which are not directly observable, such as morphological awareness. In the testing of such constructs there is a risk of measuring other constructs that are related to either the content of the tests or the execution of the tests. To assess the influence of such confounding variables it is relevant to use covariates. In this study the influence of the covariates vocabulary, which may be a confounding variable for the content of tests, and working memory which may be related to the execution of the tests, are included on the analysis.

The study I will conduct for my master thesis is in association with the «NumLit – Numeracy and Literacy development» project at the Department for Special Need Education at the University of Oslo. This project's aim is to, in short, study the development of language and

number comprehension and how different skills within these areas may be connected. NumLit is a longitudinal study going into its second year of 13. The children will be tested regularly from the last year of preschool until they are 18 years of age.

1.2 Research topic

The aim of this work is to describe how the development of morphological awareness unfolds, from the last year of preschool to first year of elementary school. In addition, it is of interest to examine whether vocabulary and working memory affects this development. On this basis I have chosen the following topic:

A longitudinal study of the development of morphological awareness across morphological domains in Norwegian children from kindergarten to first grade

In addition, the following research question will be answered:

Do working memory and vocabulary skills affect the development of morphological awareness differently across morphological domain and linguistic level?

1.3 Delimitations

The following study is based on two measurement, which may give an indication of the development of morphological awareness from kindergarten to first grade. However, analyses based on more than two measurements would generate results less affected by noise. For example, there is an age-span of 11 months within the sample, which means that there probably will be differences more related to the general development, than the development of specifically morphological awareness. This perspective is not included in the discussion because how yet more measurements might affect the results would be speculations. Having said that, the results might be a source for comparison for later studies on this topic.

1.4 Structure of this thesis

This chapter (1) presents the research question and the background and rational for this thesis.

Chapter 2 constitutes the empirical background for this thesis. Morphology's role and domains in language and morphological awareness are described. I continue with the development of linguistic awareness before I go deeper into the development of different morphological domains and morphological awareness. Subsequently, different approaches to measuring morphological awareness are presented, followed by issues related to the measurement of morphological awareness and an introduction of the covariates. Finally, the role of morphological awareness in literacy and relevant properties of morphology in the Norwegian language are presented.

Chapter 3 describes the methodological background for the study, with detailed information about the measurement tools. Validity and ethical issues are presented, followed by a presentation of the analytical method that has been performed, namely repeated measures ANCOVA.

Chapter 4 consist of a presentation of the analyses that have been performed and the results of these analyses.

In chapter 5 the results are discussed, taking the theory from chapter 2. Threats to the validity and reliability of the results will also be discussed in this chapter.

The final chapter, 6, the conclusion, summarises the findings and the limitations of this study.

2 Theory and empirical background

2.1 The role of morphology in language

According to Bloom and Lahey (1978), in language, three major components can be identified; content, form, and use. Content refers to the semantics of language, which is information or meaning. The form component consists of three categories known as phonology, morphology and syntax and these represent the tools to describe the form of utterances. The use of language implies the goals and functions of language, and also the linguistic and non-linguistic contextual adaptations made by individuals when trying to reach a goal of conveying a certain message. Based on this model of the components of language, it has been established that morphology is part of a language's form (Bloom & Lahey, 1978).

A morpheme is the smallest meaningful, grammatical unit in a language, which provide cues for meaning, spelling, and pronunciation. Furthermore, morphology refers to the mental system involved in word formation or the structure of words (Aronoff & Fudeman, 2011). Morphemes can be either bound or free. A bound morpheme is a morpheme that cannot stand alone, which means that it must be attached to other morphemes to form a word. Free morphemes, on the other hand, are forms that can stand alone (Aronoff & Fudeman, 2011). A further distinction can be made, which is relevant in morphological analysis, namely between roots and affixes. Roots are the base in words and cannot be analysed further. The root will give information about which lexeme a word originates from. In some languages root morphemes can be free, for example in English (e.g. boat and happy), whereas in other languages, roots are always bound morphemes. Affixes refer to units that can be added to a root before, in the middle or after a word and include prefixes, infixes, and suffixes (Aronoff & Fudeman, 2011). Prefixes are units attached before the root and an example of a prefix is *un-,* for instance found in "unhappy". Infixes are not frequent in English but refer to affixes added somewhere in the middle of a root. Suffixes are units attached after a root, and an example of a suffix is *-ness* found in for instance "happiness" and "awareness". Affixes can also be attached to words where other affixes are already attached, such as in "foolishness", where *-ness* is added to "foolish", which consists of "fool" and the suffix *-ish*.

In general, we differentiate between inflectional, derivational and compound morphology. Inflections involve adding grammatical forms such as tense, gender, quantity to a root morpheme (Aronoff & Fudeman, 2011). This implies that the use of inflectional suffixes to make inflected forms of words do not change their lexical meaning but do change their grammatical meaning. For instance, "jumped" is an inflected form of the verb "jump", in which the *-ed* suffix express that this action happened in the past. Derivational morphology, on the other hand, is the attachment of affixes to root morphemes that change their lexical meaning, creating a new lexeme (Bauer, 2003). For instance, when the prefix *un-* is connected to the word "true", the meaning changes from something being true to meaning that something is not true. When a word is derived it often also changes its word class, such as "sad" changing from an adjective to a noun when adding *-ness* in "sadness". The last morphological process is compounding, in which one or more lexemes are put together to form a new word. These can take the form of a single word or a pair of words. Examples of compound words are "doghouse" and "school bus".

In many languages, including English and Norwegian, written words are made up of both phonemes, units of sounds, and morphemes, units of meaning. Morphemes are as noted part of a language's form and make up a word's structure. Phonemes are the smallest sound units in words. In writing they are represented by one letter (e.g. /e/ found in "dr**e**ss") or more letters (e.g. /ð/ found in "**th**is"). The same letter sequences may represent different sounds such as in "react" and "read", where letter sequence "ea" is pronounced differently. In "react" this letter sequence corresponds to two graphemes representing the phonemes /i/ and /æ/, whereas in the infinitive form of "read" the letter sequence "ea" represents the phoneme /i/. Moreover, the pronunciation of morphemes is not consistent, and the different phonological variants of morphemes are called allomorphs (Lieber, 2016). For example, the affix marking plural in English (*-s)* represents different allomorphs. In the word "lips", for instance, the plural affix corresponds to the phoneme /s/ whereas in "balls" it corresponds to the phoneme /z/ (Aronoff & Fudeman, 2011). The pronunciation of graphemes is hence not consistent, and morphology plays a role, as it to a certain degree governs the pronunciation and spelling of words. In turn, this means that familiarity with a morpheme can help predict the pronunciation, such as in "react", which is a derivative form of "act", having the prefix *re-*. In other cases a morpheme in different forms has the same spelling, even when the pronunciation is different, which is the case in for example "sign" – "signal" (Muse, 2005). This correspondence between letters and sounds governed by both morphology and phonology in

English and Norwegian, results in the orthographic system of these languages being described as morphophonemic (Carlisle & Goodwin, 2014; Deacon & Kirby, 2004; Nagy, Berninger, & Abbott, 2006; Ragnarsdóttir, Simonsen, & Plunkett, 1999).

2.2 What is morphological awareness?

Morphological awareness refers to an explicit understanding of morphological relations between a word's form and meaning, such as grammatical inflection and productive derivation, and the ability to reflect on and manipulate the morphemic structure in words (Carlisle, 1995, p. 194). This ability is tied to reading, spelling, as well as vocabulary knowledge (Carlisle, 2003; Carlisle & Goodwin, 2014; Kuo & Anderson, 2006; Nagy et al., 2014). Morphological awareness differs from the acquisition of morphology in that morphological awareness focuses on the ability to reflect on and manipulate word formation rules in a non-communicative context. Acquisition of morphology, on the other hand, concerns the development of the ability to understand and produce morphologically complex words in natural speech (Kuo & Anderson, 2006). According to Kuo and Anderson, morphological awareness should be considered as a subordinate construct to the acquisition of morphology because morphological awareness can be considered as a more explicit representation and manipulation of implicitly acquired morphological rules.

The development of morphological awareness takes place when children are exposed to oral and written language. Berko (1958) found that children learning English as young as four years of age had developed some morphological awareness. The outset for this study was that if a child could use for example the correct plural ending to made up nouns, it demonstrates that the child had taken the rule of adding *-s* to nouns in the plural form into use. In turn, this indicates that the child will be able to inflect unfamiliar words correctly and are thus not dependent on having learned or heard a word to perform the appropriate inflection. The task Berko (1958) used in the study consisted of drawings and an examiner proclaiming one sentence with a non-word referring to a noun in a picture. The examiner simultaneously showed a picture and explained what could be seen. Next, the examiner presented the beginning of a sentence which the child was expected to finish with the non-word in an appropriate inflected form. For example, "this is a wug – now there is another one – there are two of them – there are two …". (expecting the child to say "wugs"). The four-year-old children in the study were able to make the correct inflections to some of the words, whereas

for other items there were only a few of the subjects who answered correctly. The findings of this study suggest that children have some morphological awareness in preschool, but that it is incomplete at this stage (Berko, 1958; Kirby et al., 2012). A different manifestation of the acquisition of the system of inflections is children's tendency to overgeneralise certain inflected forms in a period of time in their language development. The phenomenon of overgeneralisation refers to extending regular forms to irregular verbs, for example using "comed" when the correct form is "came". It appears that overgeneralising proceeds a period of applying the correct inflections, before the children return to the correct use of the irregular form (Marcus et al., 1992).

Derivations also seem to be taken into use at an early age. Preschool children learning languages such as English, German and French apply derivational rules to make new words according to the findings of Clark (1982). Clark used data from English, German and French children's production of transparent denominal verbs. Denominal verbs are verbs derived from nouns. One example is "I'm souping", which was said about eating soup. Based on the data, Clark proposed, that this use of familiar words and morphemes to create new words is a way of taking familiar words into use in order to express themselves. It is further suggested that the children create new forms like the mentioned example because they assume that nouns referring to an entity can also be used as verbs. Moreover, the findings by Clark (1982) could indicate that at the kindergarten level, children have developed awareness of simple derivations as well as inflections (Deacon & Kirby, 2004). However, even though it seems that children have some knowledge of derivational morphology in preschool, there seems to be a general consensus that inflectional morphology develops at an earlier stage than derivational (Kirby et al., 2012).

2.3 Terminological clarification

The term morphological awareness is often used interchangeably with other terms, such as morphological knowledge and morphological processing. However, these seem to not be the same constructs. In cases where the operationalisation of these terms has not been made explicit it can be difficult to grasp what construct the different terms refer to. In turn, the lack of consistency in how terms are used and what they are meant to signify can make it difficult to interpret and understand the concept that is being presented. The variation in how morphological awareness is used may also affect how research findings are interpreted

(Berthiaume, Bourcier, & Daigle, 2018). Nagy et al. (2014) make a distinction between tacit morphological knowledge, being the unconscious use of morphology, typical of normal language processing, and morphological awareness referring to a strategic use of morphemes. The authors chose to use "morphological knowledge" as a superordinate to tacit morphological knowledge and morphological awareness, with morphological knowledge being the global construct. This interpretation is also seen in other articles, for instance in the review by Kirby and Bowers (2017). Muse (2005) and Tannenbaum, Wagner, and Muse (2007) in comparison, refer to morphological knowledge as an individual's ability to correctly employ morphological units, for instance in a test item, but without necessarily having a conscious awareness of why these units are used in a particular way. This can be the case for a participant when the context is very informative. Morphological awareness, they claim, is based on the definition by Carlisle (1995) presented earlier, and refers to the conscious awareness of or the ability to use the smallest units of meaning in language, thus indicating that morphological knowledge is not a construct superordinate of morphological awareness.

It is possible that a line drawn between the two concepts - morphological knowledge and morphological awareness -, can involve a distinction between implicit and explicit awareness when using an understanding of the construct similar to that of Tannenbaum et al. (2007). This is different from the differentiation put forth by Nagy et al. (2014) between tacit morphological knowledge and morphological knowledge. A distinction similar to that of Nagy et al. is made by Carlisle (2004), between implicit and explicit knowledge. Implicit knowledge, according to Carlisle, comes from listening and using language and can be tapped in tasks where the child is expected to produce a correct word ending. Carlisle provided an example of a task of this sort in which a child is given a sentence; "Sally had a dress and Linda had a dress. Together they had two … "(dresses). Explicit knowledge, on the other hand, is referred to as "linguistic awareness" and can be tapped in tasks that require a higher ability of analysing language, which "may require a more explicit morphological awareness" (Carlisle, 2004, p. 323). Similarly, Muse (2005) refers to implicit morphological knowledge as the processing of linguistic information at a level of reflection that is not necessarily conscious, and a more explicit knowledge where conscious analysis and control of linguistic forms occur. Another nuance of the meaning of implicit and explicit knowledge is presented by Casalis and Louis-Alexandre (2000), who imply that only in tasks where the subjects must explain how they process, for instance an inflection, there is an indication of explicit awareness. Further they explain that a task may require cognitive control, but still mostly

target the implicit level of awareness, indicating that it is the degree of cognitive control required in a task that determines whether implicit or explicit knowledge is required.

The preceding two paragraphs indicate that there are nuances in how different terms are used, but that there seems to be a general conception that a child can progress from an implicit to a more explicit understanding of the structures in morphemes. If we go back to look at how the terms "implicit" and "explicit" have been understood in the development of linguistic skills, we can find theories by for instance Karmiloff-Smith (1986). Karmiloff-Smith proposed that access to explicit knowledge emerges progressively from implicit knowledge in a developmental course. An explicit knowledge in this model is a more accessible knowledge (Berthiaume et al., 2018).

As the relationship between morphological awareness and morphological knowledge is not quite settled and this discussion is beyond the scope of this thesis, I will only use the term "morphological awareness" throughout the thesis, based on the theoretical framework that will be presented in the following section.

2.4 The metalinguistic development of morphological awareness

In many studies, levels of morphological awareness are referred to as implicit or explicit. Even though there are variations in the terms and operationalisations used across the field (Berthiaume et al., 2018), as presented in the previous section, I will in the following present the framework for this current study, based on Diamanti, Benaki, et al. (2017) and García Grande (2018). First of all, a differentiation between morphological awareness and children's unconscious morphological operations should be made (Gombert, 1992; Karmiloff-Smith, 1986; Nagy et al., 2014). Young children producing overgeneralised forms such as "goed", is an example of unconscious use of morphology that is typical for normal language processing. This kind of overgeneralisation indicates a transitional, imperfect understanding of the rules of inflectional morphology. This is different from full-blown metalinguistic awareness, which is evidenced in non-communicative contexts requiring conscious reflection upon language elements. Measurements of morphological awareness will tap both unconscious/implicit and conscious/explicit awareness of morphology and the boundary between these is not clear (Nagy et al., 2014).

The attainment of metalinguistic awareness may not be either-or in terms of acquired level, but rather a gradual progression (Diamanti, Benaki, et al., 2017). Gombert's model, which will be described in more detail further down in this section, is based on the mentioned model by Karmiloff-Smith (1986). According to Karmiloff-Smith's model, implicit knowledge is reorganised in order to become more explicit and increasingly accessible to consciousness. During this development, knowledge goes through multiple reorganisations and explications. The reorganisations are driven by the metalinguistic judgements made by children and their spontaneous attempts to correct themselves. Gombert (1992) created a model of metalinguistic awareness that was inspired by Karmiloff-Smith's more general theorisation. In Gombert's model, the focus is on the links between children's oral language development and the learning of reading (Berthiaume et al., 2018). This model is not specific for the development of metamorphological awareness, and even though Gombert has made other specific models for the development of metalinguistic awareness for other aspects of language such as metaphonological development and metasyntactic awareness, it seems likely that the metamorphological development will follow a similar course, considering that different aspects of language are strongly connected (Carlisle, 2003; Casalis & Louis-Alexandre, 2000; Kuo & Anderson, 2006). The stages in the development of linguistic awareness proposed by Gombert are (a) the acquisition of first linguistic skills, (b) the acquisition of epilinguistic control, (c) the acquisition of metalinguistic awareness and (d) the automation of the metaprocess.

2.4.1 The acquisition of first linguistic skills

In this phase feedback from adults is important for reinforcing the production and comprehension that is adequate in order to eliminate incorrect language use.
A reorganisation is needed when the complexity and the linguistic requirements from the adults have increased, in addition to an increased length of the children's own utterances. The stabilisation and increased complexity set off the second phase. The child is required to resolve increasingly complex forms of communication and both the continuation of linguistic behaviour that has been responded with positive feedback and the reappearance of negative feedback is important in the transition. Finally, in the end of this stage the child seems to be able to use the language correctly.

2.4.2 Epilinguistic control

In this phase there is a reorganisation of the implicit knowledge from the previous phase and an addition of new knowledge to the child's knowledge which in turn becomes more and more organised and stable. A linking between the implicit knowledge takes place which leads to a functional, but un-reflected, awareness of a system of rules. This system of rules is more established, and the child acquires the ability to for example experience a dissonance when hearing an ungrammatical utterance. There is an elaboration of a stable reference point for each linguistic form, meaning that when they meet an unfamiliar linguistic context, they can rely on their knowledge of similar contexts. This is the principal characteristic of the second phase in Gombert's model (Gombert, 1992). A stable system which has a stable contextual functionality in the verbal exchanges, marks the end of the second phase. For the child to evolve into the next phase, fresh external stimuli are necessary to set off the development of meta-awareness.

2.4.3 The acquisition of metalinguistic awareness

An awareness at the meta-level is only acquired when it is demanded for accomplishing certain metalinguistic tasks. The acquisition of reading and writing skills provoke the metalinguistic awareness because reading and writing require conscious control of a number of linguistic aspects. Other metalinguistic training can also make the subject conscious on the meta-level. To reach this level, the epilinguistic control is a prerequisite; only what is mastered on a functional level can be mastered at a conscious level. On the meta level, the subject is able to, not only know the rules, but can also apply them. In this phase the child gains conscious access to explicit knowledge (Gombert, 1992).

2.4.4 The automation of the metaprocesses

We do not always consciously control our linguistic processes, and this is linked to the automation of the metaprocesses. There are two automated process whereas only one,

- the automated process – can always be replaced by meta-processes in contexts where it is required, as opposed to epi-processes that cannot be replaced by meta-processes when encountering an obstacle. As the stable epilinguistic control is assumed to be appearing towards the end of the fifth year, metalinguistic awareness will be triggered as the epilinguistic control is stable. According to Gombert, the first meta functions are identified around 6-7 years of age (Gombert, 1992).

2.5 Domains of Morphological processes

2.5.1 Inflectional morphology

Morphemes used to mark words or phrases for a certain grammatical feature are known as inflectional morphemes, and the corresponding processing domain is accordingly termed inflectional morphology (Maynard, Brissaud, & Armand, 2018). Forms created by adding the inflectional affixes have the same lexical meaning as the bare root (Pavey, 2010), which means that the words do not get a new lexical meaning when they are inflected.
Languages can be said to be rich of inflections to a varying degree. English, for instance, is a language that has relatively little inflection, while Mandarin has no inflections. Finnish, on the other hand, is categorised as a language rich in inflections (Gonnerman, 2018). There are different categories of inflections that are used in languages, and in the following, common aspects of languages related to Norwegian will be presented. Properties of Norwegian language will also be presented further at a later point.

Inflectional categories include, among others, number, gender, and tense. Number is an inflectional category which gives information about how many entities there are. In English the majority of nouns are marked for plural by adding *-s*, such as in "books". In other cases the noun form is irregular and does not take the plural *-s,* for example in "sheep" (Lieber, 2016). Languages that have grammatical gender have two or more classes in which nouns are divided into. The gendered nouns have to agree with other elements in a sentence, such as adverbs and adjectives (Lieber, 2016). An example in Norwegian is "en rød ball" – "ei rød bok" – "et rødt hus", (a red ball - a red book – a red house) where the article changes for each gender, "ei" being feminine, "en" masculine, and "et" neuter. While the adjective remains the same for feminine and masculine ("rød") it changes in the neuter form ("rødt"). Tense refers to the point in time of an event in relation to another point, such as future, present, and past. For example, the verb "talk" in the present form takes the modal "will" when the aim is to refer to something in the future ("will talk") and the suffix *-ed* when referring to something that has already happened in the past ("talked"). Other inflectional categories than those noted thus far are for instance case, voice and person. Case gives information about a word's grammatical role, for instance if the word functions as a subject, direct object, indirect object, and so forth. The inflectional category voice allows different verbs in phrases to be focused in sentence as active vs passive. Person refers to when verbs take different endings depending on

whether the subject of the sentence is in the first person, second person or someone else (Lieber, 2016).

2.5.2 Development of inflectional morphology

In the study of the development of morphological awareness, as previously noted, Berko (1958) found that children in preschool already have some knowledge of the functions of inflectional markers. In addition to finding this to be the case for marking plural, it was also found that in first grade most of the children were able to, to some extent, apply the correct allomorph in the past tense form, for instance /d/ in "spowed" and /ed/ in "motted". The data from this study also showed that this knowledge of inflections improved significantly from preschool to first grade. Berko's study has been replicated in multiple languages and the findings have been similar to those of Berko (Kuo & Anderson, 2006). According to the findings in the review of Kuo and Anderson (2006) most inflectional forms are acquired by the early elementary grades, even though there is some variation due to differences in complexity between languages. It seems that not only do children acquire the forms at around the same age, but they also seem to follow the same pattern. Results from several studies show that once children have acquired an inflection schema, they learn to distinguish regular from irregular items. After a period of applying the correct inflections for irregular verbs children tend to have a period where they overgeneralise the regular forms (Kuo & Anderson, 2006; Marcus et al., 1992). This phenomenon of overgeneralisation seems to be common in the development of inflectional morphology, including both English and Norwegian (Marcus et al., 1992; Ragnarsdóttir et al., 1999). After a period of overgeneralisation the schema becomes more elaborated and the children learn to distinguish between regular and irregular items (Kuo & Anderson, 2006). The assessment of inflectional morphology is important because it can be a good indicator of general language difficulties (Kirby et al., 2012). This is related to the difficulties children with language learning often have with learning inflectional morphemes (Smith-Lock, 1995).

2.5.3 Derivational morphology

Derivation is the process of adding prefixes or suffixes to word roots and thereby changing the meaning, and in some cases the word class. For instance, the word "behead" consist of the prefix *be-* and the root "head". In this context the prefix adds the meaning of "off", and the noun "head" is derived into a verb. The prefix *un-*, on the other hand, adds the meaning of

"not". This prefix can be found in for example "uncool" (un + cool) and "unhappy" (un + happy), and in these cases the word class, in which they belong, namely adjective, remains the same. The prefix *-un* can be found as *u-* in Norwegian in for instance "uheldig" (unfortunate). An example of a suffix is- *ify*, which is found "glorify" and "purify", where the first is derived from a noun and the latter is derived from an adjective and in these cases, both become verbs. Affixes that are used in derivations have restrictions regarding what words they can be added to. For example, as noted, the prefix *un-* can be added to adjectives, but cannot be attached to all adjectives (e.g. "uncute" is not a lexeme) (Bauer, 2003), and this concerns the use of *u-* in Norwegian as well. Another restriction in derivatives is related to what types of base morphemes they can be combined with. For instance, the suffix *-able,* can only be attached to verbs, but not to nouns, to form adjectives. Further, affixes vary with regard to productivity. This means that some affixes can only be applied to a small number of words, whereas others are more commonly used to produce new words.

2.5.4 Development of derivational morphology

The acquisition of derivational morphology starts at a later point than that of inflectional morphology and spans a longer time-period (Kuo & Anderson, 2006). Previous studies have found that first graders usually have only rudimentary knowledge of derived forms, far less than their knowledge of roots and inflected forms (Anglin, 1993). It is usually not until third or fourth grade that children start to develop a somewhat explicit awareness of the structure and meaning of derived forms, and such awareness continues to develop through the high school years, according to Tyler and Nagy (1989). In a review of the development of derivational morphology Anila, Desrochers, and Deacon (2018) found that even if knowledge of some derivational forms is acquired by third grade, for example *-er* (e.g. in "helper"), other forms are still difficult towards the end of primary school, for instance the suffix *-ness*. They suggested that exposure to printed words consolidates children's understanding of the relationship between meaning and morphological units. This idea can also be used to interpret the findings of Diamanti, Benaki, et al. (2017) with Greek children. They found that at ages four to seven years old, epilinguistic control was acquired for both inflectional and derivational morphology, but metalinguistic awareness of derivational morphology was not yet effective. This suggests that their awareness of derivational morphology was still developing.

There are some hypotheses why inflections develop at an earlier stage than derivation morphology. Among them is the fact that derivational suffixes are larger in number. Derived forms are also lower in frequency, especially in oral language, which is mainly what children encounter before they commence school (Kuo & Anderson, 2006). It may, in other words, be the case that children only encounter a limited number of the existing derived forms in kindergarten and the early pre-school years. Furthermore, in English, derived forms are transparent or opaque to a various degree. Transparent derived words are words without phonological and orthographic shifts, in comparison to their roots, for example "helpful", consisting of "help" and *-ful*. In contrast, opaque derived words have different spelling or pronunciation compared to the root, for instance "sign" – "signature", where the pronunciation changes. To investigate whether this variation plays a role in the acquisition of derivational morphology, Carlisle and Nomanbhoy (1993) looked at the effect of transparency in derivations in children's knowledge of this morphological domain. To test the hypothesis that the transparency plays a role, they used a test where the child is given a cue word of which they are expected to produce the derived form, after being supplied with a cue-sentence. In the end of the sentence the derived form of the cue-word will fit grammatically if the correct derivation has been made. An example of this is "help. Father tells me I am a good … "(the correct derivation being "helper"). "Help" - "helper" is an example of a transparent derivation. One third of the expected responses were phonologically transparent derived words, one third were opaque derived words, and one third were inflected words. The results showed that the children did significantly better on the inflected forms than on transparent-derived forms, and better on the transparent derived forms than on phonological-change forms.

2.5.5 Compound words

Compound words consist of two or more roots, forming a new word with a new meaning (e.g. rooftop). Different languages have different rules regarding the spelling of compound words. For example, in Norwegian they should always be written connected, e.g. "hundehus", in contrast to English, "dog house", where the compounds are not necessarily attached.

Children seem to develop awareness of compounds at an earlier stage than derivational morphology and it seems to continue to grow in the first school years (Kuo & Anderson, 2006). In the development of the morphological awareness tests used in this thesis, a test

targeting compound words was piloted but turned out to be too easy for the participants and was thus not used in the study (García Grande, 2018). For this reason, compound words and their development will not be discussed further.

2.6 Measurements of morphological awareness

The tools for measuring morphological awareness can be oral, written or combined oral and written. Tasks for measuring morphological awareness include judgement tasks and production tasks. In a judgement task the participant is expected to make a decision. A task of this kind does not require any manipulation of the structure of words but only using morphological principles (Kirby et al., 2012). One example of this kind of measurement is an epi-inflectional task in which the participants is expected to make a judgement of which of two sentences with words differently inflected matches a drawing illustrating one of the inflected forms (Diamanti, Benaki, et al., 2017; García Grande, 2018). An example is a drawing of a turtle playing with two monkeys along with which the examiner presents the sentences "the turtle plays with the monkeys" – "the turtle plays with the monkey", where a judgment of the first as the correct one is the right answer. Another judgment task, which targets derivational morphology, is the one used by Carlisle and Nomanbhoy (1993), in which the child judges whether a certain word is related to a similar-sounding word. For instance, "a person who teaches is a teacher" (yes/no) (Carlisle & Nomanbhoy, 1993).

In production tasks the participant is required to have abilities beyond recognition and will have to produce words by applying morphological rules (Kirby et al., 2012). These tasks can be inflectional used by for example, Berko (1958) and Diamanti, Benaki, et al. (2017). In an inflectional production task in line with Berko and Diamanti et. al's, the participants are expected to produce an inflected word, given cues in the form of one sentence with the target word in one form, and the beginning of a second sentence with contextual cues to prompt the desired inflection (optionally with matching drawings). An example of a production task with derivations is a task in which the subject is expected to supply a missing word, given either the root morpheme, or a derived form, in which case the task is to identify the root morpheme. When the root morpheme is given this is a derivation task, whereas when an already derived word is provided it is a decomposition task. Both tasks give information about the subject's knowledge of derived forms. For example, in the case of "Teach. He was a good …" (teacher)

the subject is required to derive the word “teach”. An example of a decomposition task is “The word is driver. Children are too young to … “(drive) (Carlisle, 2000; Muse, 2005).

In research we can find variations of judgement tasks and production tasks targeting different levels of morphological awareness. According to Carlisle (1995), judgement tasks provide evidence for epilinguistic control. The reasoning behind judgement tasks being appropriate for the assessment of epilinguistic skills is the little or no demand on metalinguistic manipulation and retrieval. The examiner provides the subject with the elements which are to be judged, and no production is required. Metalinguistic awareness is believed to be measured in production tasks (Diamanti, Benaki, et al., 2017) because production tasks put a higher demand on the subject by requiring both (a) a noncommunicative use of certain language objects in the test context and (b) retrieval, since the subject is expected to spontaneously produce a desired form which is based on the information and cues provided by the examiner (Diamanti, Benaki, et al., 2017).

2.6.1 Challenges with the assessment of morphological awareness

Since morphological awareness will always exist in a linguistic context there are some methodological issues related to the assessment of morphological awareness. Kuo and Anderson (2006) point to linguistic aspects that can confound the measurements targeting morphological awareness, such as vocabulary size and decoding skills in exercises where this is required. Using tasks that require decoding to measure morphological awareness can be confounding because it can be unclear whether it is the decoding skills that are measured or the morphological awareness. A potential method to avoid this is to present the tasks orally, but this puts higher demand on the short-time memory. Kuo and Anderson (2006) suggest a variant where the child is asked whether a sentence sounds correct instead of giving multiple options. This, they point out, will require enough questions to ensure a high level of reliability. Moreover, vocabulary size is highly correlated with morphological awareness in children learning various languages and it is one of the major confounding factors that need to be considered in assessing morphological awareness (Kuo & Anderson, 2006). This is a confounding factor because a response can be correct because of the vocabulary, not because of the ability to analyse the word. A way to avoid this can be to use non-words (Berko, 1958; Kuo & Anderson, 2006). In addition to this, high vocabulary skills may be a manifestation of

high overall linguistic skills. In this study working memory and vocabulary will be controlled for by being included as covariates and will therefore be introduced in the following.

Vocabulary

Vocabulary refers to the sum of words a person uses and/or knows (Burger & Chong, 2011). At five years of age children have on average 14,000 words in their vocabulary (Templin, 1957), but there are large individual differences, which have been found to be related to environmental factors, such as socio-economic status (Hart & Risley, 1995). Vocabulary is important for understanding, expressing one-self and thinking about the world. In addition to being an indicator of a person's overall level of intelligence (Sternberg, 1987), vocabulary plays an important role in language development, reading development and reading comprehension (Gathercole, Willis, Emslie, & Baddeley, 1992). Vocabulary knowledge can be divided into two categories, namely receptive and expressive vocabulary. Receptive vocabulary refers to the words a person can comprehend in for instance speech or text and includes words they cannot produce (Burger & Chong, 2011). Receptive vocabulary can be identified in recognition tasks where the subject can express their familiarity with a word. Expressive vocabulary on the other hand, refers to the words a person can use and produce in speech, written, or signed words (Burger & Chong, 2011). The assessment of expressive vocabulary requires a subject to name or explain what certain words mean and can thus give more information about the "depth" of a subject's vocabulary knowledge.

Working memory

Working memory is a construct defined differently across different fields of research, but there seems to be a consensus that it is a skill that develops over the age span and that it reaches adult capacities in the early adolescent years (DeCaro & Maricle, 2011). Working memory is assumed to be necessary for a wide range of cognitive activities and involves the temporary storage and manipulation of information (Baddeley, 2003). A central theory about the working memory is by Baddeley and Hitch (1974). They introduced a model consisting of three different subsystems of working memory, namely the phonological loop, the visuospatial sketchpad, and the central executive on which the first two depend. The phonological loop refers to verbal and acoustic information, while the visuospatial sketchpad refers to visual information. At a later point the episodic buffer was added to the model (Baddeley, 2000). This is categorised as a temporary storage system, where both visual and auditory information can be integrated together. Buffer refers to the temporality of the

episodes. The components of the working memory are believed to be connected to the long-term memory through the central executive. The central executive is thought to be related to – among other things – attention and problem solving (Gathercole, Pickering, Ambridge, & Wearing, 2004).

2.7 Morphological awareness and literacy skills

Several studies have found a relationship between morphological awareness and different aspects of literacy. This includes skills such as reading (Carlisle, 2000; Deacon & Kirby, 2004; Kirby et al., 2012; Ku & Anderson, 2003; Kuo & Anderson, 2006; Tong, Deacon, Kirby, Cain, & Parrila, 2011), which appears to be the case in a range of orthographies (Casalis & Louis-Alexandre, 2000; Elbro & Arnbak, 1996; Ku & Anderson, 2003; Kuo & Anderson, 2006; Nagy et al., 2006). Morphological awareness also seems to be related to spelling (Deacon & Bryant, 2005, 2006; Desrochers, Manolitsis, Gaudreau, & Georgiou, 2017) and vocabulary (McBride–Chang, Wagner, Muse, Chow, & Shu, 2005; Nagy et al., 2006; Sparks & Deacon, 2015; Tannenbaum et al., 2007). The relationship between morphological awareness and reading seems to concern word reading with the support of knowing base words, when reading derived forms (Carlisle & Goodwin, 2014). Vocabulary seems to be related to morphological awareness because familiarity with morphemes may facilitate the acquisition of new words (Carlisle, 2007). Spelling skills seem to be supported by knowledge of root morphemes (Deacon & Bryant, 2006). While it is well-established that phonological awareness is important in beginning reading (Melby-Lervåg, Lyster, & Hulme, 2012), it seems that grade level or age of the students affects the role of morphological awareness in the mentioned aspects of literacy (Carlisle & Goodwin, 2014). In the beginning reading children are occupied with mastering sound–letter relations and gaining fluency in the recognition of words. At a later grade level, more sophisticated linguistic knowledge may be needed for word recognition and comprehension processes, and this can mean that morphological awareness plays a more important role at these stages (Carlisle & Nomanbhoy, 1993).

Based on findings suggesting that morphological awareness is related to literacy, researchers have been interested in finding if and how this can be used in educational practice. Examples of what has been studied is whether morphological training could benefit dyslexic's spelling and reading skills (e.g. Arnbak & Elbro, 2000). As noted in the review of Carlisle and

Goodwin (2014), studies of this kind often aim to improve word-analysis strategies that the children can use when they encounter unfamiliar or words that are difficult to spell. Other studies have targeted whether morphological training in preschool benefit students at a later point when the reading instruction begins (Lyster, 2002; Lyster, Lervåg, & Hulme, 2016). Bowers, Kirby, and Deacon (2010b) on their part, found that morphological instruction is more effective when integrated with other aspects of literacy instruction.

2.8 Properties of Norwegian morphology

This study has been conducted in Norwegian, a language with inflectional and derivational morphology, and a semi-transparent spelling-sound orthographic system (Seymour, Aro, & Erskine, 2003). The grammatical categories are related to certain word classes and not all word classes are inflected in all of the grammatical categories. In Norwegian for instance, nouns are inflected in number, whereas this is not the case for verbs. Verbs, on the other hand, are inflected in time (i.e. tense), but nouns are not (Lie, 2011). Nouns and adjectives are inflected for number, gender and definiteness. There are three genders in Norwegian, feminine, masculine and neuter. Number has two forms for singular, one indefinite and one definite and this is also the case for plural. An example of how singular is inflected in the indefinite and definite form is: “ei livlig jente” – “den livlige jenta” (a lively girl – the lively girl). These examples are both in singular, the first is indefinite, the latter definite. Gender is in this example represented by the article “ei” and the inflectional suffix *-a* in the definite form. It should be noted that the feminine form can be inflected in the masculine form, and which form is chosen depends on dialect and the sociolect. Adjectives are inflected for degree for example: “han er kul” – “han er den kuleste” (he is cool - he is the coolest) where the first is the base form and the last the superlative form. Verbs are morphologically inflected for tense, mood and voice. Norwegian has both regular and irregular verbs. The regular verbs have an ending in each of the inflections and the root remains the same, whereas irregular verbs do not have an ending in simple past and usually the vowel in the root will change through the inflection. Regular verbs can be divided into two subgroups, which differ in the way they are inflected in simple past and perfect (Lie, 2011).

Derivations in Norwegian can be made with prefixes, for example *u-* in “urettferdig” (unjust) or suffixes, like *-het* in “rettferdighet” (justice) and a combination of the two (e.g.

urettferdighet – injustice). When a suffix is added the word class will in many cases change, whereas when a prefix is added, the word class remains the same (Rønhovd, 1997).
In Norwegian compounds should always be together in one word. Usually the last word in the compound will indicate the meaning of the word and determine the word class, for example *ananasringer* (pineapple rings) which are rings of pineapple.

There are no studies about the nature of the development of morphological awareness in Norwegian, but in an experimental study Ragnarsdóttir et al. (1999) looked at the nature of the development of inflectional morphology in Norwegian and Icelandic. The results of this study show that the acquisition of inflectional morphology relies on an interaction between frequency, verb type and age. According to their findings, children at the age of 4 are able to form the correct inflections on verbs from the largest group of regular verbs (taking the suffix *-et* in simple past and perfect). By the age of 8 years the children had 90% correct responses on the tasks, which were targeting both groups of regular verbs and irregular verbs. The children developed their inflectional knowledge of the smaller group of regular verbs and irregular verbs especially between the age of 6 and 8 years.

2.9 Synopsis of the theoretical section

Morphological awareness refers to the ability to reflect on and manipulate the morphemic structure in words (Carlisle, 1995), a skill that is tied to reading, spelling and vocabulary knowledge (Anila et al., 2018; Kuo & Anderson, 2006; Nagy et al., 2014).
The construct morphological awareness is related to other constructs, and in the field of research there is a variation of terms used to describe similar constructs, and it is not always clear how the different terms are operationalised (Berthiaume et al., 2018). There seems to be a consensus, however, that children develop skills that make them able to analyse the different morphological components in words, although the terminology varies. In this study the theoretical framework is based on Gombert's model of the development of metalinguistic awareness (Gombert, 1992), where the distinction between epilinguistic control and metalinguistic awareness is central. Between these levels, the child is expected to go from recognising the correct use of language to being able to apply the rules that govern the language in unfamiliar contexts. In the development of morphological awareness, the domain of inflectional morphology seems to develop at an earlier point than that of derivational morphology. This is likely to be tied to for instance higher productivity in inflectional

morphology. The development of derivational morphology also seems to span a longer period of time that that of inflectional morphology, which is thought to be related to the variations in level of transparency in derived forms and the high number of derivatives (Kuo & Anderson, 2006). Different tasks tap different aspects of morphological awareness, regarding both domain and level of awareness. In the operationalisation of epilinguistic control it is suggested that judgement tasks are appropriate, whereas production tasks are suited to tap metalinguistic awareness (Diamanti, Benaki, et al., 2017). Morphological awareness is a skill related to other linguistic skills and the measurement of it may put load on other cognitive skills. However, adjustments in design and the use of control variables may help ensure the desired skill to be tapped. In particular, vocabulary is related to other linguistic skills and a high vocabulary score may confound tasks intended to measure morphological awareness. Furthermore, oral tasks used for measuring morphological awareness may be confounded by working memory skills, by possibly posing too high demands on this ability (Kuo & Anderson, 2006). Morphological awareness is an interesting field of research because it is related to different literacy skills. Research indicates that using morphology as an approach in reading and spelling training might have a beneficial long-term effect on these skills (Bowers et al., 2010a; Lyster et al., 2016). In addition, it might be a useful tool for those struggling with reading and/or spelling.

3 Method

The aim of this study is to learn more about the development of morphological awareness, which is a field that until now has not been much explored in Norwegian. This study includes results from two measuring points; the last year of kindergarten and first grade. The same tests have been used both years. Thus, we can to compare the results whilst being certain they have measured the same skills and individuals at two time points.

3.1 Longitudinal design

The data for the study of this thesis are collected from an ongoing longitudinal research project going into its second year of 13. NumLit is a panel study, because the same individuals will be followed over time (De Vaus, 2014; Gall, 2007). In other words, the same sample of individuals will be assessed from their last year of kindergarten until they are 18 years old. Studies with a longitudinal design aim to be able to describe the development of a skill over time (Cain, 2010), and this design gives the opportunity to explore changes in specific children and look for reasons why they might have developed as they have (Gall, 2007). With longitudinal studies it is also of interest to see if certain skills can predict other skills, and if so, determine what skills are best suited to predict other skills (Cain, 2010). The NumLit study has a non-experimental design which implies that it is not in this study's purpose to change anything, but rather to study the state of something at certain points in time (Kleven, 2002b).

3.2 Sample

The sample this study is based on, is taking part in the NumLit project and originally consists of 246 children with Norwegian as their first language. They are without neurological deficits or any known learning difficulties. The age range is from being born in January to December 2012. The sample is located in municipalities around Oslo that are all considered representative of the population with respect to socioeconomic status and parents' education. The sample was recruited through kindergarten staff who informed parents about the project and asked if their children could participate. In addition, the children have been requested to participate in advance of the assessments. The analyses will be based on 187 (F=95, M=92) of

the children in the original sample because of the availability of data at the time the analyses were performed.

3.3 Data collection

The first data collection for NumLit was completed in the period from December 2017 to March 2018. The second collection took place from mid-January 2019 to the end of March 2019. In both periods the data collection was performed by master students from the Department of Special Needs Education, the Department of Education, and the Department of Psychology from the University of Oslo. The students carrying out the second data collection were not the same as in the first. The test-battery in the first measuring consisted of 26 tests, which were divided into three sessions that where to be conducted on separate occasions. Each session lasted for about 45-70 minutes (Gjerde, 2018). At the second period of data collection, there were 33 tests, divided into three sessions as well, completed one at the time, each session lasting about 45-90 minutes. All students participating as examiners received obligatory training for each of the three parts on separate occasion in advance of the testing. On all occasions has the testing taken place in a separate room ensuring a quiet environment during the assessments. In both testing rounds, the children received a sticker for each completed task and a diploma with the stickers attached at in the end of the session. Breaks were provided when needed. The testing was audio recorded on every occasion. This made it possible to go back and hear responses again if there was any uncertainty during the scoring. In addition, it works as an assurance for the parents.

3.4 The variables

The tests for measuring morphological awareness consists of three tests. These were constructed and validated for Norwegian for the purpose of being used in NumLit by Germán Garcia Grande, based on the test made for Greek by Diamanti, Benaki, et al. (2017); Diamanti, Mouzaki, et al. (2017) (García Grande, 2018). These tests have been used at both measuring points. As in the study by Diamanti, Benaki, et al. (2017) the tasks measuring awareness of inflectional morphology follow the operationalisation of Carlisle (1995). In this operationalisation, a distinction is made between different levels of metalinguistic knowledge of morphology by using a judgement task to tap epilinguistic skills and a production task to tap metalinguistic skills, as described in the theory chapter. In the following I will give a presentation of the tests and the properties of the items. I will also present the instructions and

the administration for the different tasks. The covariates vocabulary and working memory are also a part of the NumLit dataset and were chosen among other tests tapping the same skills, and these tests will also be presented. British Picture Vocabulary Scale II (BPVS II) was preferred to use over the vocabulary test from Wechsler Preschool and Primary Scale of Intelligence IV (WIPPSI IV) (Wechsler, 2012) since the latter is a test that requires the subject to have verbal expression skills, in addition to vocabulary knowledge. The Backward Digit Span test was chosen over Listening Recall by preference, on the basis of close to equal reliability coefficients.

3.5 Tests tapping morphological awareness

3.5.1 Epi-inflectional awareness judgement task

The aim of this test is to measure the children's epilinguistic control and thus targets their unconscious linguistic knowledge (Diamanti, Benaki, et al., 2017; García Grande, 2018). It is a two-alternative forced task in which the child has to judge which of two sentences is produced with the grammatical inflection that is appropriate for a displayed image. The target words that are the objects of inflection, are in this task non-words, (word-like forms with no semantical meaning). The task is carried out by presenting the child with a picture of a turtle conducting an action. Simultaneously the examiner is presenting two similar sentences using finger puppets, that represent one sentence each. There is a target word in the sentences that appears with different inflections in the two sentences, and the task is to judge which of the inflections corresponds to the displayed picture. The child expresses the answer by pointing at the puppet thought to represent the sentence matching the picture. One example is a picture of a turtle bouncing a ball and the sentences presented by the examiner is "the turtle bounces the ball" – "the turtle bounces the balls" (in Norwegian: skilpadden spretter kebelen/skilpadden spretter keblene. Kebelen is the non-word used to replace ball). The finger doll saying "the turtle bounces the ball" is consequently the correct response in this example. The task consists of 16 items and four examples. The non-words that are used in the items are similar to Norwegian words in phonology and spelling structure to reduce the cognitive load. Different word classes are represented in the test namely nouns, verbs, and adjectives and the properties of the items from each word class are presented below (García Grande, 2018).

Property of verb items

In the epi-inflectional control judgement task there are five verb items, plus two among the examples. All the verb item sentence pairs have one sentence in simple present and the other sentence in simple past and all the verb items included in this task classify as regular verbs. This means that irregular verbs in which the past tense indication follow special rules, such as phonological shift in the root, are not included (e.g. “they leave” – “they left”). Another type of a verb tense that are not represented in this task are verbs that are dependent on modal verbs. In Norwegian, as in English, the tense present perfect is an example in which a modal verb is required, namely the verb “have” in simple present. In Norwegian, regular verbs are characterised by having the suffix *-er* as an indicator of simple present. In the simple past tense, regular verbs can be divided into two inflectional categories: class one include the suffixes *-a* and*-et* in the simple past and the second, smaller, class include the suffixes *-t, -te*, *-de* and *-dde* in simple past (Rønhovd, 1997). Both of the inflectional categories are represented in the items, with the suffix *-et* from the first class, in three of the items (of which one is an example item), and the suffix *-te* from the second class in four of the items (including one example item). In order to minimise the probability that there would be uncertainty whether a picture depicted an action in the present or past, the adverbs now for present, and yesterday for past, are included at the beginning of each sentence. The sentences are made appropriate for Norwegian by having the verb is in the second place in the sentence (García Grande, 2018). This makes the structure of the verb sentences “now building the turtle a house” – “yesterday built the turtle a house” (in Norwegian: nå vukker skilpadden et hus – igår vukket skilpadden et hus).

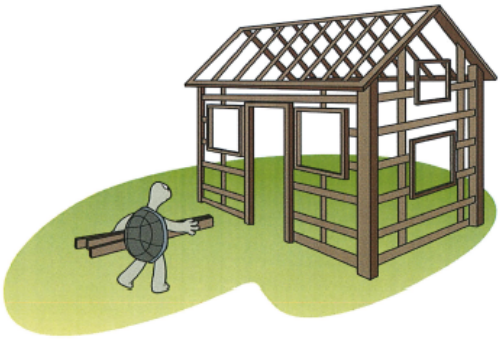


***Figure** 1 Epi-inflectional verb item in simple present*

Example:

1st puppet: **Skilpadden vukker et hus**
2nd puppet: Skilpadden vukket et hus

Properties of noun items

There are six noun items in this task, in addition to one example item. In the Norwegian language, morphological suffixation in nouns convey information about quantities; singular and plural and about gender; feminine, masculine or neuter. As is the case in English, indefinite forms have an article before the noun. For example, “a map” is “et kart” in Norwegian. The indefinite singular form of nouns is included in this task to contrast the plural definite form, which takes an inflectional suffix. The gender of a noun is conveyed in the indefinite singular article “ei” for feminine, “en” for masculine and “et” for neuter. In the indefinite plural forms the genders have the same inflection as a general rule, and this suffixation is – *er*. There are exceptions to the rule of this suffixation, in which the plural indefinite has no suffixation, and this is not unusual to find in neuter nouns (Lie, 2011). However, in this test, the noun items in plural are with suffixation. In the definite singular the suffixes are *-a* for feminine words, *-en* for masculine words and *-et* for neuter words. This means that the definite singular form of, for example, “the picture” is “picturethe” (bildet). The general rule for suffixation of plural nouns in the definite form is *-ene,* independent of gender.

In the items in this task, three sentence pairs include a plural indefinite noun with a pairing sentence is the indefinite singular form including one example item. The remaining five

sentence pairs have one sentence with a noun in the definite singular form *-en* for masculine or *-et* for neuter and a noun in definite plural *-ene* in the matching sentences. In this task, only masculine and neuter nouns are included. This is partly because feminine nouns can be ambiguously suffixed; as both feminine and masculine. In addition, the use of masculine suffixes is considered to be more conventional and is more encountered in written form. There are five sentence pairs that have a masculine noun, plus the example, and one sentence pair with a neuter noun.

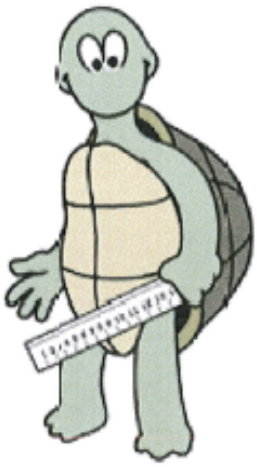


Figure 2 *Epi-inflectional noun item in indefinite singular*

Example:

1st puppet: **Skilpadden holder en væsp**
2nd puppet: Skilpadden holder to væsper

Properties of adjective items

There are five adjective items in the epi-inflectional task in addition to one among the example items. Words that belong in the word class adjectives indicate something's properties or features, such as being nice, kind and patient. In different inflected forms, in both English and Norwegian, adjectives indicate to which degree the properties manifest, such as in "big" - "bigger" – "biggest". In this task there are only adjective indicating properties of nouns, meaning that there are no adjectives with the suffixation for the comparative and superlative forms (bigger, biggest). Adjectives can be either simple or compound in Norwegian, as in

English. Compound adjectives have their origin in a noun or a verb, such as beautiful, with the suffix *-ful*, making the noun beauty the adjective beautiful. Simple adjectives are words that are not derived from a noun or a verb, such as long, tall, and short. Only simple adjectives are used in the items in this task. The adjective items are targeting the potential understanding of how suffixes in adjectives operate and how they affect the nouns they are related to. Like nouns, adjectives are inflected depending on quantity, gender, and definite and indefinite determiners in Norwegian. For instance, most adjectives do in relation to a neuter noun get the suffix *-t* in the indefinite singular form. For instance, will "høy" become "høyt" (high). When adjectives are related to feminine or masculine nouns in the indefinite singular form, their structure is not affected, since the adjective remains without further gender suffixation. Adjectives related to plural indefinite nouns, get the suffixation *-e,* independent of the gender of the noun. Adjectives related to nouns in the definite, singular form, take the suffix *-e*. Sentences with a noun in the definite form with an adjective is required to have a supplementary definite article in front of the adjective both in cases of singular and plural. This means that the noun in singular definite form "flasken" (bottlethe) needs to be preceded by the article "den" to be grammatically correct when presented with an adjective. Thus, the grammatically correct structure should be "den fine flasken" (the nice bottlethe). The definite article varies dependent of a noun's gender and is "det" in neuter form, "den" in feminine and masculine form. In plural definite form the article preceding the adjective and noun is "de", regardless of gender. Furthermore, adjectives' suffix in the plural definite form is the same as in singular definite and plural indefinite form, taking the suffix *-e*.

In the items of this task, the nouns related to the adjective items are all masculine. The structure of the sentences in the adjective items can be distinguished into two. The first is a comparison of a singular and plural definite adjective such as the "the turtle is throwing the big shoes" (skilpadden kaster de kvyre skoene) – "the turtle is throwing the big shoe" (skilpadden kaster den kvyre skoen). The second type of items have an adjective in the singular indefinite form versus one in the plural indefinite form. An example of a sentence pair of this kind is: "the turtle is eating a sour lemon"– "the turtle is eating sour lemons" (skilpadden spisen en prin sitron - skilpadden spiser prine sitroner).

Figure 3 *Epi-inflectional adjective item in indefinite plural*

Example:

1st sentence: Skilpadden spiser en prin sitron
2nd sentence: **Skilpadden spiser prine sitroner**

Instructions and administration for the epi-inflectional control judgement task

The instructions for this test consist of the examiner telling the child that they are going to look at some pictures of turtles together, followed by explaining that two puppets will describe what the turtles do in their own special language. The examiner explains to the child that only one of the puppets is expressing the action displayed on the picture correctly, and that this is not the same puppet every time. The test starts off with four examples and the examiner gives positive feedback after a correct answer and corrective feedback if the child chose the wrong puppet. The examiner can repeat the item sentences once if the child requires it. The scoring is 1 for correct answer and 0 for incorrect. The correct answer is marked in bold font in the scoring sheet, so the examiner can easily see if the child has given the correct answer.

3.5.2 Meta-inflectional awareness production task

The aim of this task is to measure the children's morphological meta-awareness, which refers to a more conscious control of the different aspects of language, than in epilinguistic control. This is assumed tapped by having a production task (Diamanti, Benaki, et al., 2017). In the

completion of the test, the child is presented with two pictures of a turtle conducting different actions and a sentence with an inflected target word matching the action of the turtle in one of the pictures. The examiner then presents the beginning of a second sentence and the child is expected to produce the ending of the second sentence with the target word inflected so that it matches the second picture. An example is "the turtles greet the monkeys" – "the turtle greets the" … (expecting the child to produce "monkey") (in Norwegian, skilpadden hilsen på ådene - skilpadden hilser på åden). In the first picture the turtle is greeting multiple monkeys and in the second the turtle greets one monkey, making "monkey" the correct answer in this item. The same non-words are used in this task as in the epi-inflectional test. The use of the same non-words and the same sentence pairs as in the previous test is expected to facilitate the exercise, since they have heard the inflections before. In addition, the same pictures where used in this task as in the epi-inflectional task, but for this task pairing pictures are used as well in order to have a corresponding picture for each sentence. By using non-words in this task, we can be fairly certain that if a child answers correctly, he or she has internalised the working system of grammatical inflections. This means that the child is able to generalise to words in the same word class, even if the word is unfamiliar and thus aware of the inflection independently of the lexicon (Berko, 1958). Nouns, verbs, and adjectives are the word classes represented and I will present the properties of the items in each word class in the following. There are 16 items in total; six are nouns, five adjective and five verbs and four are examples.

Properties of the verb items

There are five verb items in this task plus two verb items among the examples, and they are the same as in the epi-inflectional task. This means that the sentence pairs consist of one sentence in simple present, and the other in simple past. When the targeted inflection is in simple past, it is either with the suffix *-et* or *-te*. In the task the examiner points to one of two pictures and presents a sentence such as "now the turtle walks slowly". The examiner then presents a second picture and introduces the second sentence by saying "yesterday the turtle…" and giving the child the opportunity to produce "… *walked* slowly" (In Norwegian, nå gudder skilpadden sakte – I går …*guddet* skilpadden sakte). In Norwegian present continuous is not indicated with the suffix -ing, in the same manner as in English. To refer to an action happening in the now in Norwegian, simple present is normally employed, which means that there is no distinction between "walk" and "walking". In Norwegian, present continuous and simple present is differentiated by using time adverbs such as now and yesterday, which is what is done in the items in this task and in the epi-inflectional task. The

sentences are consequently “now the turtle is walking slowly – yesterday the turtle walked slowly”. The target sentence varies being the sentence in simple present and the sentence in simple past.

Figure 4 *Meta-inflectional verb item in simple present and simple past*

Example:

1st sentence: Nå gudder skilpadden sakte
2nd sentence: **I går guddet skilpadden sakte**

Properties of the noun items

There are six noun items in this task, plus one in the example. As for the verb items, the noun items are the same as in the epi-inflectional task. The children are expected to make an inflection of a non-word-noun, in order to describe a picture illustrating either plural or singular the indefinite or the definite form. In the same manner as for the verb items, the examiner provides one sentence that corresponds to one picture and then commences on the second sentence that the child is expected to finish. For the sentence pairs and the example item that has an inflection in plural indefinite form, the first sentence provided by the examiner is singular indefinite form, thus the sentence pair is for instance “the turtle plays with a cat” – “the turtle plays with three …”, expecting the child to produce the inflected form “cats” based on the contextual cues. There are three sentence pairs of this sort, and as in the epi-inflectional task, two of these are among the examples. In the remaining sentence pairs, there are one sentence in the singular definite form and one sentence in the plural definite form. The examiner either presents the sentence in the singular or plural form, which means that the inflection the child is expected to produce varies. An example of a sentence pair of this sort is “the turtle carries the bags” – “the turtle carries the…” (…bag) (skilpadden bærer

favene – skilpadden bærer faven) where the definite suffix -*ene* reveals the determination of the noun (the= the suffix -*en* in singular and -*ene* in plural).

Figure 5 *Meta-inflectional noun item in plural definite and singular definite form*

Example:

1st sentence: Skilpadden bærer favene
2nd sentence: **Skilpadden bærer faven**

Properties of the adjective items

There are five adjective items in the task and one example. The same items as in the epi-inflectional task are used here. This results in the sentence pairs either consisting of nouns in the definite form or of nouns in indefinite form, with associated adjectives. The first type of sentence pair in definite form includes the sentence "the turtle throws the new shoes" – the turtle throws the" … (...new shoe) (skilpadden kaster de kvyre skoene – skilpadden kaster den... kvyre skoen). An example of the of the remaining four items in the indefinite form is "the turtle is drinking from a nice cup" – "the turtle is drinking from…" (…nice cups) (skilpadden drikker av en flei kopp – skilpadden drikker av... fleie kopper). For the adjective items it is expected that the child produces both the adjective and the noun it is related to. This is because in the case of the comparison between adjective singular and plural definite nouns, the adjectives get the suffix -e in both forms (kvyre skoen – kvyre skoene) and the inflection of the noun will reveal whether the child has produced the expected inflection. The examiner is to encourage the child to do this in order to be able to confirm that the child is fully aware of the inflection and not just repeating the first sentence.

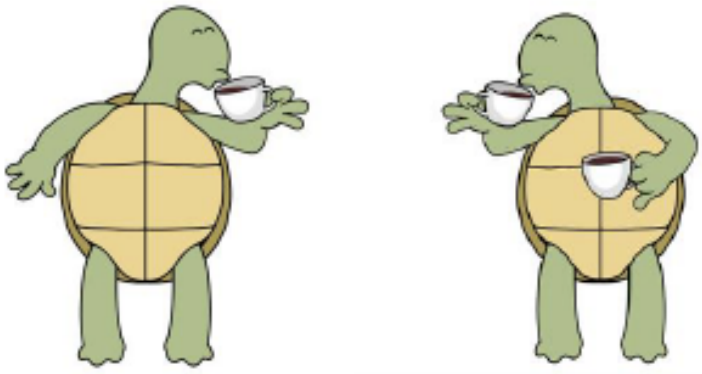


Figure 6 *Meta-inflectional adjective item in indefinite singular and indefinite plural*

Example:

1st sentence: Skilpadden drikker av en flei kopp
2nd sentence: **Skilpadden drikker av fleie kopper**

Instructions and administration of the meta-inflectional awareness production task

In the instructions associated with this task the examiner explain that they will continue to look at some pictures with the turtle doing different things, but this time the examiner will describe the first picture and the child will describe what happens in the second one. The examiner points out that they will continue using the turtle's special language to describe the pictures in the same way the puppets did in the first task. There are four examples, where one is a noun item, one is an adjective item and two are verb items. The examiner gives feedback to ensure that the child has understood the task. The examiner can repeat the item sentences once if this is required. The children get 1 point for the correct answer and 0 for incorrect. In terms of the scoring there are some remarks. First of all, points can be given even if a word is not phonologically correctly pronounced, but correctly inflected. Secondly the child can get a point as long as a grammatically correct inflection is performed, for instance if the child's response is in indefinite form, when the presented inflection in the first sentence is in definite form. For example, if the child responds, "two balls" (to kebler) instead of "the balls" (keblene). There is an attached table of the accepted forms associated with this task provided for the examiners.

3.5.3 Meta-derivational production task

Like the previous tests, this test targets the metalinguistic skills of the children, where manipulation of morphemes is the task (Diamanti, Benaki, et al., 2017). In this task the children are asked to add a corresponding derivational suffix to a target word. A sentence, including the target word, is presented first, with an appropriate drawing, followed by a presentation of the next sentence except the word of which the child is expected to produce the derived form. An example is: "Maria writes in class" – "Maria has a class of …" (…writing) (Maria skriver i timen. Maria har en time med … (skriving)). In the items, the syntactic structure of the sentence varies from the first sentence to the second. This is because deriving words often give them a different meaning and thus, an adjustment of the sentence is necessary to provide sufficient contextual cues to fit the targeted derived form. In this task only real words were used, which according to Diamanti, Benaki, et al. (2017) is because the derivational knowledge/awareness develops at a later point than inflectional, making it too difficult if non-words were to be used. In addition, derivatives are more based on meaning than inflections. A challenge with using real words is the question whether one is testing word knowledge or level of awareness, which could confound the results (Diamanti, Benaki, et al., 2017). These authors concluded with the limitation of using real words not being too influential as they found the meta-derivational task with real words appeared be more difficult than the meta-inflectional task with non-words. In turn, they suggest that this indicates that awareness of derivations has not developed as much as that of inflections. Whether this may be the case in Norwegian as well will be discussed in the discussion chapter.

Properties of items

A derivation often changes the word class a word originally belongs in, and in this task, this was the case in the largest portion of the items. Derivations can be made both by adding prefixes and suffixes, but in this task, there are only suffixed items. In this task a stem morpheme is used as the prompting sentence and is either a noun or a verb, except in two items, which are decomposition tasks where the root morpheme is the target word, and the prompting word is in the derived form. In Norwegian, as in English, there are a number of suffixes that can be used to make derivations, some of which are used in this task. The first group of items consists of roots that become nouns after being derived. These derivations are verbs in an inflected form in the root form. In these cases, they get the suffix *-else* such as in "beveger" – "bevegelse" (moves - movement) and *-ing* in "skriver" – "skriving" (writes -

writing). The remaining items are adjectives derived from different word classes. Some are derived from verbs, such as when they have the suffix -*ende* as in "trøster" – "trøstende" (comforts - comforting). An exception is the example item, "forvirret" – "forvirrende" (confused-confusing), where the suffix -*ende* is from an adjective already derived from the verb "forvirre" (confuse). Other adjective items have the suffix -*ete* in the derived form. In these cases, some roots are nouns and some that are verbs. The verbs in these cases are already derived from nouns such as in the verb "fleiper" (joking) that has its root in "fleip" (joke). This item is consequently "fleiper" – "fleipete" (joking – cheeky). One of the items has the suffix -*som* which in this case makes an adjective a new adjective. The root morpheme in this case is originally the noun "slit" (toil). This is "sliten" – "slitsomt" (tired-tiering) in the task.

***Figure** 7 Meta-derivational item with suffix -else*

Example:

1st sentence: Nikolai beveger seg til musikken
2nd sentence: **Nikolai gjør en bevegelse**

Instruction and administration of the meta-derivational task

In the instruction for this task the examiner informs the child that they will look at some pictures together and that the examiner will describe what they see in each picture, succeeded by the examiner expressing the beginning of a second sentence. There are four examples in this task and in the presentation of the examples the examiner is to stress the target word in the first sentence. The child is then expected to finish the second sentence with one word, being the derived form of the target word. In the examples, the examiner is to give feedback by either giving positive feedback in cases of a correct answer and correcting feedback in cases of a wrong answer. Throughout the whole test, the sentences can be repeated once if required. The items are scored with one point when the derivation is correct and zero if it is incorrect.

See appendix (1) for the full tests. Further reading about the development and validation of the tests used, see García Grande (2018)

3.6 Covariates

3.6.1 British Picture Vocabulary Scale II (BPVS II)

BPVS II is a test which aims to measures a child's receptive vocabulary. This test is translated and adapted from the British version (Dunn, Dunn, Whetton, & Burley, 1997) into Norwegian norms based on 884 children by Lyster, Horn, and Rygvold (2010). The targeted group are children between 3 and 16 years. In this task the examiner presents a word and the child is expected to choose the corresponding picture out of four alternatives by pointing or expressing the number of the picture. There is an increasing level of difficulty from high frequency to more infrequent and abstract words. The items are divided into 12 sets consisting of 12 items each. When 8 or more mistakes are made within one set, the stop criterion has been met. In the data collection the results of this test are based on, the test was electronic for the purpose of being used in NumLit. The pictures were presented on a computer and a recorded voice presented the words. The answer was given by the child pointing at the word they believed was correct followed by the examiner clicking on the chosen picture. A new item automatically appeared after the response had been administered by clicking. The program automatically shut when the number of allowed mistakes is set was reached.

3.6.2 Backward Digit Span

The alpha reliabilities (excluding zero-scoring children and zero-variance items) are .80 for Digit Span and .84 for Listening Recall, which is the other working memory test in the NumLit test battery. Since these values do not favour any of the tests to a large degree, I chose Digit Span to use as a covariate in this study.

Digit Backward Span is originally found in the test battery *Working Memory Test Battery for Children* (WMTB-C) (Pickering & Gathercole, 2001). The task for the subjects in this test is to repeat a series of numbers, provided by the examiner, backwards. For example, if the examiner says "9-2" the subject is expected to respond "2-9". There are six series of tasks which consists of six items in each. In the first item series there are two digits and for each series the number of digits to remember increases by one. When the subject provides the correct response in four items in a series, the examiner moves on to the next. The stop rule is three incorrect responses within one series.

3.7 Validity and reliability

The validity system I use as reference in the following was introduced by Cook and Campbell (1979).

3.7.1 Internal validity

Internal validity refers to the question of causal relationships; to what degree are the conclusion of a causal relationship probable based on the data analysis. A randomised effect study with a random sample receiving an intervention with a matching random sample that functions as a control group is a design that could generate valid conclusions about a cause and effect relationship. In this study the design is not experimental and there is only one group and it has not received any intervention as part of the study. In addition, the condition of a random sample has not been met in this study, as only a limited portion of the population had the opportunity to participate and these were not chosen randomly.

3.7.2 Statistical validity

Statistical analysis is the tool for analysing the results in this study, hence it is relevant to consider the question of statistical validity. Statistical validity refers to what degree one can make a general assumption that there is a relationship between the dependent and independent

variable, or the observed tendency, based on the results being statistically significant (Lund, 2002). In this study it is relevant to consider whether the potential significant findings regarding the differences in the development across the different morphological awareness tests are likely to have been affected by threats related to statistical validity.

With a high statistical validity, the risk of making type errors is reduced, which means making the wrong conclusion based on the statistical results. The type errors are the type I-error, where a correct null-hypothesis is discarded, and type II-error where a false null-hypothesis is accepted (Lund, 2002). The probability or risk of making a type I error is related to the decided significance level, e.g. if the significance level is $p < 0,5$, the risk of making a wrong conclusion in less than 5%. Considering the risk of making a type II error, one has to take the statistical strength into account, which is related to the size of the sample and the found effect.

3.7.3 Construct validity

In measurements an important aspect is the construct validity; to what degree one has succeeded in operationalising the theoretical constructs to indicators that can be measured (Kleven, 2002a). Considering that morphological awareness is a compound in the broader construct linguistic awareness, a challenge can be how other abilities, both linguistic and cognitive, affect how well the targeted construct is measured. In the validation of the tests this study is based on, García Grande (2018) found through an exploratory factor analysis that these tests seem to be a good measurement of the same construct.

3.7.4 Reliability

The reliability of an instrument depends on whether the results can be interpreted consistently across situations (Field, 2016) and refers to the extent data is free of measuring errors (Kleven, 2002a). In the context of testing we are concerned with whether the tests consistently measure what it should measure, in other words: if the measurements are consistent a person will score nearly the same with repeated testing, given that they have not changed or developed. For considering the reliability one can use statistical tools in addition to considering the existence of random measurement errors in the test-administration.

3.7.5 External validity

External validity can help in evaluating whether we can generalise results beyond the participating population and time and context the research took place in (Lund, 2002). To consider whether results can be generalised, we need to consider how representative the sample is, and the time and the actual context the results are based on. In the case of this study, the goal is to able to say something about how children with Norwegian as their first language develop morphological awareness, thus the question about external validity and generalisation is important. To consider whether a sample is representative for a population, one has to consider the sampling process, the homogeneity within the sample and the sample size (Lund, 2002).

3.8 Ethical considerations

The NumLit project is approved by the Norwegian Center for Research data (NSD) and has collected parental consent from the participating children's parents. Each participant gave an oral assent at both measurement points. The examiners both last year and this year are students at either the Department of Education, Department of Special needs education or Department of Psychology at the University of Oslo who had received training in advance of the testing. Testing children as a group puts higher demand on the researcher to consider the ethical aspects of the participation. In the case of NumLit the examiners were all students who are familiar with children. The tests were specifically made and chosen for their age, which is in line with the national guidelines of ethical practice in research (NESH, 2016). The demand of giving the research objects information they need to give an informed consent can be difficult with children, but the information is to be adjusted to the age of the participants. The children are also to be informed that they at any point can withdraw from the project (NESH, 2016). To protect their privacy the names are replaced by ID-numbers on the test-documents and in the datafiles used for plotting the results.

3.9 Repeated-measures design

3.9.1 ANOVA

Analysis of variance, ANOVA in short, is used when we want to compare more than two means. If we were to compare only two means, a t-test is the appropriate test to perform whereas when there are more than two means the risk of making type I-errors is increased

when using a t-test. For instance, performing three separate t-tests will increase the probability of making a type I-error from 5% to 14,3%, which is a level that is too high according to criterion accepted in social science (Field, 2009). Similar to the t-test, the null-hypothesis in ANOVA is that there is no difference between the means, but an ANOVA will produce the F-statistic instead of the t-statistic. The F–ratio must be at least 1 for the result to be significant, because the mean of the mean square between groups must be larger than the mean square error (F=MSG/MSE). MSG is the variance between groups and the MSE in the variance within each group (Cetinkaya-Rundel, Diez, & Barr, 2013).

3.9.2 Repeated measures ANCOVA

In a repeated measures ANOVA the different levels of independent and dependent variables have to be defined. In this study the independent variable is time and in the analyses this is labelled as grade level. This is the one of the main levels and consists of two levels in this study; kindergarten and first grade. The other main level is the dependent variable, namely the results of the morphological awareness tests, which consists of three tests, and thus make up three levels. In this repeated measures analysis, we are interested in finding out whether there is a significant difference between the means of the tests, whether there is a significant difference in means across grades and whether grade and test interact significantly with each other. Further, we are also interested in the influence of working memory and vocabulary on the two main levels and whether the working memory and/or vocabulary affect the interaction between the two main levels. Working memory and vocabulary will thus be used as covariates, making the analyses a repeated measures ANCOVA. If the scores on the covariates vary systematically with the differences in the means on the dependent variables, it means that some of the variation within the morphological awareness variables may be significantly related to these covaries.

3.9.3 Assumptions for performing repeated measure ANCOVA

As ANCOVA is a parametric test based on the normal distribution there are some assumptions that must be true.

Independence of cases in the assumption that the data generated from one participant should be independent of that from another participant. However, in a repeated measure we expect the data from one person will not be independent in the different measures, since they are the

same person, but the data between the participants must nevertheless be independent of each other (Field, 2009). Normality is the assumption of having data that are normally distributed. Even if ANCOVA is relatively robust regarding the assumption of normality, deviations may affect the results (Field, 2009). Homogeneity of variance refers to the variance being stable across variables, which means that this assumption is not met when there is large variance across the variables in a study. ANCOVA is a test that is relatively robust to violations of this assumption when the sample sizes are equal (Field, 2009), which is the case in the present study. The assumption of homogeneity will for this reason not be attended further. An assumption that is not relevant in a regular ANCOVA, but relevant in repeated measures ANCOVA, is sphericity. This refers to the equality of variances of the differences between independent levels (Field, 2009). Sphericity can be checked in SPSS Mauchly's test, which is done simultaneously as the repeated measure ANCOVA. If this test's statistic is significant ($p < .05$) it should be concluded that the condition of sphericity is not met (Field, 2009). When the assumption of sphericity is not met, corrections can be used, which are also provided in the output in SPSS when doing repeated measures ANCOVA.

4 Analysis

4.1 Descriptive statistics of the different variables

The final sample used in this study consists of 187 subjects. Among these there were some missing values which have been excluded from the analyses by filtering them out. This left 174 cases, which the analyses in the following are based on.

A descriptive analysis provides an overview of the different variables and their distribution. Kurtosis and skewness are important to determine the normality of the distribution. A positive kurtosis means that there is an overweight of values in the tails and a negative kurtosis means that there are too few values in the tails. In a distribution that is skewed there is an overweight of either low or high values. In cases where there is an overweight of low values, this is a positive skew and the distribution will have a longer right tail. A negative skew, on the other hand, will have a longer left tail, as a result of too many high values in the distribution. A value of zero on kurtosis and skew reflect a normal distribution, and the higher these values deviate from 0, the less normal is the distribution (Field, 2016). Outliers, which are values either very much higher or lower than the other values, have in cases where they are thought to have an effect on the results been windsorised; a technique in which the outliers' score is changed to the next highest score plus one (Field, 2009). Information about outlier that have been adjusted in included in the summary of the variables.

Cronbachs alpha (α) and Revelles omega (ωt) are measures of internal consistency in the different variables. These measures give information about whether the tasks within a test correlate. The reliability coefficients are values between 0 and 1, and a value of .7 to .8 is considered an appropriate for Cronbachs Alpha (Field, 2009), which is the same as for Revelles Omega total. The reliability coefficients presented in the following are the results from the reliability analyses from the previous measurement and is therefore based on the sample of 246 cases. The reliability coefficients from the latest measurement is also based on the full sample because these were acquired at a later point than the main analyses.

Table with the overview of the variables

Variable	*N*	*Med*	*M*	*SD*	*Skew*	*Kurtosis*	*Cronbachs α*	*Revelles ωT*
MA epi-inflectional K	174	11	11.14	2.45	-0.310	-0.470	.55	,67
MA meta-inflectional K	174	7	6.95	3,78	0.263	-0.618	.83	.85
MA meta-derivational K	174	5	4.61	2.15	0.419	0.068	.61	.65
MA epi-inflectional G1	174	13,5	13.14	1.81	-1.119	1.469	.55	.66
MA meta-inflectional G1	174	11	10.29	1.11	-0.747	-0.317	.87	.89
MA meta-derivational G1	174	5	5.45	2.40	0.123	0.083	.61	.67
BPVS_K	174	61.5	62.82	13.05	-0.551	2.223	.90	.95
Backward digit span_K	174	6	5.54	3.39	0.009	-0.340	.88	.92

Table 1 *Descriptive statistics*

4.1.1 The variable epi-inflectional from kindergarten

The histogram in figure 8 show the distribution of the scores from the epi-inflectional test in kindergarten.

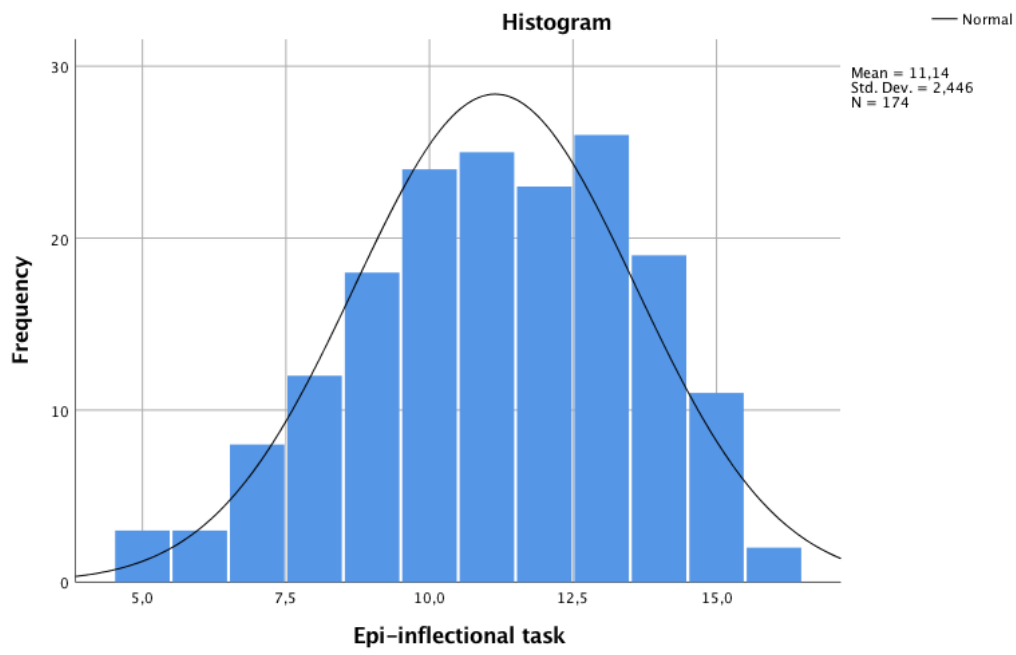


Figure 8 *Histogram of the Epi-inflectional task in kindergarten*

The histogram shows that the distribution has slight skew to the left -0.310. The mean is 11.14 and the median 11. The kurtosis is -0.470 and as we can see from the histogram, there are few values in the left tail. The reliability coefficients are .55 with Cronbachs Alpha and .67 with Revelles Omega total.

4.1.2 Epi-inflectional, first grade

Figure 9 illustrating the distribution of this variable show a clear tendency of higher scores and this distribution has a left skewe, with a value of -1.119. This reveals a ceiling effect in this variable.

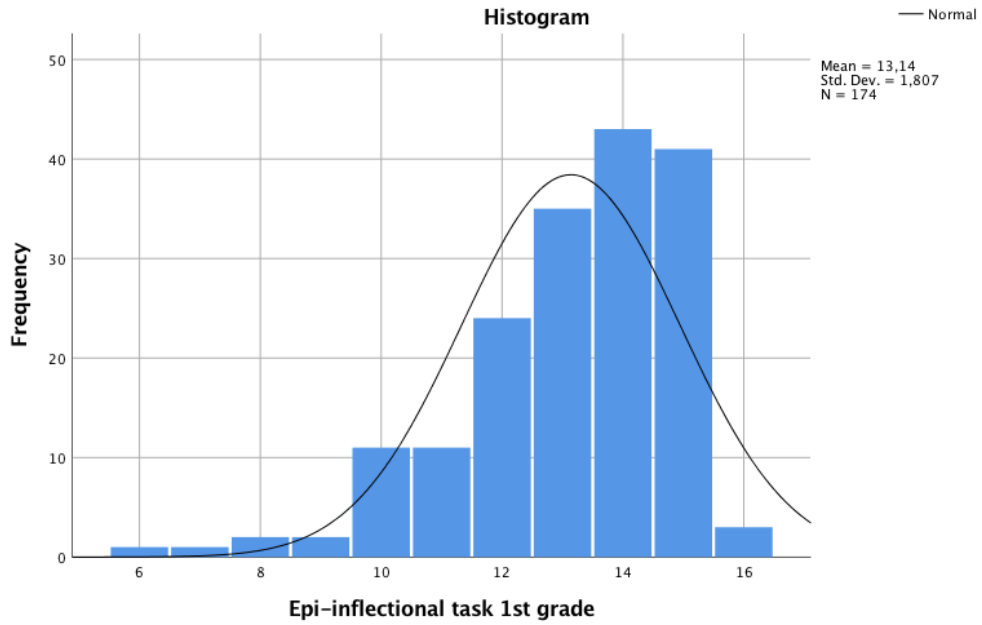


Figure 9 *Histogram of the Epi-Inflectional task in 1st grade*

The kurtosis is 1.469. The mean is 13.14 and median is 13.5, which is close to the maximum score of 16. The mean has increased with two points compared to the previous measurement. The reliability coefficients for this test in first grade are .55 with Cronbachs Alpha and .66 with Revelles Omega total.

This variable was considered transformed because of the ceiling effect, but a transformation of one variable would require a transformation of all the variables since the variables have to be in the same format in an ANCOVA. In turn, this would lead to results that would not be interpretable.

4.1.3 The variable meta-inflectional, kindergarten

In the histogram in figure 10, we can see a distribution where the total span of scores is represented. The values are slightly centred in the middle and the kurtosis is -0.618.

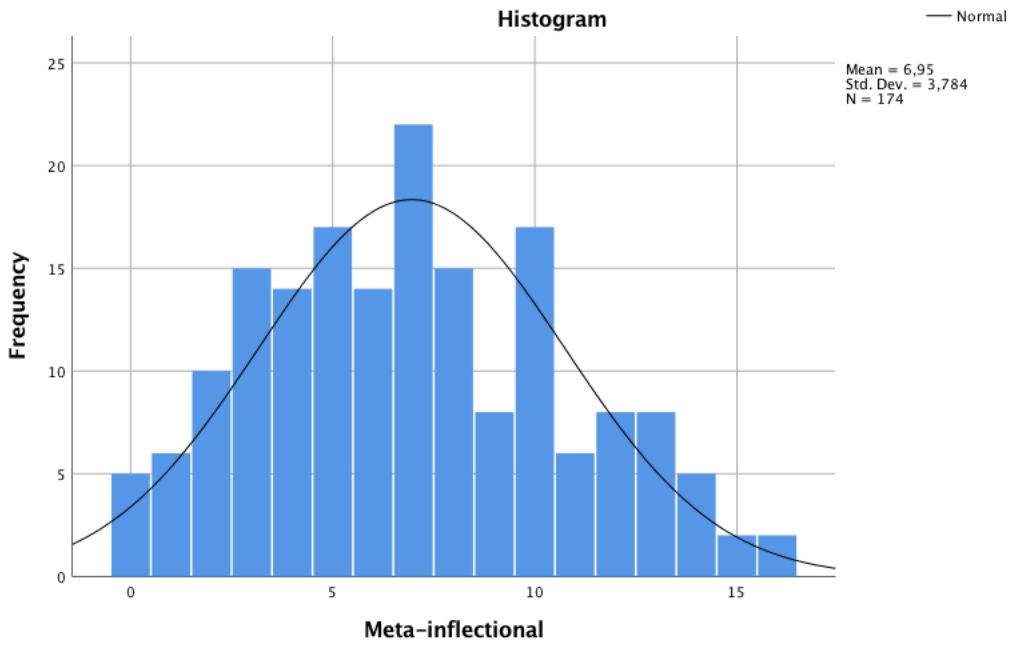


Figure 10 *Histogram of the Meta-inflectional task in kindergarten*

The mean on this variable is 6.95 and the median 7. The skew is 0.263, which indicates a slight right skew. The Cronbach's Alpha value is .83 and the Revelles omega total is .85 on this test (García Grande, 2018).

4.1.4 Meta-inflectional first grade

In figure 11, we can see a tendency of high scores in the distribution of the meta-inflectional test in first grade.

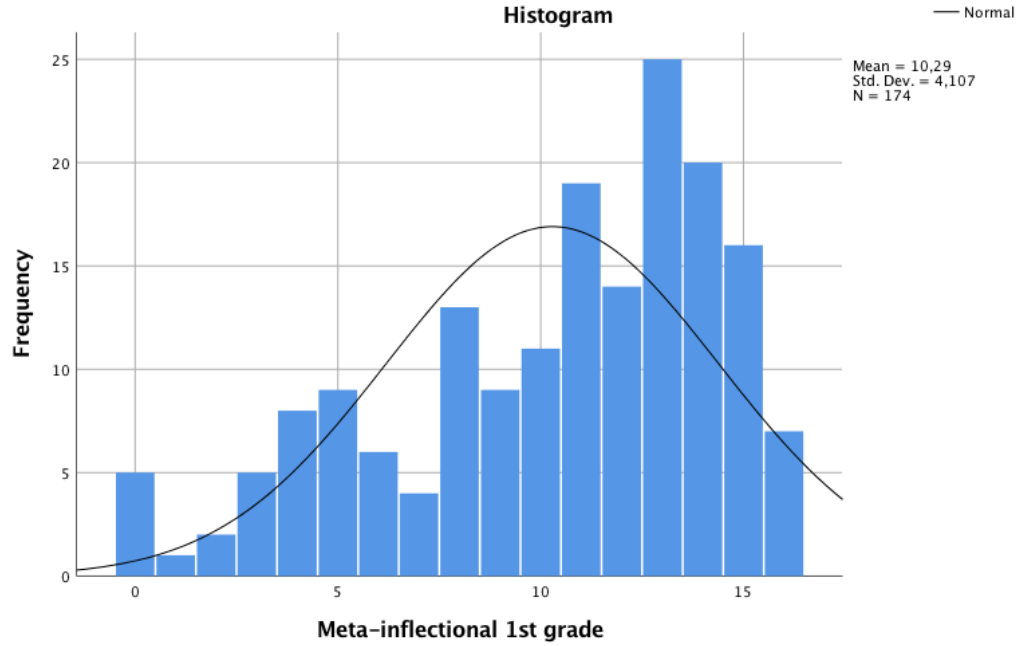


Figure 11 *Histogram of the Meta-inflectional Task in 1st grade*

The mean on this task has increased from 6.95 in kindergarten to 10.29 in first grade and the median is now 11. There is a left skew of -0.747, and the kurtosis is -0.317. The tests of reliability gave a Cronbachs Alpha value of .87 and .89 with Revelles Omega total.

4.1.5 Meta-derivational kindergarten

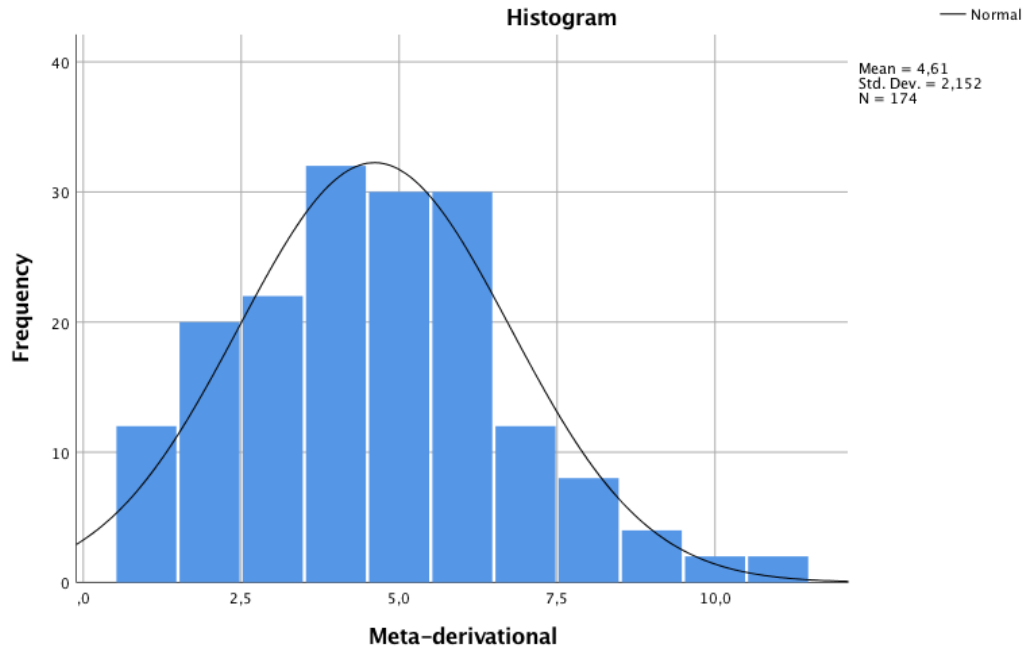


Figure 12 *Histogram of the Meta-Derivational Task in kindergarten*

The results from the meta-derivational test, seen in figure 12, show a span in scores from 1-11 points. This means that none of the children scored the maximum of 14 points. The mean is 4.61 and the median is 5. There is a slight right skew 0.419 and the kurtosis is 0.068, which indicate that the values are quite centered. This test has a Cronbach's Alpha value of .61 and a Revelles Omega total value of .65 (García Grande, 2018).

4.1.6 Meta-derivational first grade

The histogram showing the distribution of this variable in figure 13 shows that the maximum amount of points acquired is 12, which means that there are still none who has gotten the maximum score of 14.

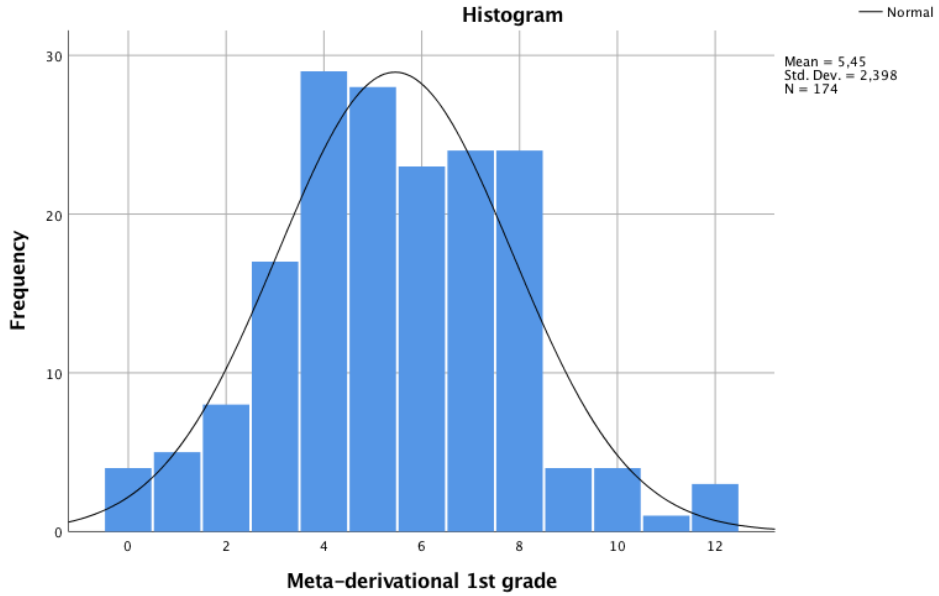


Figure 13 *Histogram of the Meta-Derivational Task in 1st grade*

The mean of this variables is 5.45, which is an increase of 0.84 points from the mean in kindergarten. The median is still 5. The skew is 0.123 and kurtosis 0.083, which are values that indicate that this distribution is close to normal. The Cronbachs Alpha reliability test indicate a value of .61 and with Revelles Omega total a value of .67.

4.1.7 BPVS in kindergarten

Figure 14 shows the distribution of the scores on the vocabulary test BPVS in kindergarten. The mean is 63.16 and the median 61.5. The skew in this distribution is 0.220

and the kurtosis is -0.657, which indicate that there are relatively many values centered in the middle. As there were two outliers with very low scores, these were windsorised in order to reduce the impact of these scores. Reliability values are .90 with Cronbachs Alpha and .95 with Revelles omega (Wimmer, 2018)

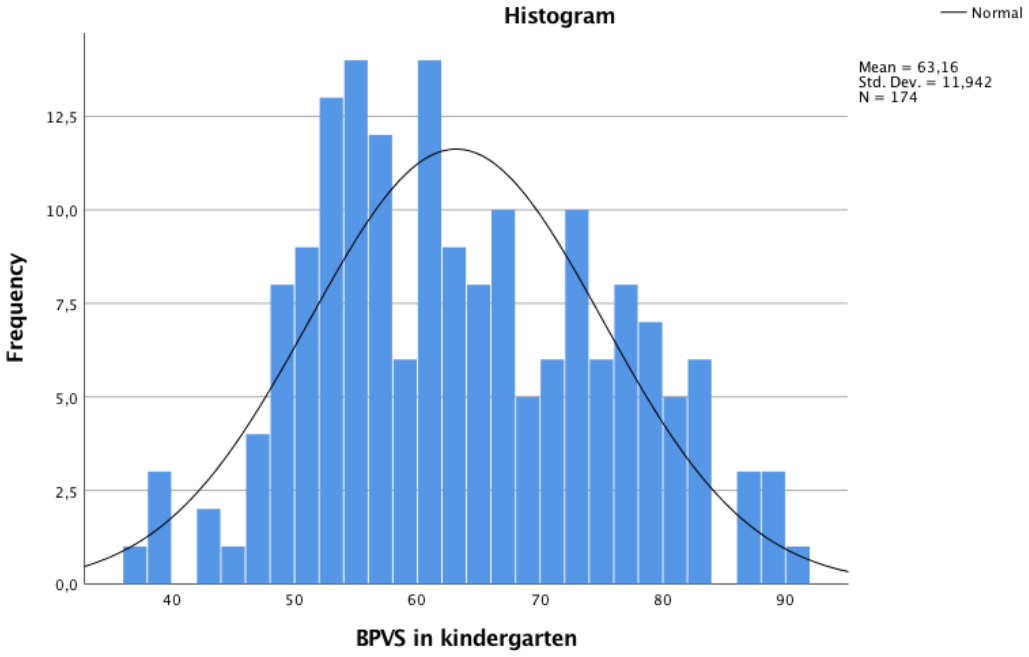


Figure 14 *Histogram of BPVS in kindergarten*

4.1.8 Backward digit-span

In figure 15 we can observe the distribution of the scores on Digit Span Backwards in kindergarten.

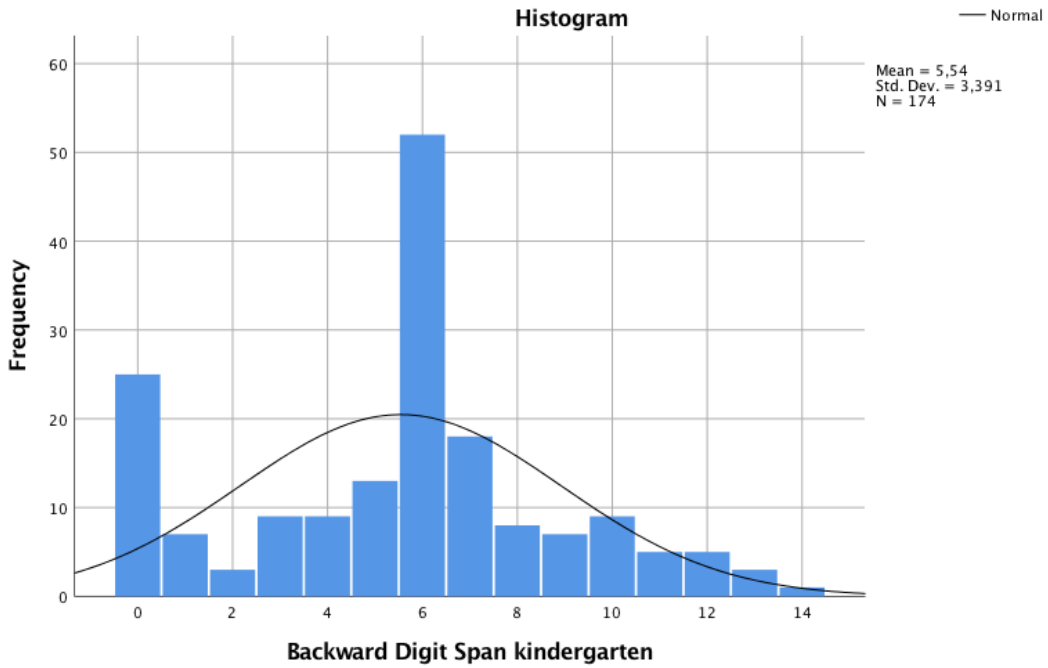


Figure 15 *Histogram of Backward Digit Span in Kindergarten*

The mean is 5.54 and the median is 6 and as we can see in the histogram, many acquired the score of 6. The skew is 0.009 and the kurtosis is -0.340. Reliability tests provided the values .88 with Cronbachs Alpha and .92 with Revelles Omega total (Wimmer, 2018).

4.1.9 Test of normality

Table 2 shows the Shapiro-Wilk test for all of the variables and give information about whether the variables are normally distributed according to this test.

	Shapiro-Wilk		
Test	**Statistic**	**df**	***p***
Epi-inflectional_K	0.970	174	.001
Meta-inflectional_K	0.976	174	.004
Meta-derivational_K	0.962	174	<.001
Epi-inflectional_G1	0.894	174	<.001
Meta-inflectional_G1	0.924	174	<.001
Meta-derivational_G1	0.976	174	.005
BPVS_K	0.952	174	.001
Backward digit span_K	0.937	174	<.001

Table 2 *Shapiro-Wilk test of normality*

As all the p-values show significant values ($p < 0.05$), the assumption of normality is not met. The presented distributions also show deviation from the normal distribution to various degrees. In general ANOVA and the associated F-value, are relatively robust to non-normality when the group sizes are equal (Field, 2009), which they are in this study.

4.2 Analyses

In this study repeated measures ANCOVA is the main analysis. The aim of this analysis is to examine how the different superset levels in the ANCOVA interact within and with each other.

The three morphological awareness tests have different number of items, therefore raw scores from them lie on different scales (0-16 vs. 0-14). In order to perform a repeated measures analysis, however, the scores have to be equally proportioned across all levels of the dependent variables. Therefore, prior to the analyses, the scores were divided by the number of items in the tests. Furthermore, the covariates were centred. This is a process where a variable is transformed into deviations around a fixed point (Field, 2009). The BPVS and Backward digit span scores were transformed into z-scores, where 0 is the fixed point and mean. Centring is used in regression in order to have more interpretable results of the interactions, and as ANCOVA is a version of regression (Field, 2009), centring is also relevant here.

4.2.1 Mauchly's Test of Sphericity

Table 3 shows that the assumption of sphericity only holds for the test level. For this reason, I will use the Huyhn-Feldt correction in the analyses where the assumption of sphericity is violated.

Within-subjects effect

	df	***p***
MA tests	2	0.076
Grade	0*	.
MA tests*Grade	2	0.018

*the sphericity cannot be measured when there are less than three levels, which is the case for grade

Table 3 *Mauchly's test of sphericity*

4.2.2 Tests of within-subjects effects

The table of within-subjects effects (table 4) shows the main results of the ANCOVA. When the appropriate correction is chosen, the column *p* indicates whether the results are significant. If the *p*-value is less than 0.05, the means of the specific levels are significantly different. The superset levels are the morphological awareness tests and grade level. The morphological awareness tests consist of three subset levels, since this number is equal to the number of tests. Grade level includes kindergarten and first grade, and thus consists of two subset levels. In the analyses the first level that is analysed is the test-level. This means that the overall mean for each test are compared with each other. In addition, this analysis of difference will indicate whether the difference in the tests is affected by any the covariates. Grade-level is the comparison of the total mean of all three tests from kindergarten and first grade. In the interaction analysis where the subset levels from both superset levels are compared, will indicate whether grade level has affected the different tests differently or whether the increase in test score means is equal for all three tests.

Level	***Correction***	**df**	***F***	**p**	η_p^2
MA tests	Sphericity Assumed	2	538.156	<.001	0.759
MA tests * BPVS_K	Sphericity Assumed	2	17.202	<.001	0.091
MA tests * Digit span_K	Sphericity Assumed	2	4.494	0.012	0.026
Grade	Huynh-Feldt	1	166.021	<.001	0.493
Grade * BPVS	Huynh-Feldt	1	1.866	0.174	0.011
Grade * Digit Span	Huynh-Feldt	1	1.607	0.207	0.009
MA tests * Grade	Huynh-Feldt	1.956	29.244	<.001	0.146
MA tests * Grade * BPVS_K	Huynh-Feldt	1.956	0.279	0.752	0.002
MA tests * Grade * Digit Span_K	Huynh-Feldt	1.956	0.346	0.703	0.002

Table 4 *Tests of within-subjects effects*

Table 4 shows, as noted, determine which effects are significant. We will go into further detail about which levels the significant effects hold after presenting the within-subjects contrasts (table 5). First of all, in table 4, we can see that there are significant differences among the overall means of the three morphological awareness tests, indicating that the three tests are not equally difficult. These differences between scores in the morphological awareness tests (averaged across grades) are associated with differences in kindergarten

vocabulary and working memory scores. Specifically, children who scored higher on working memory and vocabulary, scored on average higher on morphological awareness tests than those with a lower working memory and vocabulary score. I will return to this observation in section 4.2.3. Further, the grade level affects the overall mean of all the tests, this means that the overall mean for morphological awareness is significantly different between the two grade levels. This is to be expected, as normally developing children should have some development of their linguistic abilities from kindergarten to first grade. Moreover, there is an interaction between the differences seen between test and grade level, which means that the increase in score from kindergarten to first grade is not equal for all the three tests. This interaction is not significantly affected by differences in vocabulary or working memory scores which implies that the difference in how much the different morphological awareness test scores improved between kindergarten and first grade did not depend on working memory or vocabulary level.

4.2.3 Test of within-subjects contrast

The analysis of within-subjects contrasts compares different levels to each other, in order to have more information about between which subset levels there is an effect. In the analyses the contrast “difference” was chosen. This contrast gives sufficient information to further interpret the significant effects found in the within-subjects effect.

Level	*Test*	*Grade*	*df*	*F*	p	η_p^2
MA tests	Meta-inflectional vs. Epi-inflectional		1	282.31	<.001	0,623
	Meta-derivational vs. Meta-inflectional		1	886.332	<.001	0.838
MA tests * Digit Span_K	Meta-inflectional vs. Epi-inflectional		1	1.817	0.179	0.011
	Meta-derivational vs. Meta-inflectional		1	8.138	0.005	0.045
MA tests * BPVS_K	Meta-inflectional vs. Epi-inflectional		1	28.593	<.001	0.043
	Meta-derivational vs. Meta-inflectional		1	1.702	0.194	0.010
Grade		First grade vs kindergarten	1	166	<.001	0.493
Grade * Digit Span_K		First grade vs kindergarten	1	1.607	0.297	0.009
Grade *BPVS-K		First grade vs kindergarten	1	1.866	0.174	0.011
MA tests * Grade	Meta-inflectional vs. Epi-inflectional	First grade vs kindergarten	1	17.481	<.001	0.091
	Meta-derivational vs. Meta-inflectional	First grade vs kindergarten	1	42.615	<.001	0.199
MA tests * Grade * Digit Span_K	Meta-inflectional vs. Epi-inflectional	First grade vs kindergarten	1	0.001	0.972	0
	Meta-derivational vs. Meta-inflectional	First grade vs kindergarten	1	0.739	0.391	0.004
MA tests * Grade * BPVS_K	Meta-inflectional vs. Epi-inflectional	First grade vs kindergarten	1	0.251	0.617	0.001
	Meta-derivational vs. Meta-inflectional	First grade vs kindergarten	1	0.311	0.578	0.002

Table 5 *Tests of within-subjects contrasts*

From table 5 it can be seen that there is a significant difference in performance (averaged across kindergarten and first grade), between the tests meta-inflectional and epi-inflectional and between meta-derivational and meta-inflectional. The contribution of working memory is significantly associated with the difference between the overall means in meta-derivational and meta-inflectional tests, but not with the difference between meta-inflectional and epi-inflectional. Figure 16 illustrates this interaction between the morphological awareness tests and working memory. This figure shows that the meta-inflectional test score is significantly more affected by a higher working memory score compared to the meta-derivational test

score. In order to create this graph for visualising and further interpret this significant effect the scores from the covariate was recoded into two a new variable via a median split, with low scores as one group (1) and high scores as another (2).

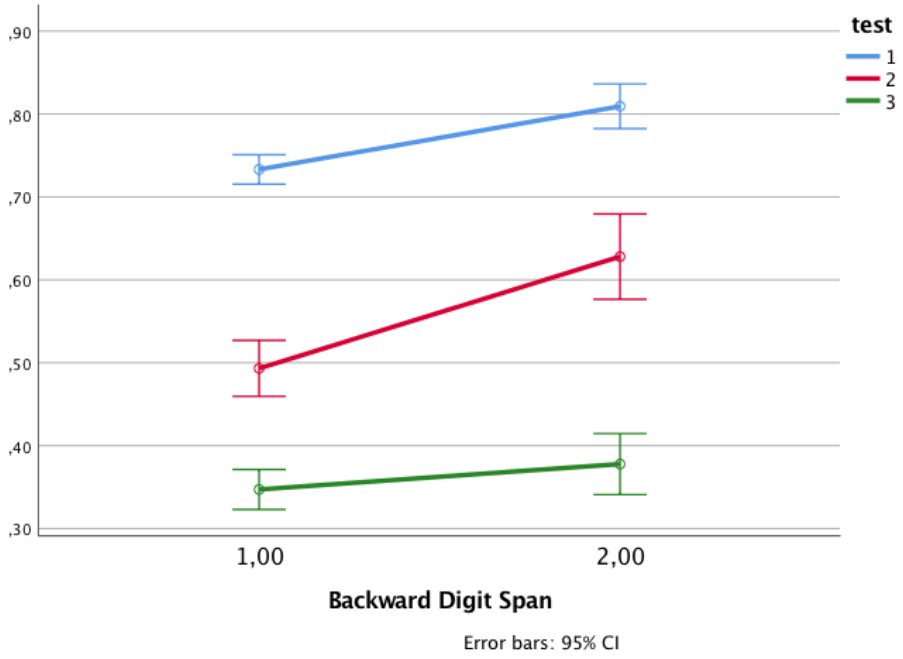


Figure 16 *Graph illustrating the interaction between morphological awareness tests and Backward Digit Span. The blue line is the epi-inflectional, red is the meta-inflectional and blue is the meta-derivational test.*

Vocabulary is significantly associated with the difference in mean performance between meta-inflectional and epi-inflectional, but not with the difference between meta-inflectional and meta-derivational. In figure 17, a graphic illustration of this interaction can be seen, which shows that the meta-inflectional test score is significantly more affected by a high vocabulary score than the score on the epi-inflectional test. This variable was also recoded into two groups representing low (1) and high (2) scores via a median split.

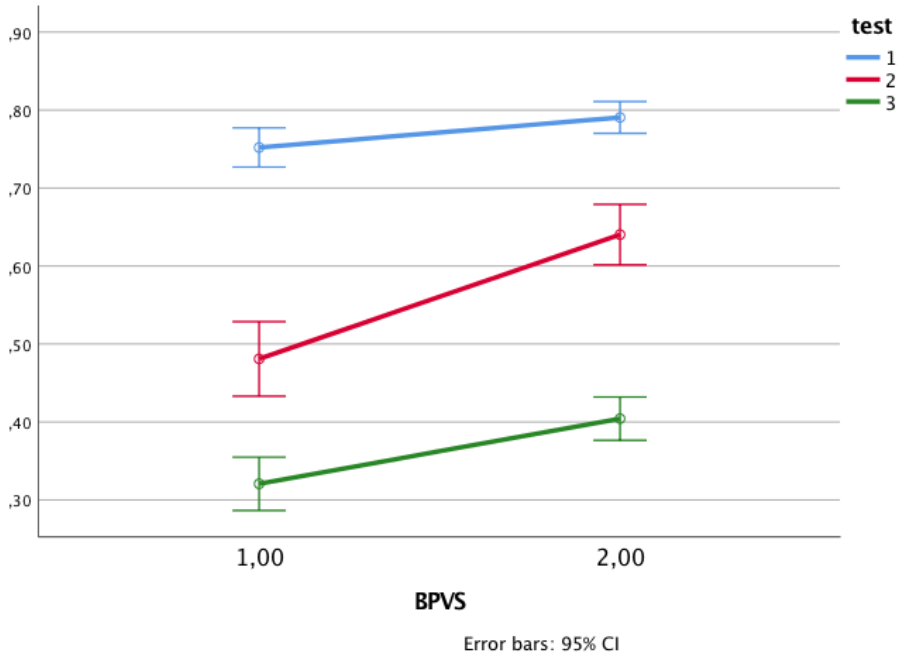


Figure 17 *Graph illustrating the interaction between the morphological awareness tests on BPVS. The blue line is the epi-inflectional, red is the meta-inflectional and blue is the meta-derivational test.*

Furthermore, in table 5, we can see that the interaction between test and grade level is significant the for meta-inflectional and epi-inflectional test scores and for the meta-inflectional and meta-derivational test scores. This means that there is a significant difference in how much the means of these tasks changed between grades. A graph will in this case provide further information about the significant differences by illustrating which test-mean in the comparisons increased the most. In figure 18, we can observe that grade level has had a larger effect on the meta-inflectional test score, compared to both the other test levels; (a) the epi-inflectional test score and (b) the meta-derivational test score.

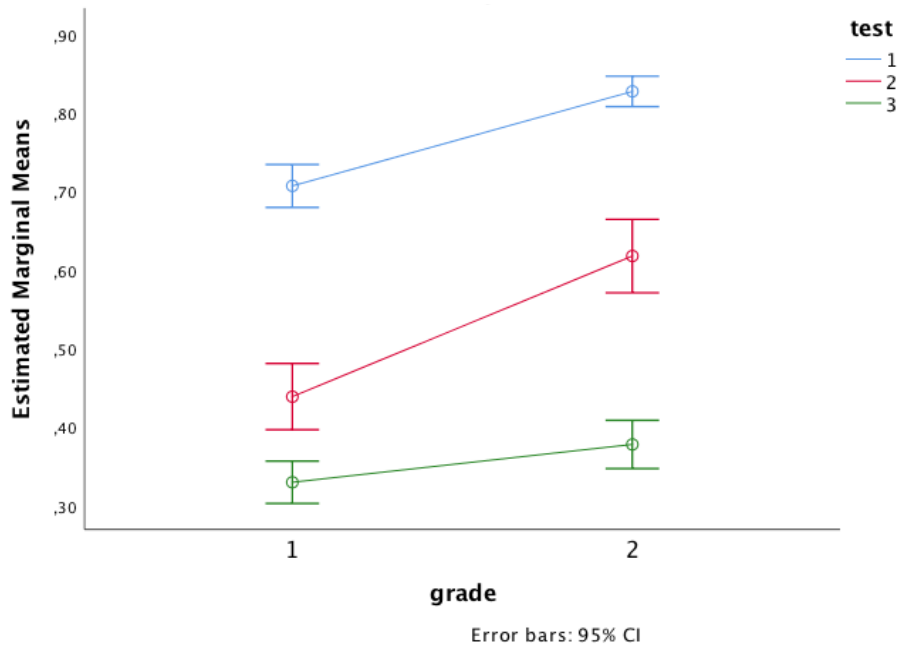


Figure 18 *Graph illustrating the interaction between tests and grade. The blue line in this graph represents the epi-inflectional test, the red the meta-inflectional test and the green line represents the meta-derivational test. On the horisontal axis, 1 represents kindergarten, and 2 represents first grade.*

5 Discussion

5.1 Results in light of theory

5.1.1 Development of morphological awareness

The aim of this study was to examine whether different domains and levels of morphological awareness develop to the same extent from kindergarten to first grade. It appears that the mean score of the meta-inflectional task has increased significantly more than the means of both the epi-inflectional and meta-derivational tasks. The difference in improvement between the epi- and meta-inflectional level may be related to the epi-inflectional task measuring a level of awareness that precedes the meta-level (Gombert, 1992). The difference in the increase of mean score between the meta-inflectional and meta-derivational task is also in line with previous research conducted in other languages. Awareness of derivational morphology has been found to develop at a later age and over a longer time-span than inflectional morphology (Berninger, Abbott, Nagy, & Carlisle, 2010; Diamanti, Benaki, et al., 2017; Kuo & Anderson, 2006; Selby, 1972). The findings in this study give reasons to believe that this is also the case in Norwegian, based on the specific materials used here for measuring these skills.

Epilinguistic control vs. metalinguistic awareness

The information about the development of different levels of morphological awareness in Norwegian is provided by the difference in increase between the epi-inflectional and meta-inflectional task. This result indicates a difference between the epilinguistic control task and the meta-awareness task indicate that meta-awareness develops at a later stage than epilinguistic control, at least according to the measurements used in this study. This is different from what Diamanti, Benaki, et al. (2017) found in Greek children, that both levels of awareness of inflections was developed at the time of measurement (which was from four to seven years of age). Diamanti, Benaki, et al. (2017) did, however, find a similar difference between the epi- and meta-level on derivational morphology, which indicated that the meta-awareness of derivations succeeds that of inflectional. Combined, these findings support the hypothesis that morphological awareness may exist on different metalinguistic levels. Moreover, reasons why the results from this study differ from the study on Greek children with regard to the inflectional tasks may be related to meta-awareness of inflections

developing at an earlier age in their sample, compared to the present sample. This, in turn, may be related to the fact that Greek children receive informal instruction of morphological awareness and phonological awareness in kindergarten (Diamanti, Benaki, et al., 2017), which seems to be an emerging focus in Norwegian kindergartens (Kunnskapsdepartementet, 2016). It may also be related to differences in the nature of the two languages, or perhaps a combination of the two. Finally, the possibility that the apparent difference in the levels of awareness in inflectional morphology observed in Norwegian and not in Greek is related to other aspects cannot be excluded. For example, differences in difficulty of the particular test items used may have contributed to these reported differences.

Regarding the significant difference in the increase in scores in the two inflectional tests, an explanation of why the epi-inflectional test did not increase as much as the meta-inflectional may be related to the observed ceiling effect in first grade on the epi-inflectional test. A ceiling effect may reduce the representativeness of the measured domain of interest. This test had an overall mean of close to 70%, hence there may have been smaller chances of increasing this mean further, compared to the meta-inflectional task where no ceiling effect was found. In any case, the finding suggests that meta-awareness of inflectional morphology develops at a later stage than epilinguistic control but does not give sufficient information about whether the epi-level of awareness is still under development.

Inflectional vs. derivational morphology

The significant difference between the development of the meta-derivational awareness and the meta-inflectional tests indicates that the development of awareness of derivational morphology lags behind that of inflectional. This result is similar to what Diamanti, Benaki, et al. (2017) found for Greek children. Further, this finding is supported in previous research, indicating that the acquisition of derivational morphology takes place later than the acquisition of inflectional morphology and that it develops over a longer time span (Anila et al., 2018; Kuo & Anderson, 2006; Tyler & Nagy, 1989). Some of the theories for why derivational morphology appears to have a later developmental pattern compared to that of inflectional include the large number of derivational suffixes, compared to inflectional ones (Kuo & Anderson, 2006). If this is also the case for Norwegian, it may be a reason for these results. Additionally, it seems plausible that children have not encountered a wide range of derivational forms at this young age. The theory of transparent derived forms being acquired before the more opaque forms, proposed for English (Carlisle & Nomanbhoy, 1993), seems

less relevant in Norwegian, but studies investigating the nature of transparency in Norwegian derivatives may be required to better understand this.

The substantial growth in meta-awareness of inflections in Norwegian, compared to the other tests, especially the meta-derivation test, may be supported by the findings by Ragnarsdóttir et al. (1999). These authors found that by eight years of age the past tense inflection was mastered for both regular classes of verbs, as well as for the strong verbs at over 90% of the items in the study. Between four and six years of age the number of correct responses increased from 51% to 72%, indicating that there is an increase in the number of verbs mastered with respect to past tense inflection in this age-span. In our study, the items in the meta-inflectional test included five verbs, of which three belong to the smaller verb group with the suffix and two belong to the larger verb group taking the inflectional suffix *-et* in the simple past. Even though information about the distribution of correct and incorrect responses and the differences from kindergarten to first grade is not included in this study, it might be that the acquired level of the knowledge of different inflectional patterns has influenced the ability to recognise the correct inflections of non-word verbs, with which the children were not familiar.

An alternative explanation of why the development of meta-derivational morphological awareness appears to lag significantly behind meta-inflectional awareness could be related to the nature of the items in the test targeting meta-derivational awareness. Based on previous research it can be assumed that some derived forms will be acquired later than others, and since this has not been previously explored in Norwegian, we lack the insight to know which is acquired around what age. In the validation of these tests, García Grande (2018) found in a qualitative analysis of the proportion of the correct answers that there were more correct responses on the items requiring the suffixes *-ing* and *-ete*, than on the items requiring the suffixes *-else* and *-ende*. García Grande (2018) suggested that this could be related to some derived forms being acquired prior to others. The point is that other derivational affixes might have given a different result. In any case, it will be interesting to see how awareness of meta-derivational morphology develops in the coming years of the NumLit study.

The effect of vocabulary and working memory skills

The significant interaction of high vocabulary with the difference between the epi-inflectional and meta-inflectional test may on one hand be surprising. One might rather have expected a

significant association of high vocabulary with differences from the meta-derivational test instead, considering that this is the only test where real words were used. The absence of such an interaction may however be taken as corroborating the finding that the meta-derivational task is more difficult because it measures a less developed domain of morphological awareness (Diamanti, Benaki, et al., 2017). The observation that a high vocabulary seems to explain more if the difference in the score of the meta-inflectional test than in the epi-inflectional test could be related to the observed ceiling effect on the epi-inflectional test. As previously noted, there is a possibility that the ceiling effect conceal a development of epilinguistic control, which in turn may have affected the result of a high vocabulary explaining the difference in meta-inflectional test significantly more.

Working memory was significantly associated with the difference between the overall means in meta-derivational and meta-inflectional tests. A high score on working memory seems to explain more of the difference in the score on the meta-inflectional test. This could be related to the non-words putting a higher demand on the working memory, since they are unfamiliar. On the other hand, this task was supplemented with pictures to reduce this possible load on the working memory, which in turn may indicate that this result is related to other unknown aspects.

The lack of a significant effect of vocabulary and working memory on the difference in longitudinal development between the three morphological awareness tests indicates that a high vocabulary or working memory cannot explain why these tests develop differently. This may indicate that performance in these tasks is not severely affected by vocabulary and working memory, which may further indicate that the measures have succeeded in tapping the desired domain, namely morphological awareness.

5.1.2 Synopsis of results in light of theory

The discussion indicates that the results from the present study are in line with previous research, which has found that inflectional morphology develops at an earlier stage than derivational morphology. This is also true for the levels of awareness, where the epi-level of awareness is assumed to precede that of meta-awareness. The results also indicate that the children are more aware of the morphological structure and that the found differences cannot

be explained by a general improvement in language skills, assessed by proxy via vocabulary, or because their ability to hold and handle material in memory has improved.

5.2 Results in light of validity

In order to further interpret the results, it is necessary to consider different aspects of the study that pose threats to inferring that the findings are valid. This also includes exploring alternative explanations to why the results are what they are. The validity system used for refence in the following is the one by Cook and Campbell (1979)

5.2.1 Internal validity

Internal validity refers to which degree one can draw conclusions about the causal relationship between the independent and the dependent variables in a study. This study has a non-experimental design, which means that this will have a lower internal validity than what can be acquired in experimental designs, which is the only design that can give clear answers about causal relationships between variables (Kleven, 2002b).

In a study with an experimental design maturation is considered a threat to the internal validity. This is related to the question regarding whether the found results are related to the intervention or maturation. However, in this study maturation cannot be regarded a threat in the interpretation of the results. Considering that the aim of this study is to understand how morphological awareness develops, maturation could rather be said to be the independent variable through time. Maturation in this context, may be related to metalinguistic development. Since metalinguistic awareness is required for reading and spelling (Gombert, 1992) and the first graders have had reading instruction, it is possible that this affects the development of morphological awareness, in addition to general language development. Based on the results of this non-experimental study, however, this is difficult to know for certain. Having said that, we can assume that the children have received a similar reading instruction, and hence not expect that different results are due to differences in the instruction, and thus not expect that it affects the results overall in terms of the threat of history to the internal validity.

Other threats to the internal validity relevant to consider in this study include testing and instrumentation. Testing refers to the retest effect and indicates that one has to consider

whether the subjects may have learned something from previous measurements (Lund, 2002). The subjects in this study were measured with the same tests in both kindergarten and first grade. However, considering that a year has passed, it seems unlikely that the results will be affected by the children having learned something from the first testing that they could apply in the second. Instrumentation is an aspect of the tests that may be responsible for the generated results (Lund, 2002). Ceiling and floor effects and systematic changes in the scoring criteria are factors that affect the internal validity. The results indicated ceiling effect on the variable epi-inflectional, and thus there is a possibility that this test does not represent the true results for the subjects.

In general, the internal validity is not considered to be high as a result of the design of this study. Among the threats that may have affected the results the ceiling effect on the epi-inflectional test appears to be most relevant.

5.2.2 Statistical validity

Statistical validity is related to the conclusions drawn about the relationship between the independent and the dependent variable. This coherence between the variables depends on the statistical significance and power. Statistical power is affected by factors such as sample size, the decided level of significance, directionality and effect size. Prior to statistical analyses, a statistical significance level is chosen, which in this case is $p < 0.05$. When the analysis indicates statistical significance, one has to consider the possibility of making a type I-error, in which cases a true null hypothesis has been rejected. The probability of making a type I-error is equal to the chosen significance level, which means that if the significance level is $p < 0.05$, there is a 5% chance that the null-hypothesis has been wrongfully rejected. In cases of non-significant effects, the risk of making a type II-error has to be considered in order to avoid rejecting a true alternative hypothesis.

In this study it is relevant to consider the risk of making a type I-error in cases where significant effects have been found. On the test of the difference in means across tests the results are significant within the chosen significance level. We could further see in the tests of within-subjects contrasts that the differences in the tests were significant for both contrasts, the epi-inflectional vs meta-inflectional and the meta-derivational vs meta-derivational. This finding is supported by the intention that the tests are meant to measure different aspects of

morphological awareness and be of different difficulty. I therefore consider this result unlikely to have appeared by chance.

The differences in mean performance on the morphological awareness tests across grade was significant at the chosen level. Considering that maturation should be a factor affecting the results, it would be suspicious if the overall mean result did not improve significantly in one year. For this reason, it seems that the risk of rejecting a true null-hypothesis in this case is also low. The interaction between test and grade also appear to be significant. As previously discussed, this is similar to what has been found in other studies of the development of morphological awareness, and therefore, the risk of having made a type I-error is considered low.

The risk of making a type II-error and rejecting a potentially true alternative hypothesis should be considered. With a larger sample, effects which have not been detected in this study might be found. As the non-significant results are all related to the effect of vocabulary and working memory these are the relevant effects to consider in this case. In line with the previous discussed potential confounding effect of working memory and vocabulary, which indicated that the design and difficulty of the tests may have reduced this effect, there does not appear to be a strong case for assuming that a type II-error has been made.

In the review of the assumptions for doing the analyses conducted in this study, we saw that not all of them were met. For example, the assumption of normality was not met for any of the variables. If we on the other hand look at the skewness and kurtosis of the variables, it was found that only one of the variables had a skew and kurtosis over 1. It should still be kept in mind that the violation of this assumption may have influenced the results, even though ANCOVA is relatively robust to violations of the assumption of normality.

Furthermore, another potential threat to the statistical validity is the relatively low reliability in two of the morphological awareness tests, namely the epi-inflectional task (Cronbachs Alpha .53 and Revelles Omega total .67 in kindergarten and Cronbachs alpha .55 and Revelles Omega .66 in first grade) and meta-derivational task (Cronbachs alpha 0,61 and Revelles omega 0,65 in kindergarten and Cronbachs Alpha .61 and Revelles Omega .67). This is related to the statistical validity because low test reliability reduces the statistical strength and consequently this should be considered a threat to the statistical validity in this study.

5.2.3 Construct validity

The construct validity in a study depends on the degree to which the measurements have succeeded in measuring what they are supposed to measure. It is related to the interpretation of the measurements because it is not the test itself that holds the validity (Kleven, 2002a). Construct validity is closely related to the reliability, which is concerned with whether the measurement has succeeded in measuring what it is supposed to measure but is not occupied with whether operationalisation has been successful or not. When discussing construct validity and reliability it is relevant to consider possible random and systematic measurement errors. Random measurement error is related to the reliability and systematic measurement errors to construct validity. However, it may be difficult to separate the two types of errors from each other, as reliability and construct validity are overlapping (Kleven, 2002a).

As noted, an issue with measuring morphological awareness is that it is a non-observable phenomenon and thus cannot the measured directly. Theories of which measurable traits represent a construct form the basis of the operationalisation of a construct. As presented in the theory chapter, there is some uncertainty related to what exactly morphological awareness includes. and maybe especially how implicit and explicit knowledge is manifested. García Grande (2018) found in his validation of the morphological awareness tests that they are associated with a single construct. The results from the reliability analysis showed that the items are related to one construct by 39%, while what the remaining 61% of the variance corresponds to is uncertain. This unexplained variance may be related to the different tests tapping different linguistic skills in addition to morphological awareness, which may not be correlated with each other. This is in line with the finding of the different effect of the covariates on the different tests. Overall it appears that we can assume that the morphological awareness tests tap the intended construct to some extent. Nonetheless, it is relevant to discuss what potential threats to construct validity can be identified in this study.

As previously mentioned, the reliability analysis indicates that there are other constructs beyond morphological awareness that are tapped by the morphological awareness tests. This is expected, as Kuo and Anderson (2006) point out that it might be impossible to obtain a measurement that is limited to only testing morphological awareness, since it is so closely related to other linguistic skills. They point to syntactic awareness, vocabulary, and working

memory as potential confounding variables in morphological awareness measurements and consequently aspects that may be a source of systematic measurement error. The role of syntactic knowledge is especially related to derivational morphology. In this study there are no means for getting information about the role syntactic knowledge may have played in the results of meta-derivational task. The confound possibly caused by vocabulary, on the other hand, has been addressed both by using non-words in the inflectional tasks and by using a vocabulary measure as a covariate in the statistical analyses. It proved to not influence the difference in the increase of test-scores across grade and thus does not seem to explain a significant portion of the development of morphological awareness with these morphological awareness tests. The load on working memory was tried to be reduced by using pictures in all of the tests. Considering that working memory was not associated with differences in the rate of development of performance in the three morphological awareness tests, this may indicate that the use of pictures had an effect. Furthermore, this could also indicate that the tests have succeeded in measuring the targeted construct. Overall, central confounding factors have been tried to be reduced, but for instance the effect of syntactic awareness is not known. It is also unknown whether results from other tests used as covariates would have given a different result.

Moreover, there is a chance that some of the items have been a source of measurement errors. First of all one item in the epi-inflectional task has a very low correlation with the other items, according to García Grande (2018). An explanation to why this item has a low reliability might be related to the picture not clearly depicting past, which it is supposed to, and rather than present, as it may be interpreted as (item 12). When the same picture and sentences are used in the meta-inflectional task the first sentence provided by the examiner and the picture gives the children a contrast to the target inflection. Since the item is used in both the epi-inflectional and meta-inflectional tasks, this could be the reason for keeping the item in the test even if it showed a low correlation with the other items in the first test.

A systematic error related to the meta-inflectional task may be pronunciation of the non-words. Even though these were made similar to the sound and composition of real Norwegian words, they were unfamiliar to the children. For instance, a child with some anxiety related to performance may be reluctant to give an answer when uncertain about the pronunciation. In such cases this may affect both measurement points. This is a question of whether the role of motivation could lead to a systematic error, but hopefully such errors were avoided in this

study. Finally, the possibility of variations in how the tests were conducted depending on the examiner, also pose a reliability threat. Even if all the examiners did receive mandatory training, potential difference in how the tests were administered cannot be excluded.
The construct validity in the covariates is considered to be high but taking into account that they are only meant to measure aspects of the larger construct they are a part of, they can be considered to be limited. BPVS only targets receptive vocabulary, but according to the manual it holds high construct validity (Dunn et al., 1997) on this component of the construct vocabulary. Backward Digit Span is standardised and included in the Wechsler Intelligence Scale for Children IV (WISC IV), and therefore considered to have a high construct validity.

5.2.4 External validity

External validity is related to the generalisability of the findings of the study. We are interested in which contexts, time and people the results are generalisable to. Factors that represent threats to the external validity is the sampling procedure, the sample's homogeneity and the interaction between the independent variable and individuals, situations and time.

In this study the sampling procedure was not random, which is a way to ensure the representativeness of the population. The sampling was done by having parents sign up their children through kindergartens in municipalities around the capital of Norway. If the population intended for this study is Norwegian kindergarten and first-graders with Norwegian as their first language, we have to consider whether the sample can be assumed to be representative for this population. According the project description of the longitudinal study, the area from which the children are located representative of the general population in Norway regarding parents' education level and income. Information about the background of the parents whose children are participating is however, not available, hence the possibility of selection bias cannot be excluded. Therefore, the sampling may reduce the generalisability of the study. When that is said, we have seen that the results are supported by previous research, which means that we cannot assume that the results we have found here will not be true for the same age-group in the larger population.

A high level of homogeneity in a sample reduces the representativeness of a sample, as opposed to a heterogenous group. In this study the sample is a result of convenience sampling (Gall, 2007). This refers to a sampling process where the sample was found on the basis of

convenience, in this case for example the location relatively close to the University of Oslo and the parents signing up for their children to participate. Further, there is no reason to believe that this is an especially homogeneous sample, but we cannot be certain about this. In addition, the sample can be categorised as relatively large, and this is also a factor that strengthens the external validity.

For these reasons the external validity is considered to be satisfactory and we can assume the results may be generalisable to the larger population. But, having said that, one should be careful generalising based on one the finding from one study, when the demands required to perform logic inferential statistics have not been met (Gall, 2007).

6 Conclusion

6.1 Implications of this study

Overall, the findings in this study indicate that Norwegian children follow a similar developmental course as what is found to be the case in other languages. The awareness of meta-inflectional morphology appears be the skill that develops most from kindergarten to first grade. Considering that awareness at the epi-level is hypothesised to be developed preceding awareness at the meta-level, the tests may indicate that this is the case. On the other hand, the development on epi-inflectional awareness may have been left undiscovered because of the observed ceiling effect on this measurement. The small increase in the meta-derivational test may provide supplementing evidence that awareness of derivational morphology has a development which spans over more years than inflectional. However, information about the nature of the development of derivations in Norwegian could give a clearer picture of whether other derived forms than those included in the used tasks would give a different result. Finally, the influence of vocabulary and working memory appears not to be related to the development of any specific morphological domain or level more than the others.

In conclusion, this study gives some information about the development of morphological awareness from kindergarten to first grade in Norwegian children and comparing studies will be necessary to further explore how robust these findings are.

6.2 Limitations

In addition to the noted possible limitations regarding the ceiling effect overserved on one test, the possibility that a different operationalisation of the construct morphological awareness and the different levels of awareness would have generated different results cannot be excluded. More research is needed to understand this better, in order to enable its use as a tool for potentially discovering developmental delays in language. This is also the case for the different domains of morphology. For instance, would more information about when the different inflectional forms in general are acquired possibly give a better background for identifying when this development is delayed.

Moreover, this study has not given any information about the development of the different levels of morphological awareness in the domain of derivational morphology. Further research could consider including this as a measurement to see if similar findings as for the inflectional domain, where epilinguistic control seem to develop before the meta-awareness. It would also be interesting to see whether epi-derivational control developed closer to epi-inflectional or meta-inflectional morphological awareness.

References

Anglin, J. M. (1993). *Vocabulary development : A morphological analysis*. Chicago: University of Chicago Press.

Anila, F., Desrochers, A., & Deacon, S. H. (2018). The Acquisition of Derivational Morphology in Children. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological Processing and Literacy Developement* (pp. 112-132). London & New York: Routledge, Taylor & Francis.

Arnbak, E., & Elbro, C. (2000). The Effects of Morphological Awareness Training on the Reading and Spelling Skills of Young Dyslexics. *Scandinavian Journal of Educational Research, 44*(3), 229-251. doi:10.1080/00313830050154485

Aronoff, M., & Fudeman, K. (2011). *What is morphology?* (2nd ed.). Malden, Mass: Wiley-Blackwell.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? In (Vol. 4, pp. 417-423).

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of communication disorders, 36*(3), 189-208.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation, 8*(C), 47-89. doi:10.1016/S0079-7421(08)60452-1

Bauer, L. (2003). *Introducing linguistic morphology* (2nd ed. ed.). Edinburgh: Edinburgh University Press.

Berko, J. (1958). The Child's Learning of English Morphology. *Word, 14*(2-3), 150-177. doi:10.1080/00437956.1958.11659661

Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in Phonological, Orthographic, and Morphological Awareness in Grades 1 to 6. *Journal of Psycholinguistic Research, 39*(2), 141-163. doi:10.1007/s10936-009-9130-6

Berthiaume, R., Bourcier, A., & Daigle, D. (2018). Morphological Processing Tasks and Measurement Issues. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological Processing and Literacy Development* (pp. 48-88). London & New York: Routledge, Taylor & Francis group.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: John Wiley & Sons.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010a). The Effects of Morphological Instruction on Literacy Skills: A Systematic Review of the Literature. *Review of Educational Research, 80*(2), 144-179. doi:10.3102/0034654309359353

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010b). The Effects of Morphological Instruction on Literacy Skills:A Systematic Review of the Literature. *80*(2), 144-179. doi:10.3102/0034654309359353

Burger, A., & Chong, I. (2011). Receptive Vocabulary. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development* (pp. 1231-1231). Boston, MA: Springer US.

Cain, K. (2010). *Reading Development and Difficulties* Chichester: Bps Blackwell.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological Aspects of language processing* (pp. 189-209). Hillsdale, NJ: Lawrence Erlbaum Associates

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing, 12*(3), 169-190. doi:10.1023/a:1008131926604

Carlisle, J. F. (2003). Morphology Matters in Learning to Read: A Commentary. *Reading Psychology, 24*(3-4), 291-322. doi:10.1080/02702710390227369

Carlisle, J. F. (2004). Morphological processes that influence learning to read. In C. A. Stone (Ed.), *Handbook of language literacy: Development disorders* (pp. 318-339). New York: Guilford Press

Carlisle, J. F. (2007). Fostering Morphological Processing, Vocabulary Development, and Reading Comprehension. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary Development - Implications for reading comprehension* (pp. 78-103). New York and Londin: The Guilford Press

Carlisle, J. F., & Goodwin, A. P. (2014). Morphemes Matter - How Morphological Knowedge Contributes to Reading and Writing In C. A. Stone, E. R. Silliman, B. J. Ehren, & G. P. Wallach (Eds.), *Handbook of Language and Literacy - Development and Disorders* (Second ed., pp. 265-283). New York: The Guilford Press.

Carlisle, J. F., & Nomanbhoy, D. M. (1993). Phonological and Morphological Awareness in First Graders. *Applied Psycholinguistics, 14*(2), 177-195. doi:10.1017/S0142716400009541

Casalis, S., & Louis-Alexandre, M.-F. (2000). Morphological Analysis, Phonological Analysis and Learning to Read French: A Longitudinal Study. *Reading and Writing: An Interdisciplinary Journal, 12*(3-4), 303-335. doi:10.1023/A:1008177205648

Cetinkaya-Rundel, M., Diez, D. M., & Barr, C. D. (2013). OpenIntro Statistics: an Open-source Textbook.

Clark, E. V. (1982). The young word maker: a case study of innovation in the child's lexicon. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 390-425). Cambridge: Cambridge University Press.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation : design & analysis issues for field settings*. Boston: Houghton Mifflin Co.

De Vaus, D. A. (2014). *Surveys in social research* (6th ed. ed.). London: Routledge.

Deacon, S. H., & Bryant, P. (2005). What young children do and do not know about the spelling of inflections and derivations. *8*(6), 583-594. doi:doi:10.1111/j.1467-7687.2005.00449.x

Deacon, S. H., & Bryant, P. (2006). This turnip's not for turning: Children's morphological awareness and their use of root morphemes in spelling. *British Journal of Developmental Psychology, 24*(3), 567-575. doi:10.1348/026151005X50834

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics, 25*(2), 223-238. doi:10.1017/S0142716404001110

DeCaro, M. S., & Maricle, D. E. (2011). Working Memory. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development* (pp. 1579-1581). Boston, MA: Springer US.

Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2017). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing, 31*(8), 1-25. doi:10.1007/s11145-017-9772-y

Diamanti, V., Benaki, A., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Development of early morphological awareness in Greek: Epilinguistic versus metalinguistic and inflectional versus derivational awareness. *Applied Psycholinguistics, 39*(3), 545-567. doi:10.1017/S0142716417000522

Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool Phonological and Morphological Awareness As Longitudinal Predictors of Early Reading and Spelling Development in Greek. *Front Psykol, 2039*(8). doi:10.3389/fpsyg.2017.02039

Dunn, L. M., Dunn, L., Whetton, C., & Burley, J. J. W., UK: NFER-Nelson. (1997). British picture vocabulary scale II.

Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Ann. Dyslexia, 46*, 209-240.

Field, A. (2009). *Discovering statistics using SPSS : (and sex and drugs and rock 'n' roll)* (3rd ed. ed.). Los Angeles: SAGE.

Field, A. (2016). *An adventure in statistics : the reality enigma*. Los Angeles: Sage.

Gall, M. D. (2007). *Educational research : an introduction* (8th ed. ed.). Boston, Mass: Allyn and Bacon.

García Grande, G. (2018). *Morphological awareness in Norwegian Pre-schoolers* (Master), University of Oslo, Oslo.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology, 40*(2), 177-190. doi:10.1037/0012-1649.40.2.177

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*(5), 887-898. doi:10.1037/0012-1649.28.5.887

Gjerde, T. (2018). *Sammenheng mellom lytteforståelse og eksekutive funksjoner hos enspråklige norske femåringer.* (Master), University of Oslo.

Gombert, J. É. (1992). *Metalinguistic development*. New York: Harvester Wheatsheaf.

Gonnerman, L. M. (2018). A Linguistic Analysis of Word Morphology. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological Processing and Literacy Development* (pp. 3-15). New Tork: Routledge.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*: Paul H Brookes Publishing.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23*(2), 95-147. doi:10.1016/0010-0277(86)90040-5

Kirby, J. R., & Bowers, P. N. (2017). Morphological instruction and literacy. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (Vol. 15, pp. 437-461). Amsterdam: John Benjamins Publishing Company.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Journal of Educational Psychology, 25*(2), 389-410. doi:10.1007/s11145-010-9276-5

Kleven, T. A. (2002a). Begrepsoperasjonalisering. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 141-182). Oslo: Fagbokforlaget.

Kleven, T. A. (2002b). Ikke-eksperimentelle design. In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 265-284). Oslo: Fagbokforlaget.

Ku, Y.-M., & Anderson, R. (2003). Development of morphological awareness in Chinese and English. *An Interdisciplinary Journal, 16*(5), 399-422. doi:10.1023/A:1024227231216

Kunnskapsdepartementet. (2016). *Tid for lek og læring — Bedre innhold i barnehagen*. Oslo: Regjeringen Retrieved from https://www.regjeringen.no/no/dokumenter/meld.-st.-19-20152016/id2479078/sec1

Kuo, L.-j., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist, 41*(3), 161-180. doi:10.1207/s15326985ep4103_3

Lie, S. (2011). *Norsk morfologi* ([2. utg.]. ed.). Oslo: Ling.

Lieber, R. (2016). *Introducing morphology* (Second edition. ed.). Cambridge: Cambridge University Press.

Lund, T. (2002). Metodologiske prinsipper og referanserammer In T. Lund (Ed.), *Innføring i forskningsmetodologi* (pp. 79-125). Oslo: Unipub

Lyster, S.-A. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing, 15*(3), 261-294. doi:10.1023/a:1015272516220

Lyster, S.-A. H., Horn, E., & Rygvold, A.-L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British picture vocabulary scale, (BPVS II). *Spesialpedagogikk, 74*(9), 35-43.

Lyster, S.-A. H., Lervåg, A. O., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing, 29*(6), 1269-1288. doi:10.1007/s11145-016-9636-x

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4), 1-182. doi:10.2307/1166115

Maynard, C., Brissaud, C., & Armand, F. (2018). The Acquisition of Inflectional Morphology in Children In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological Processing and Literacy Development* (pp. 91-111). New York & London: Routledge, Taylor & Francsis group.

McBride–Chang, C., Wagner, R., Muse, A., Chow, B. W.-Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics, 26*(3), 415-435.

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological Skills and Their Role in Learning to Read: A Meta-Analytic Review. *Psychological Bulletin, 138*(2), 322-352. doi:10.1037/a0026744

Muse, A. E. (2005). *The Nature of Morphological Knowledge.* Florida State University

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology, 98*(1), 134-147.

Nagy, W., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological Knowledge and Literacy Acquisition. *Journal of Learning Disabilities, 47*(1), 3-12. doi:10.1177/0022219413509967

NESH, D. n. f. k. (2016). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi

Pavey, E. L. (2010). *The structure of language : an introduction to grammatical analysis*. Cambridge: Cambridge University Press.

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*: Psychological Corporation.

Ragnarsdóttir, H., Simonsen, H. G., & Plunkett, K. I. M. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: an experimental study. *Journal of Child Language, 26*(3), 577-618.

Rønhovd, J. (1997). *Norsk morfologi* (2. rev. utg. ed.). Oslo: Ad notam Gyldendal.

Selby, S. (1972). The development of morphological rules in children. *British Journal of Educational Psychology, 42*(3), 293-299. doi:10.1111/j.2044-8279.1972.tb00722.x

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol., 94*, 143-174.

Smith-Lock, K. (1995). Morphological usage and awareness in children with and without specific language impairment. *Annals of Dyslexia, 45*(1), 161-185. doi:10.1007/BF02648217

Sparks, E., & Deacon, S. H. (2015). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics, 36*(2), 299-321. doi:10.1017/S0142716413000246

Sternberg, R. J. (1987). Most vocabulary is learned from context. In *The nature of vocabulary acquisition* (Vol. 89, pp. 89-105). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Tannenbaum, K. R., Wagner, R. K., & Muse, A. E. (2007). *Vocabulary acquisition : implications for reading comprehension*. New York: Guilford Press.

Templin, M. C. (1957). *Certain language skills in children : their development and interrelationships*. Minneapolis: Univ. of Minn. Press.

Tong, X., Deacon, S. H., Kirby, J. R., Cain, K., & Parrila, R. (2011). Morphological awareness: A key to understanding poor reading comprehension in English. *Journal of Educational Psychology, 103*(3), 523-534.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language, 28*(6), 649-667. doi:10.1016/0749-596X(89)90002-8

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence—fourth edition*. Bloomington, NM: Pearson.

Wimmer, A. M. (2018). Eksekutive funksjoner og reseptiv grammatikk. En kvantitativ studie om sammenhengen mellom eksekutive funksjoner og reseptiv grammatikk hos norske 5-åringer. In.

Appendix 1 Morphological Awareness Test- Final Version

Epi-inflectional Awareness Judgment Task-Final Version

NB! Se i manualen for uttale av non-ord

Instruksjon til barnet: "Vi skal se på noen bilder av skilpadder (vis heftet med bilder til barnet). Dukkene vil si hva skilpaddene gjør på hvert bilde på sitt eget, rare språk. Men bare én av dukkene sier det riktig. Du må høre godt etter og peke på den dukken som sa det riktig. Det er ikke den samme dukken hver gang. La oss prøve sammen*"*:

Eksempler:

1. Skilpadden leker med en *duss*	**Skilpadden leker med tre *dusser***
Etter barnets respons, si: *"Veldig bra. Denne dukken sa det riktig"* ***eller*** "*Hør godt etter".* Repetér den riktige setningen og si*: "Denne dukken sa det riktig fordi skilpadden leker med tre dusser*". *"Nå prøver vi noen fler".* Gjør det samme med de andre eksemplene. **Viktig:** Dersom barnet spør om testleder kan gjenta setningen, kan du *kun* gjøre det én gang. Sett ring rundt den setningen barnet velger for hvert bilde.	
2. Skilpadden ser en *føls* klovn.	Skilpadden ser flere *følse* klovner.
3. I går *vukket* skilpadden et hus	**Nå *vukker* skilpadden et hus**
4. Nå *gjummer* skilpadden tørsten med vann	I går *gjumte* skilpadden tørsten med vann

Det **korrekte** svaralternativet står med **fet skrift**. Skriv 1 for riktig svar og 0 for galt svar.

Item	*Forslag: Dukke høyre hånd*	*Dukke vestre hånd*	*Riktig 1*	*Galt 0*
1.	**Skilpadden hilser på *ådene***	Skilpadden hilser på *åden*		
2.	Skilpaddene går ombord i *jægen*	**Skilpaddene går ombord i *jægene***		
3.	Skilpadden bærer *favene*	**Skilpadden bærer *faven***		
4.	**Skilpadden holder en *væsp***	Skilpadden holder to *væsper*		
5.	Skilpadden fargelegger *åmtet*	**Skilpadden fargelegger *åmtene***		
6.	**Skilpadden spretter kebelen**	Skilpadden spretter keblene		
7.	Skilpadden sklir ned de *prosse* skliene	**Skilpadden sklir ned den *prosse* sklia**		
8.	Skilpadden spiser en *prin* sitron	**Skilpadden spiser *prine* sitroner**		
9.	**Skilpadden ligger på en *kræs* pute.**	Skilpadden ligger på *kræse* puter.		
10.	Skilpadden drikker av en *flei* kopp	**Skilpadden drikker av *fleie* kopper**		

11.	**Skilpadden kaster de *kvyre* skoene**	Skilpadden kaster den *kvyre* skoen		
12.	Nå *pyrer* skilpaddene i mål.	**I går *pyrte* skilpaddene i mål.**		
13.	**Nå *såmerer* skilpadden en pakke**	I går *såmerte* skilpadden en pakke		
14.	I går *terdret* skilpadden på butikken.	**Nå *terdrer* skilpadden på butikken.**		
15.	**Nå *gudder* skilpadden sakte.**	I går *guddet* skilpadden sakte.		
16.	I går *kræste* skilpaddene sammen.	**Nå *kræser* skilpaddene sammen.**		

Meta-inflectional Awareness Production Task- Final Version

Instruksjon til barnet: "Tidligere har vi sett at skilpadden gjør forskjellige ting. Nå skal vi se på bildene igjen. Jeg vil si hva som skjer på det ene bildet, og så kan du si hva som skjer på det andre bilde. La oss prøve."

Eksempler:

Testleder viser det første bildet og sier: "*Skilpadden leker med en duss*". **Testleder peker så på neste bilde og sier:"** *Skilpadden leker med tre...*" (Om barnet ikke svarer, kan det tilføyes: Hva ville skilpadden sagt?) **Barnet fortsetter setningen med nonordet:** *"dusser"*	
Etter barnets respons, si: *"Veldig bra"* **eller** "*Skilpadden leker med en duss. Skilpadden leker med tre dusser"*. Legg trykk på nonordet! **Viktig:** Dersom barnet spør om testleder kan gjenta setningen, kan du kun gjøre det én gang. **Testleder sier:** *" La oss se på noen fler!"*	
2. Skilpadden ser en *føls* klovn. Skilpadden ser flere ...	*følse klovner.*
3. I går *vukket* skilpadden et hus. Nå ...	*vukker skilpadden et hus*
4. Nå *gjummer* skilpadden tørsten med vann. I går ...	*gjumte skilpadden tørsten med vann*

OBS! Noter barnets **eksakte svar** dersom barnet svaret noe annet enn svaret oppgitt i tabellen. Feil uttale er tillatt så lenge barnet bøyer ordet korrekt. F. Eks. barnet sier *gukket* istedenfor *vukket*, eller *tølse* istedenfor *følse.*

Item	Bilde 1	Bilde 2	Annet svar:	Riktig 1	Galt 0
1.	Skilpadden hilser på *ådene*. Skilpadden hilser på ...	*åden*			
2.	Skilpaddene går ombord i *jægen.* Skilpaddene går ombord i...	*jægene*			
3.	Skilpadden bærer *favene.* Skilpadden bærer ...	*faven*			
4.	Skilpadden holder en *væsp.* Skilpadden holder to...	*væsper*			
5.	Skilpadden fargelegger *åmtet.* Skilpadden fargelegger ...	*åmtene*			
6.	Skilpadden spretter *kebelen.* Skilpadden spretter ...	*keblene*			

7.	Skilpadden sklir ned de *prosse* skliene. Skilpadden sklir ned den ...	*prosse* sklia			
8.	Skilpadden spiser en *prin* sitron. Skilpadden spiser ...	*prine* sitroner			
9.	Skilpadden ligger på en *kræs* pute. Skilpadden ligger på ...	*kræse* puter			
10.	Skilpadden drikker av en *flei* kopp. Skilpadden drikker av ...	*fleie* kopper			
11.	Skilpadden kaster de *kvyre* skoene. Skilpadden kaster den ...	*kvyre* skoen			
12.	Nå *pyrer* skilpaddene i mål. I går ...	*pyrte* skilpaddene i mål.			
13.	Nå *såmerer* skilpadden en pakke. I går ...	*såmerte* skilpadden en pakke			
14.	I går *terdret* skilpadden på butikken. Nå ...	*terdrer* skilpadden på butikken			
15.	Nå *gudder* skilpadden sakte. I går ...	*guddet* skilpadden sakte			
16.	I går *kræste* skilpaddene sammen. Nå ...	*kræser* skilpaddene sammen.			

Meta-derivational Awareness Production Task-Final Version

Instruksjon til barnet: "*Vi skal se på noen bilder sammen* (vis heftet med bilder til barnet). *Først skal jeg fortelle hva jeg ser. Etterpå skal jeg begynne den andre setningen og så vil jeg at du si resten av setningen med ett ord. La oss prøve først."*
OBS! Legg trykk på ordet barnet må bøye når eksemplene presenteres.

Eksempler:

1. Jenta **tegner**. Derfor kan vi si at jenta lager en.... (*tegning)*
Etter barnets respons til hvert eksempel, sier testleder: *"Veldig bra".* Dersom barnet svarer feil, sier testleder: "*La oss si det en gang til. Jenta tenger, derfor kan vi si at jenta lager en ... tegning!* Minn barnet på at det må svare kun med ett ord. **Testleder sier:** *Nå prøver vi noen fler*". **Viktig:** Dersom barnet spør om testleder kan gjenta setningen, kan du kun gjøre det en gang.
2. Lukas er glad i ***svømming.*** Lukas er glad i å .....(*svømme*)
3. Hun blir ***forvirret*** av spillet. Hun synes spillet er .... *(forvirrende)*
4. Jenta sitter og ***smiler.*** Jenta sitter .... *(smilende)*

OBS! Dersom barnet svarer **noe annet** enn svaret i tabellen, noter barnets eksakte svar.

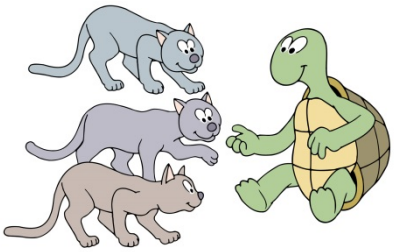
Item			**Annet svar:**	**Riktig 1**	**Galt 0**
1.	Katten har mye ***hår***.	Katten er veldig …(*hårete)*			
2.	Erik ***trener*** mye.	Erik har vært mye på .....(*trening)*			
3.	Såpen ***skummer***.	Såpen er veldig .....*(skummete)*			
4.	Mari ***skriver*** i timen.	Mari har en time med .....(*skriving)*			
5.	Ole har en genser med ***striper***.	Ole har en genser som er .....(*stripete)*			
6.	Nicolai ***beveger*** **seg** til musikken	Nicolai gjør en.....(*bevegelse)*			
7.	Anne ***føler seg*** trist.	Anne har en trist .....(*følelse)*			
8.	Moren ***trøster*** jenta	Moren er ..... (*trøstende)*			
9.	Gutten er glad i ***lesing***	Gutten liker å .....*(lese)*			

10.	Han ***maler*** veggen	Han dekker veggen med .....*(maling)*			
11.	Han blir ***sliten*** av dansing	Han synes at dansing er..... *(slitsomt)*			
12.	Parfymen ***dufter*** godt	Parfymen er .....*(duftende)*			
13.	Fuglen ***flyr***	Fuglen kommer .....*(flyvende)*			
14.	Jenta ***fleiper*** med de andre	Jenta er veldig .....*(fleipete)*			

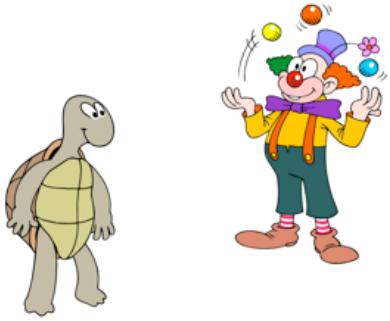
Appendix 2 Pictures of the Morphological Awareness Test - Final Version

Epi-inflectional Awareness Judgement Task Pictures

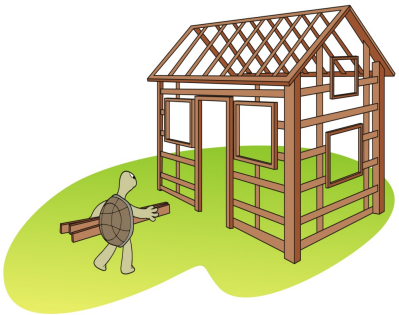
Eksempel 1



Eksempel 2

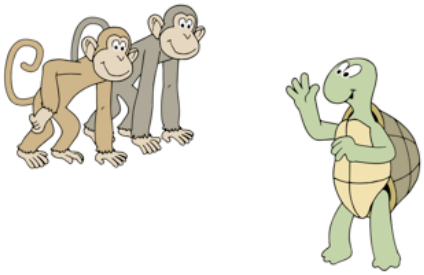


Eksempel 3

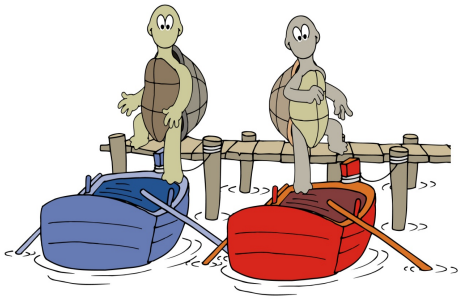


Eksempel 4

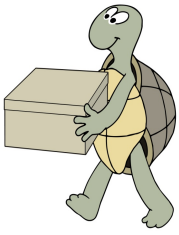
Item 1



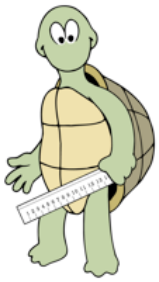
Item 2



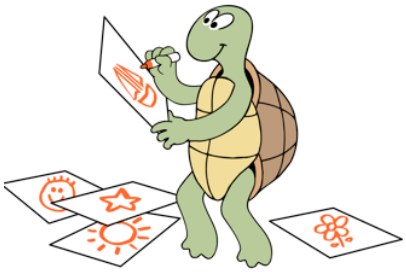
Item 3



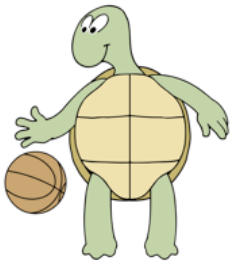
Item 4



Item 5

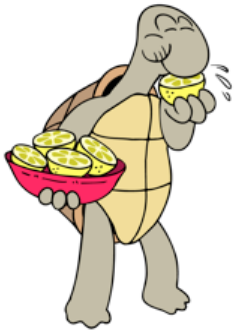


Item 6

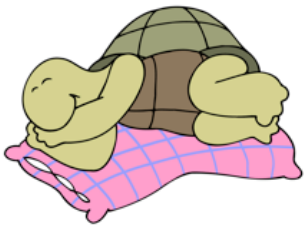


Item 7

Item 8

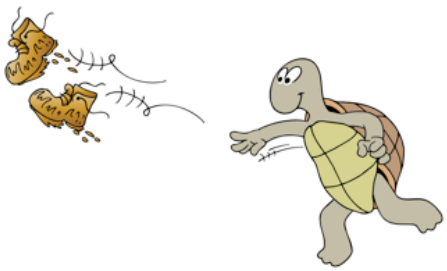


Item 9

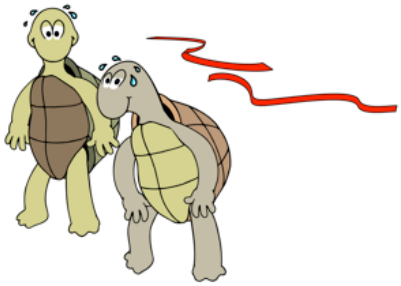


Item 10

Item 11



Item 12 (deleted item)

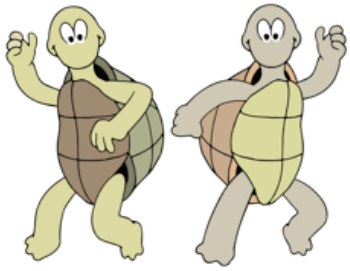


Item 13

Item 14

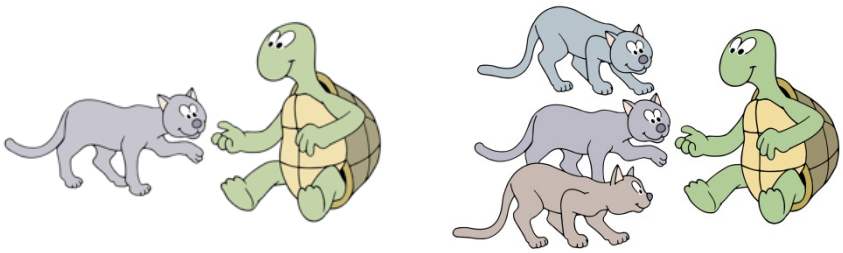
Item 15

Item 16

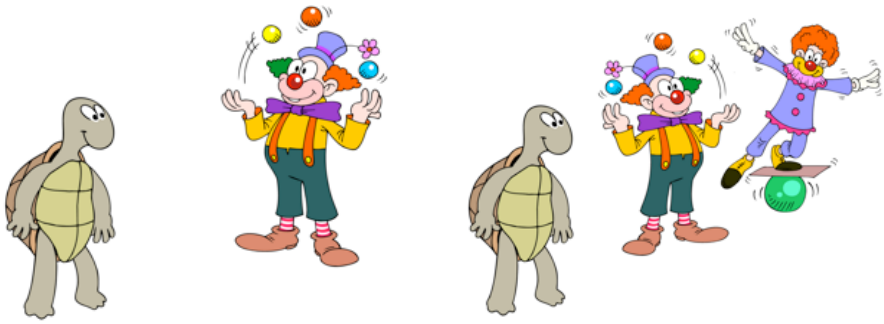


Meta-inflectional Awareness Production Task Pictures

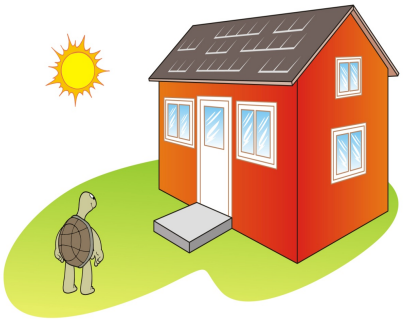
Eksempel 1

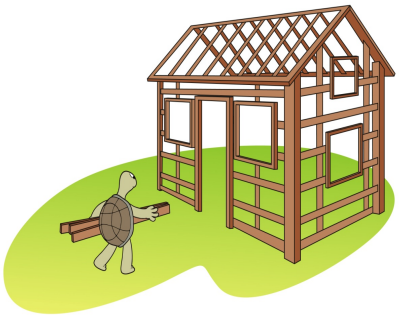


Eksempel 2



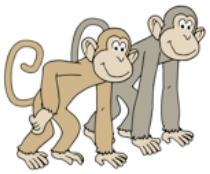
Eksempel 3



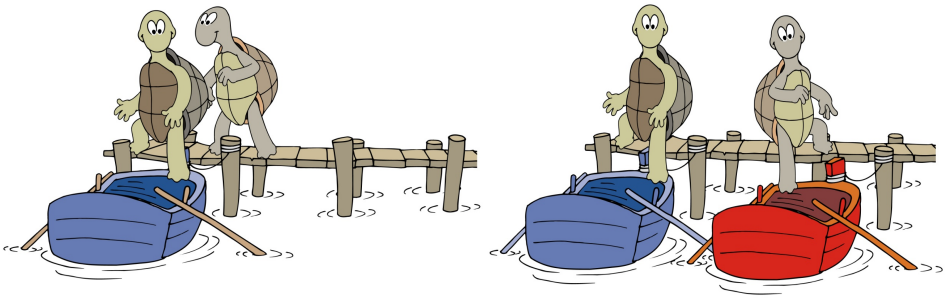


Eksempel 4

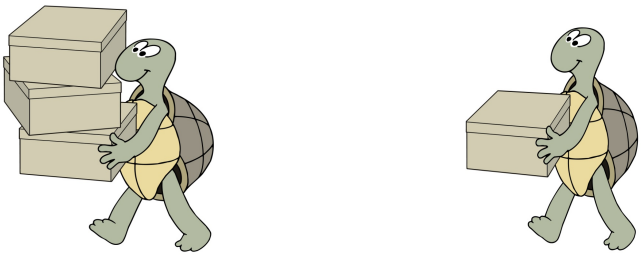
Item 1



Item 2

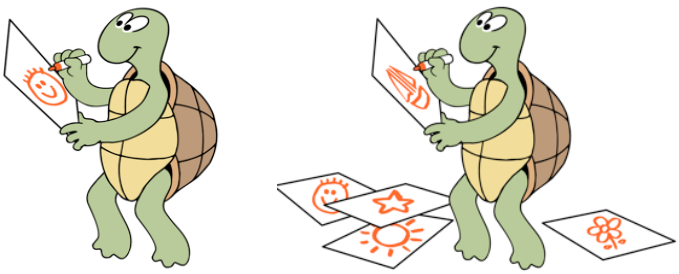


Item 3

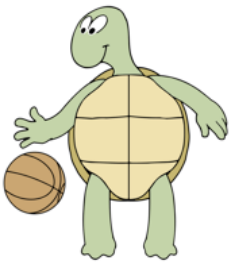


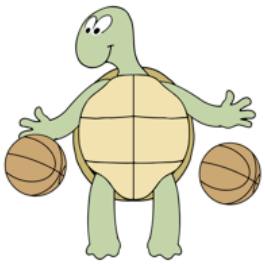
Item 4

Item 5

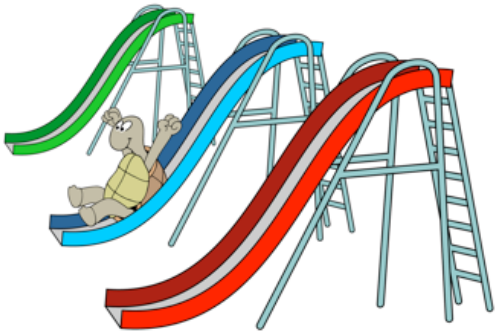


Item 6

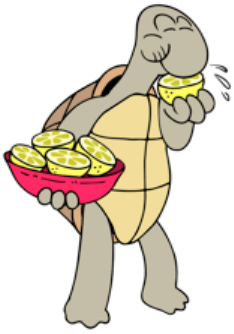




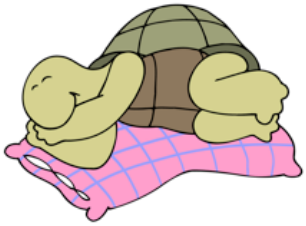
Item 7



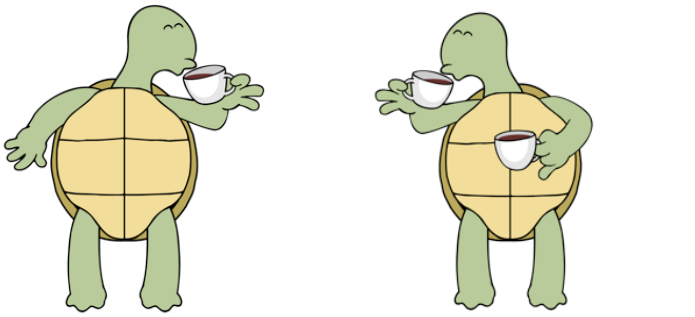
Item 8



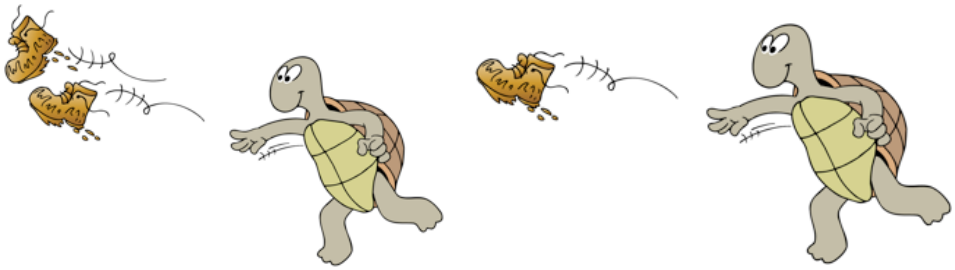
Item 9



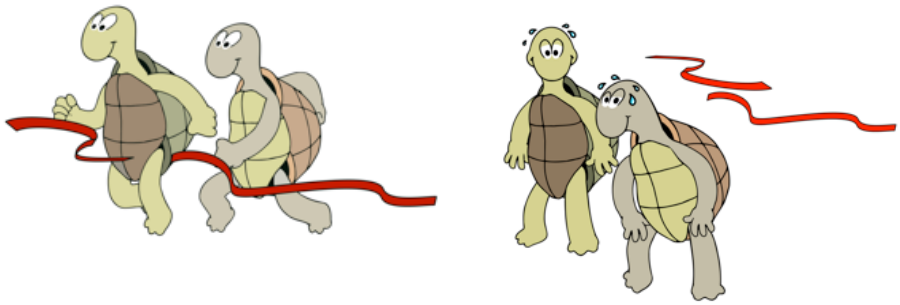
Item 10



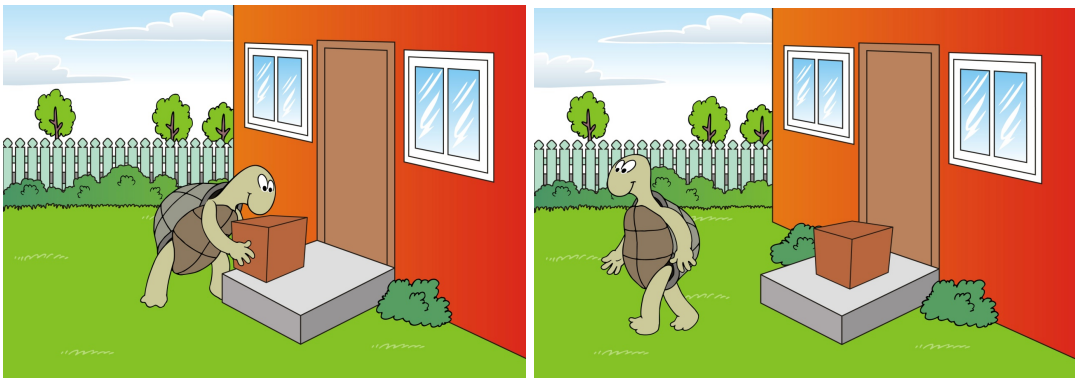
Item 11



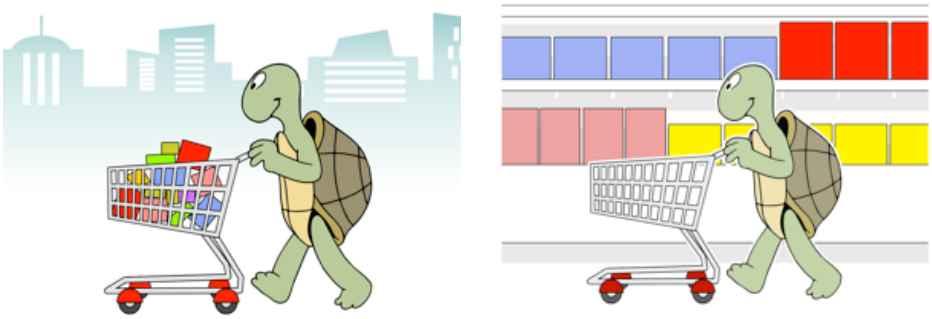
Item 12



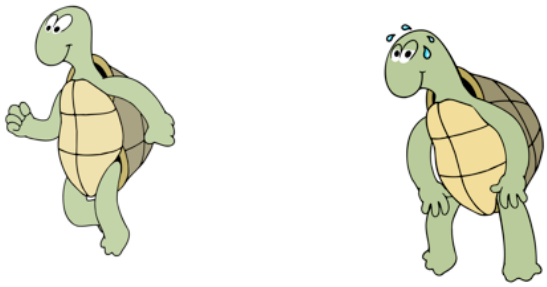
Item 13



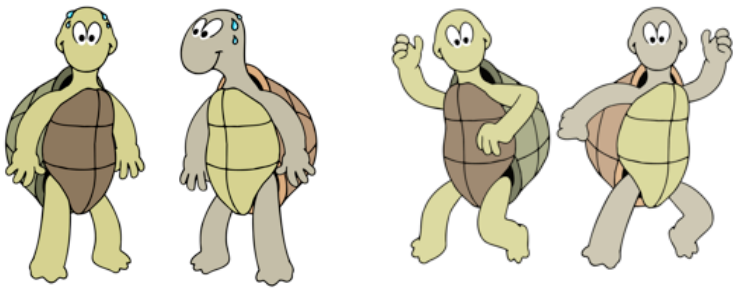
Item 14



Item 15

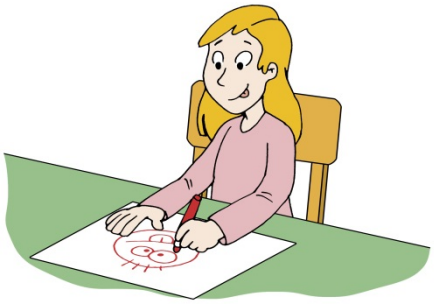


Item 16

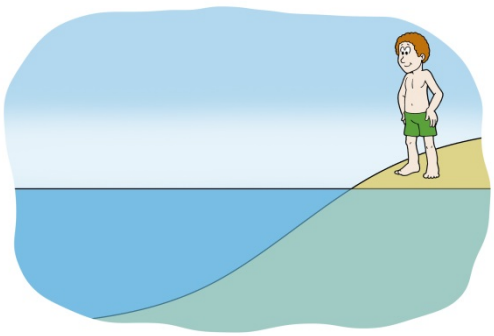


Meta-derivational Awareness Production Task Pictures

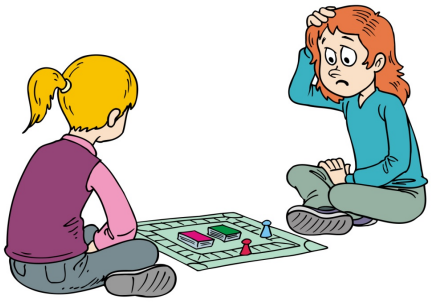
Eksempel 1



Eksempel 2

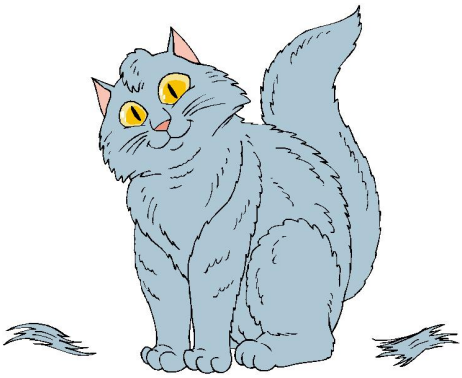


Eksempel 3



Eksempel 4

Item 1



Item 2

Item 3

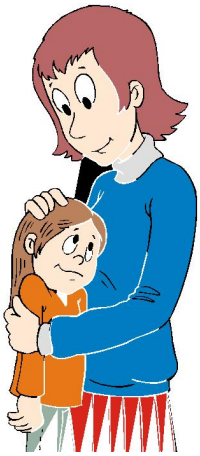
Item 4

Item 5

Item 6

Item 7

Item 8



Item 9

Item 10

Item 11

Item 12

Item 13

Item 14

Appendix 3 Guidelines for the Pronunciation of Non-words

Epi-inflectional & Meta-inflectional Awareness Task

Uttale av non-ord

Følgende liste beskriver kort hvordan non-ordene må uttales. I kolonnen «uttale», finnes det et eksempel på et ekte ord som ligner på non-ordet. Det er viktig at du respekterer normene for uttale av non-ord ettersom målet er at alle testledere sier ordene på samme måte.

Eksempel	**Non-ord**	**Uttale**
1.	Duss	Uttales med «u». Trykk som i ordet «tuss».
2.	Føls	Trykk som i ordet «løs».
3.	Vukker	Uttales med «u». Trykk som i ordet «vugger».
4.	Gjummer	Uttales med «o». Trykk som i ordet «gjemmer».
Item	**Non-ord**	**Uttale**
1.	Ådene/åden	Trykk på «å» som i ordet «åsene».
2.	Jægen/jægene	Uttales med «æ» og ikke med «e».
3.	Favene/faven	Uttales som i ordet «pave» eks. pavene/paven.
4.	Væsp/væsp	Uttales med «æ» og ikke med «e». Unngå å si «vesp».
5.	Åmtet/åmtene	Uttales «åmte». Uten siste «t-lyden» som i et intetkjønnsord. Dersom «t» uttales kan non-ordet forveksles med et adverb.
6.	Kebelen/keblene	Uttales lik som i ordet «kabel». Eks. kabelen/kablene
7.	Prosse	Uttales med «å», Altså «pråsse».
8.	Prin/prine	Uttales lik som i ordet «prim» eller «rim».
9.	Kræs/kræse	Uttales med «æ» og med lang vokal. Altså «krææs».
10.	Flei/fleie	Som i ordet «lei» og «leie».
11.	Kvyre	«k» uttales som i ordet «kva»
12.	Pyrer	Samme trykk som i ordet «fyre»
13.	Såmere	Samme trykk som i ordet «levere» eller «summere»
14.	Terdret/terdrer	«er» uttales som i ordet «herder». Non-ordet erstatter ordet «handler» derfor legges til en «r» mellom «d» og «e». Altså «Terd**r**er»
15.	Gudder	Uttales med «u». Samme trykk som i ordet «jogger».
16.	Kræste/kræser	Uttales med «æ» og ikke med «e».

Appendix 4 Guidelines of Scoring and Administration

Kriterier for skåring av Meta-inflectional Awareness Production Task

Den følgende listen gir en oversikt over de ulike svarene som anses som korrekt og derfor får 1 poeng. Avvik fra disse svaralternativene vil gi 0 poeng.
Barnet får også 1 poeng dersom barnet erstatter ordlyder i stammen av ordet, men bruker riktig bøyning (-en, -et, -ene, er osv.) For eksempel om barnet sier «åben» eller «åmen» istedenfor «åden».
Kolonnen «målord» inneholder non-ordet som fungerer som stimulus og i kolonnen «tillatte svar» finnes ulike svaralternativer som får 1 poeng. Item 7-11 er markert med en stjerne(*) ettersom disse svarene krever at barnet sier både non-ordet og substantivet i riktig bøyningsform for å få 1 poeng. Det vil si at dersom barnet kun sier «prosse» og ikke «prosse sklia/sklien» får barnet 0 poeng.
I kolonnen «opprinnelig ord» står det ekte ordet som non-ordet er basert på kun som en referanse for testlederen.

Item	Målord	Opprinnelig ord	Tillatte svar
1.	Ådene	Apene/apen	Åden
2.	Jægen	Båten/båtene	Jægene
3.	Favene	Boksene/boksen	Faven
4.	Væsp	Linjal/linjaler	Væsper
5.	Åmtet	Arket/arkene	Åmtene
6.	Kebelen	Ballen/ballene	Keblene
7.	Prosse	Bratte	Prosse sklia/prosse sklien*
8.	Prin	Sur/sure	Prine sitroner/fem prine sitroner/ flere prine sitroner*
9.	Kræs	Myk/myke	Kræse puter/flere kræse puter/ seks kræse puter*
10.	Flei	Hvit/hvite	Fleie kopper/ flere fleie kopper/ to fleie kopper*
11.	Kvyre	skitne	Kvyre skoen*
12.	Pyrer	Løper	Pyrte/Pyret/pyra
13.	Såmerer	Leverer	Såmerte/såmeret
14.	Terdret	Handler	Terdrer
15.	Gudder	Jogger	Guddet/gudda
16.	Kræste	Trener	Kræser

***item 7- 11** barnet må både si og bøye substantivet i tillegg til non-ordet som oppgitt i tabelen for å få poeng.