The interface between language abilities and cognitive abilities, how much does it contribute to the assessment of bilingual children. A study on Norwegian Bilingual 8-year-olds.

Lilian Haugereid and Anne Dalland Coldwell

Candidate number 101

Master's Thesis

Master's programme in health sciences – Logopedics

The Faculty of Psychology

Department of Biological and Medical Psychology

University of Bergen

Spring 2018

Acknowledgments

We would first like to express our sincere appreciation to our main supervisor, Turid Helland and our second supervisor, Wenche A. Helland, both of the Department of Biological and Medical Psychology, University of Bergen, for their guidance, enthusiastic encouragement, and useful critiques of this research work. We would especially like to thank Turid Helland for her advice and assistance in keeping the progress on schedule and her valuable technical support on this project. Our sincere thanks go to Wenche A. Helland for her active involvement during the data collection for this study.

Secondly, we would like to thank all the school children involved in this study. We appreciate their patience and their cooperation. We would also like to thank the schools and teachers that helped us with the practical implementation of the testing.

We express our very profound gratitude to our family, particularly our spouses, Petter Haugereid and Kris Coldwell, for providing us with unfailing support and continuous encouragement throughout our years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without their support. Thank you! We would like to thank our children, Naomi, Sara, John, Even and Maia, for the patience they had for us during the period of our Master's studies and during the writing of this thesis.

Finally, we would like to thank Hanna Nordbø and Marianne Dahl for their contribution to this study during the collection and scoring of data, and for their cooperation.

Thank you all!

Lilian and Anne, May 23rd, 2018.

TABLE OF CONTENT

ARTICLE

APPENDICES

Sammendrag

Hovedformålet med denne studien var å kunne skille mellom språkvansker som skyldes en forbigående mellomspråklig vanske, og vansker som skyldes SLI og /eller dysleksi hos tospråklige 8 -åringer med Norsk som andrespråk. Utvalget i studien var 20 tospråklige 3.-klassinger fra en større norsk by. Vi hadde en kontrollgruppe på 42 enspråklige barn, fra en tidligere studie. Testbatteriet var basert på både språklige (L2) ferdigheter og nevrokognitive faktorer, i henhold til de tre nivåene (symptomatisk, kognitivt og biologisk) hentet fra den differensialdiagnostiske modellen til Morton og Frith. Vår hypotese var at tospråklige barn med språkvansker kan identifiseres av Morton og Frith sin modell for differensialdiagnostisering. Vi ønsket også å se om tospråklige barn har noen nevrokognitive fordeler i forhold til den enspråklige gruppen. Vår hypotese var at enspråklige vil ha bedre resultat på de språklige testene i forhold til L2-gruppen, men det var ikke forventet noen forskjeller på gruppene med hensyn til nevrokognitive evner.

Som ventet, viste funnene dårligere resultat på språktestene for L2 -gruppen i forhold til den enspråklige gruppen. De kognitive testene viste at de tospråklige hadde bedre resultat på de testene som omhandlet visuelle og auditive evner sammenlignet med L1-gruppen. Resultatene fra studien viser at det er mulig å bruke den differensialdiagnostiske modellen til å identifisere tospråklige barn som er i risikosone for å ha språkvansker og/eller dysleksi og å skille mellom dem og de som har en mellomspråklig vanske.

Nøkkelord: Tospråklighet, Mellomspråklig vanske, Språkvansker, Dysleksi, Differensialdiagnose, Nevrokognitive fordeler.

Abstract

The main aim of this study was to separate transient between language problems from problems caused by SLI and/or dyslexia among bilingual 8-year olds having Norwegian as their second language (L2). The sample in the study was 20 bilingual-3rd grade school children from a larger Norwegian city. We had a control group of 42 monolingual children, from a previous study. Assessment tools were based on both linguistic (L2) skills and neurocognitive factors, according to the three levels (symptomatic, cognitive and biological) of the differential diagnosis model by Morton and Frith. We hypothesized that bilingual children with language impairment can be identified by Morton and Frith's differential diagnostic model. We thus aimed to explore whether the differential diagnosing model, can serve as a method in defining language impairment in bilingual children. We also wished to see if bilingual participants exhibit some cognitive advantages over their monolingual counterparts. We hypothesized that monolinguals will outperform bilinguals on language abilities, but no group differences were expected with respect to neurocognitive abilities.

As predicted our findings showed poor bilingual performances in language tests compared to monolingual performances. The neurocognitive tests showed bilingual strengths in visual and auditory processing compared to the L1 group. Findings from our study clinically imply that the differential diagnostic model can be used to identify bilingual children at risk of having language impairment and/ or dyslexia, from children with between language problems.

Keywords: Bilingualism, Between language problems, Language impairment, Dyslexia, Differential diagnosis, Neurocognitive advantage.

Background of the study

Globalization is an increasing process worldwide and so is bi-/multilingualism. Norway is one of the countries that attract large numbers of immigrants. Statistics Norway (SSB) defines immigrants as persons born abroad of two foreign-born parents, who have four foreign-born grandparents; and children of immigrants are referred to as Norwegian born with immigrant parents (Sandnes, 2017). SSB further reports that at the end of the year 2017, 884,000 persons had immigrant background (16.8% of the entire population), of which 159,000 are Norwegian born with immigrant parents (3% of the entire population) (Sandnes, 2017). With regards to minority language statistics, there is no recent official coordinated language statistics in Norway. Statistics from 2013 show that 311 different languages were registered in Norway (Wilhelmsen, Holth, Kleven & Risberg, 2013). In 2016, out of the 282,600 kindergarten children, 46,300 of them had minority language background (SSB, 2017b). According to SSB, statistics from elementary schools are not individual-based, therefore it is not possible to say exactly how many immigrants or Norwegian-born children with immigrant parents are registered in primary schools. However, according to the Education Act (opplæringsloven) all children and young people have the right and obligation to primary education; so, we can therefore assume that most children aged 6-15 years attend school (SSB, 2017b). Section 2.8 of the Education Act (Opplæringsloven § 2.8) says that language minority pupils in primary schools with another mother tongue than Norwegian and Sami are entitled to special tuition, and if necessary, they also have the right to mother tongue education, bilingual teaching or both. SSB found out that the number of pupils that receive special Norwegian language teaching is increasing. In 2016, 45,300 pupils received special tuition in Norwegian language (SSB, 2017b). Results from 2016 national tests show that both immigrants and Norwegian-born with immigrant parents score lower than other pupils in most

of the tests. Minorities in Norway often have some other spoken language at home. If the problems lie in the language of instructions and the fact that minority school pupil might not have the same language competence in Norwegian as their monolingual fellow pupil, then it is not surprising that school tests scores are lower for these groups of pupils. Bialystok, Luk, Peets, and Yang (2010) write that bilingual children need to "distribute their language-learning time across two languages, and it is likely that some words occur in a context in which they only use one of their languages." Bialystok, Luk et al. (2010) further point to many studies (Adams, 1990; Kastner, May & Hildman, 2001; Ouellette, 2006; Ricketts, Nation & Bishop, 2007; Rohde & Thompson, 2007; Swanson, Rosston, Gerber & Solari, 2008) that show evidence that vocabulary size is important for "academic achievement and literacy acquisition."

A survey, done by the Norwegian Education Directorate (Utdanningsdirektoratet) in 2016 indicates that teachers who teach pupils with Norwegian as a second language (L2), do not have sufficient education to teach the target group (Rambøll Management, 2016). This indicates that teachers are not always able to evaluate bi-/multilingual children if language impairment is involved.

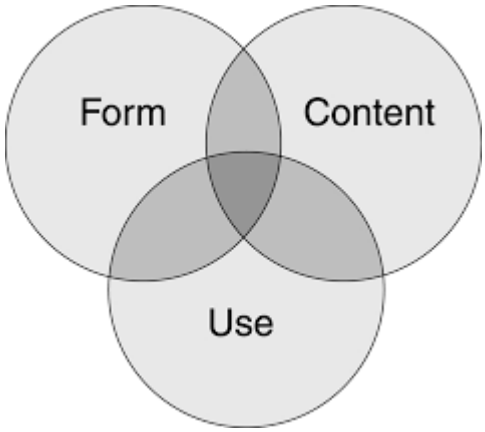
Theoretical framework

What is language

Language can be defined differently depending on the type of study it concerns. Bloom and Lahey (1978) mention that the definition of language depends on the context by which one asks the question "what language is" (Bloom & Lahey, 1978, p. 4). Language can, therefore, be seen as a mode of communication in a sociological definition. Language can also be seen as a system of communication used by a set of people or community, or country; in which case, we can refer to different types of languages such as Norwegian, English, French or Urdu. Anderson and Shames (2014, p. 6) define language linguistically as "a conventional

system for representing concepts through the use of arbitrary symbols and rule-governed combinations of those symbols". The language model proposed by Bloom and Lahey (1978) has three intersecting components of language; the content, form, and use.

Figure 1. Illustration of Bloom and Lahey's language model.



*"Language consists of some aspects of **content** or meaning that is coded or represented by linguistic **forms** for some purpose or **use** in a particular context. This three-dimensional view of language is basic to describing the development of language and for understanding language disorders"* (Bloom & Lahey, 1978, p. 11)

Bilingualism

***Bilingual development*:** Helland (2012) writes that although there are major individual differences in language development, we can describe typical development patterns common to most children. Children (monolinguals) can produce key phrases by the age of 2 to 3, and by 3 they produce sentences and experiment on inflecting nouns and verbs (Helland, 2012, p. 31). According to Kohnert (2013), the speech of typically developing (TD) monolingual children have grammatical complexity and are intelligible to listeners familiar to the child, but some children may have "normal nonfluencies" (Kohnert, 2013, p. 82). For TD bilinguals, development and skills in languages may depend on factors such as the age of acquisition of the languages, the degree of stimulation and language experience, language social status, the motivation for learning languages et cetera (Egeberg, 2016, p. 19-20). Paradis, Genesee, and

Crago (2011) mention that bilingual TD children tend to have a smaller vocabulary than monolinguals in each of their languages but a combination of their vocabularies from both languages is bigger than those of their monolingual counterparts. Where there is the absence of other disabilities, like hearing impairment, syndrome et cetera, bilingual children develop their languages in principle in the same way as monolingual children (Egeberg, 2016, Kohnert, 2013). According to Hoff (2014), the acquisition of language by minority children seems to follow the same time (pattern) as language development by monolingual children. Children with minority language (L1 and an L2) often spend more time learning each of the languages than monolingual children whose language is the language of instruction, they use their first language to learn their own language. In other words, it appears that children exposed to, and learning, two languages have a slower development of each of their languages than monolingual children (Hoff, 2014).

What is bilingualism: A bi-/multilingual person is widely known as a person with knowledge in two or more languages. But how comprehensive a person needs to be in the languages in order to be considered bi-/multilingual is debatable. Øzerk (2016, p. 73-74) discusses researchers' moderate and strict perceptions of bilingualism. Among those with moderate perception on bilingualism is Hall (1952) who defines a bilingual as a person with "some knowledge and control of the grammatical structure of the second language" (as cited in Øzerk 2016, p. 72). Researchers with strict perspective on bilingualism have strict requirements to the linguistic skills in both languages in order to characterize someone as bilingual (Øzerk, 2016, p. 73). According to this view, a bilingual person is a person who has monolingual competence in both languages, and this competence is confirmed by measuring the person's language abilities through language tests (Øzerk, 2016, p. 73-78). Egeberg (2016) has a moderate definition of multilingualism; the term often includes all those who relate to and have command of "*behersker*" several languages (Egeberg, 2016, p. 11). Helland

(2012, p. 180) points out that most people are bi-/multilingual and being monolingual is the exception. Citing Archibald (2011), Helland (2012, p. 180) points out that bilingualism is a development opportunity for everyone, both children, and adults.

***Types of Bilingualism*:** Egeberg (2016) refer to bilingual children who acquire one language (L1) from birth and the other (L2) at a later age, as sequential bilinguals, while Kohnert (2013) refer to them as successive bilinguals. Paradis et al. (2011) prefer the term "dual language" in referring to children who acquire two distinct languages simultaneously. Øzerk (2016, p. 54) presents the principal distinctions between categories of bilingualism in Norway. He groups them into minority bilinguals and majority bilinguals. Minority bilinguals include bilinguals with immigrant, indigenous (urfolk), Nynorsk and Kven backgrounds (Øzerk, 2017, p 54).

Furthermore, Paradis et al. (2011) and Kohnert (2013) distinguish between subtractive and additive bilingualism. According to Paradis et al. (2011), subtractive bilingualism (a term they borrowed from Lambert, 1977) occurs when children acquire a majority language at the expense of their native language. This type of bilingualism is common among immigrant children or grandchildren of immigrants where they must learn the dominant language (Paradis et al., 2011, p. 49 -50). Kohnert (2013, p. 93) explains that this type of bilingualism is subtractive because the language of the community is promoted at the disadvantage of the child's first language (L1). On the other hand, additive bilingualism occurs when both languages are promoted by the community, and the promotion of the second language is to "supplement and not supplant" the second language (L2) (Kohnert, 2013, p. 93). Øzerk (2016, p. 74-79) outlines four different types of bilingualism based on a linguistic and competence-oriented approach: well-developed (balanced) bilingualism, mother tongue dominant bilingualism, L2 dominant bilingualism and unsatisfactory bilingualism. A well-developed bilingual is proficient in two languages and her linguistic skills in both languages are at a high

level. Meanwhile, a mother-tongue dominated bilingual has high linguistic skills in L1 but rather low linguistic skills in L2. L2 dominated bilinguals, on the other hand, have high linguistic abilities in L2. The last of the four, unsatisfactory bilingualism, refers to a bilingual development situation where the child does not score high in any of its languages, or when the child does not have a native level either in L1 or L2. In all four cases, linguistic competence in the two languages is measured by monolingual norms in each language (Øzerk, 2016, p. 74-79).

The brain, cognition, and bilingualism

Kohnert (2013, p. 217 - 218) writes that modern research on well-functioning brains in monolingual adults generally supports the lateralization and localization view of language functions. Thus, for most monolinguals, language is lateralized or located in the left hemisphere; especially language form, lexical system and rapid processing of linguistic units (Kohnert, 2013, p. 218). Abutalebi, Cappa, & Perani (2005) and Paradis (2004) point out that the left cerebral hemisphere is also the dominant hemisphere for bilinguals (cited in Kohnert, 2013, p. 220). Results from a meta-analysis on lateralization of language between monolinguals and bilinguals by Hull and Vaid (2005), show that monolinguals and late bilinguals (L2 acquisition after age 6) use the left hemisphere and that early bilinguals use both hemispheres more equally (stated in Kohnert, 2013, p. 220). These findings by Hull and Vaid (2005) also show gender differences. Neural activation patterns during language processing for both female monolinguals and bilinguals were in the left hemisphere. Male bilinguals showed lateralization for both hemispheres whereas their monolingual counterparts showed left lateralization (Kohnert, 2013. p. 220).

Paradis et al. (2011) refer to Laura Berk's explanation of what cognition is. It states that cognition *"includes all mental activity - attending, remembering, symbolizing, categorizing, planning, reasoning, problem solving, creating and fantasizing"* (Paradis et al.

2011, p. 39). Bialystok (2001) discusses the interaction between language and cognition by first pointing out two polarized views. One view is that children learn language in order to express acquired concepts. Another view is that learning a lexicon "signals the occasion to create a concept, thus "the linguistic features of words" will guide their "conceptual conceptions" (Bialystok, 2001, p. 189). Irrespective of whether the interaction is from concept to words or words to concept, Bialystok (2001) adds that it is undoubtable that there is an interaction between the way monolingual children learn language and the way they learn concepts. The evidence for this interaction can be found in the way children build up their linguistic and conceptual structures when they learn different languages (Bialystok, 2001, p. 189). Bialystok (2001) exemplifies this by discussing research done on the acquisition of spatial prepositions such as *in* and *on* in English and Korean. In Korean, the difference between *in* and *on* is whether the relationship between the referent and the object is tight fitting or loose fitting. In English, however, the relationship between object and the referent is containment or surface attachment (Bialystok, 2001, p. 189).

What are language impairment and dyslexia

According to Helland (2012, p. 64), a child has language impairment when the child has difficulties speaking, when others have difficulties understanding the child when she speaks, or when the child has difficulties understanding language. It is also the case when the child does not follow an anticipated progress in her language development. Helland (2012) also points out that language impairments occur most frequently in preschool age and that about 10-15% of all children have delayed language development (2012, p. 64). As discussed above, according to the language model by Bloom and Lahey (1978), a good and normal language development involves the successful integration of the components in the language model; that is, language content (meaning), form (structure and grammar) and use (interaction and communication with others). However, if there is a "glitch" in one or more of these

components, or in the interaction between them it will result in some form of language impairment (Bloom & Lahey, 1978). Helland writes that specific language impairment (SLI or primary language impairment - PLI) is a term used for children whose language development is considerably below age-norm, without this being related to any clear and unambiguous cause (2012, p. 64). Helland further points out that when language disorders cannot be attributed to organic deviations, neurological injuries, brain injuries or any kind of diagnosis, it is usually called SLI.

There is often a link between SLI and reading impairment (Paradis et. al., 2011; Helland, 2012). Thus, children with SLI often exhibit difficulties in learning how to read.

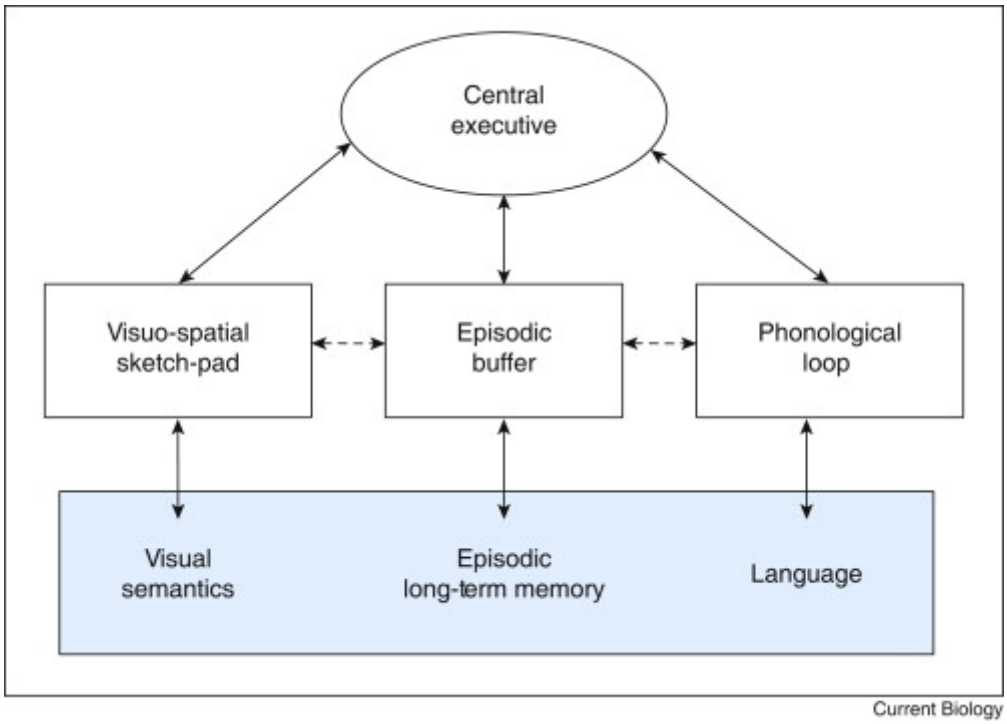
Executive function, visual-spatial skills, and working memory/short-term memory

Following Pennington and Ozonoff (1996) and Henry (2012), Armon-Lotem, de Jong and Meir (2015, p. 278) look at the five main components of executive functions; switching; fluency; planning; interference inhibition; working memory/updating. According to Best and Miller, (2010); Henry, (2012) and Miyake et. al., (2000), *switching* is "the ability to navigate fluently between different sets of instructions, mental states or tasks, which also involves moving attention from one thing to another" (cited in Armon-Lotem, de Jong and Meir, 2015, p. 279). *Fluency,* on the other hand, requires the ability to generate elements based on certain rules, and fluency tasks are often "associated with verbal tasks such as rapid naming and semantic fluency" (Armon-Lotem, de Jong, & Meir, 2015, p. 280). *Planning* is how we organize our thoughts and actions to meet a goal (Armon-Lotem, de Jong & Meir, 2015). And *interference inhibition* is the ability to delay a certain behavior which is often a previously learned response (Armon-Lotem, de Jong & Meir, 2015, p. 280). Empirical studies proving bilingual advantages in interference inhibition is discussed later in this writing.

Working memory refers to "the system or systems that are assumed to be necessary in order to keep things in mind while performing complex tasks such as reasoning,

comprehension and learning" (Baddeley, 2010). According to Baddeley, (2010) working memory concept "evolved" from the concept of short-term memory, which he defined as "temporary storage of small amounts of material over brief periods of time". Baddeley (2010) further reports that children with specific language impairment (SLI) usually have poor short-term memory and are slow in acquiring new words/vocabulary.

Figure 2. Illustrates Baddeley and Hitch's memory theory (borrowed from Current biology journal)



The proposal of working memory is from earlier work by Baddeley and Hitch (1974) who divided it into three subsystems; the phonological loop (PL), the visuospatial sketchpad (VS) and the central executive. The theory proposes that short-term memory is composed of these three components, which work together as a part of a "unified working memory system that serves the function of facilitating the performance of a range of complex tasks" (Baddeley, 2003, p. 190). The PL processes verbal and acoustic information while VS provides visual information; but both the PL and the VS are dependent on the central

executive which is an attentionally-limited control system (Baddeley, 2003). A fourth component, the episodic buffer, combines visual and auditory information because of its ability to hold multidimensional episodes (Baddeley, 2010).

According to Baddeley (2003), the three components of short-term memory (that is, PL, VS and central executive) comprise "verbal-acoustic" storage system which is important for tasks like immediate retention of digits (Baddeley, 2003, p. 190). Evidence showing that impairment in the phonological loop can lead to difficulties in acquiring a second language (L2) is provided by Service (1992) on Finnish (L1) and English (L2) learners (mentioned in Baddeley, 2003). Service (1992) found out that children with good immediate verbal memory were better at learning L2 than those with short spans (Baddeley, 2003).

According to Helland (2012, p. 139), phonemes and graphemes are stored in the short-term memory. According to the working-memory model by Baddeley and Hitch, the processing of the phonemes occurs in the phonological loop. The passive phonological storage site is located in the Wernicke's area, while the active retrieving system is in both the Brocas area and the motorway. Furthermore, this information should be controlled by the central executive, which itself does not have storage capacity, but determines how the information is to be processed (stated in Helland, 2012, p. 139). Helland (2012) further points out that if one or more of these elements, that is, passive phonological storage and the active retrieval system, have low capacity, it will result in a less effective short-term memory and working memory. Regarding the cognitive aspect of language, Helland (2012) writes that testing for linguistic cognition involves testing phonological, morphological and syntactic skills. Thus, low scores on phonological awareness, vocabulary and linguistic comprehension can relate to failure within the phonological loop (Helland, 2012, p. 142).

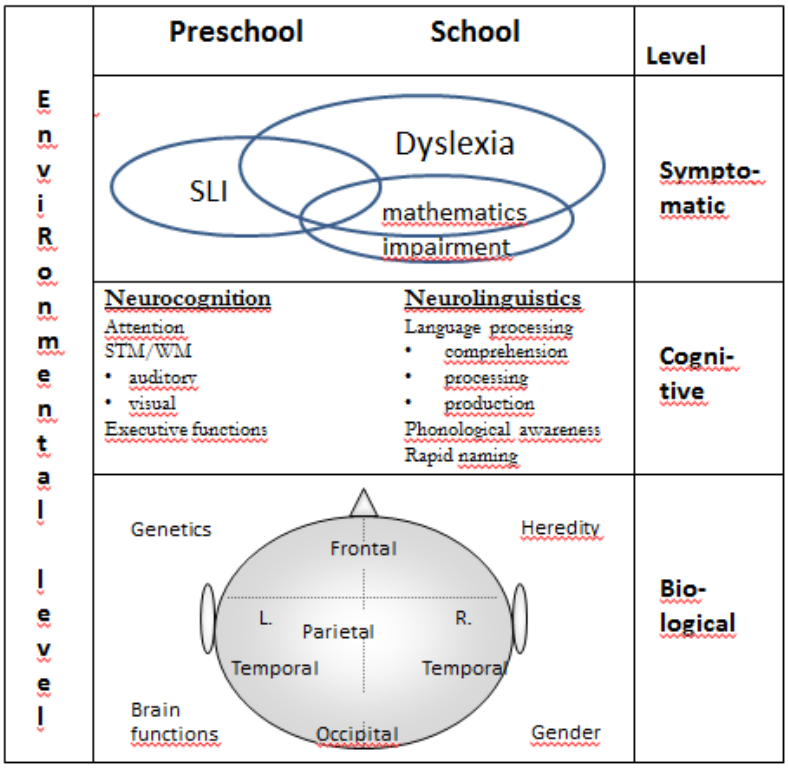
Detecting Language impairment and/or dyslexia in bilinguals

There are many different standardized assessment tools for testing the different aspects of language (as outlined by Bloom and Lahey, 1978) on monolinguals when language development does not go as expected. But there is lack of such tools for bilinguals. Challenges speech and language pathologists (SLPs) face, when working with bilingual children, include lack of standardized tests in their L1, less training and little experience in working with bilingual children, and "lack of mentors to guide the complex process of language assessment in developing bilinguals" (Kohnert, 2013, p.146). Kohnert warns that this may over-identify, under-identify, or misidentify developing bilinguals. Paradis et al., (2011; p. 213) advise that the use of monolingual standardized tests for assessing bilinguals may lead to biased assessment. But they also point out that some of these test batteries have subtests that test "language-specific" abilities in contrast to other subtests that test "language-general" abilities; and the subtest that tests the latter "do not entirely depend on accumulated language-specific knowledge" (Paradis et al. 2011, p.215). De Jong (2008) refers to assessment reviews by Gutiérrez-Clellen (1996) who communicates her skepticism towards the use of monolingual test batteries on bilinguals. Such tests lack specificity (the ability to identify those without disorders) and sensitivity (the ability to identify those with a disability) when tested on bilingual children (de Jong, 2008).

Armon-Lotem, de Jong and Meir (2015, p. 95) citing Conti-Ramsden et al., (2001) state that sentence repetition tasks (SRep task) are found to be very sensitive and specific in identifying language impairment among monolinguals. SRep involves listening to sentences and repeating them verbatim (Armon-Lotem, de Jong & Meir, 2015, p. 98). But so far, there are no assessment tools for bilingual or multilingual children, there is also a lack of SRep tasks for this same group of children (Armon-Lotem, de Jong & Meir, 2015, p. 96).

Helland (2012) suggests the adaptation of the causal model by Morton and Frith (1995) when collecting information on a disorder. Helland (2012, p. 124) states that this model, which is the core of the investigation work, can help to differentially diagnose persons with language impairment and/or dyslexia. The causal model suggests the collection and investigation of information concerning biology, cognition, and symptoms on persons under assessment for language impairment. The model also states that environmental issues also should be considered when assessing the three different levels. This involves collecting information at a symptom level (language difficulty symptoms the child has), biological level (information on health, heredity, genetics, brain function, hearing, gender and motor skills), cognitive level (information about the child's cognitive abilities) and at a level that looks for environmental causal factors. This model can guide an SLP to correctly assess a bilingual child suspected of language impairment and/or dyslexia.

Figure 3. Illustration of Morton and Frith's model done by Helland (2002)



Symptoms at behavioural/symptomatic and neuro-cognitive levels, and illustration of the cortical lobes at the biological level, Adapted from Morton & Frith 1995 (Helland 2002).

Bilingualism advantageous or disadvantageous

Paradis et al. (2011) discuss the connection between language and cognition by looking at two ways they connect. One aspect considers the cognitive capacity of bilingual children and the other considers cognitive consequences of learning two languages. They conclude that scientific evidence suggests that children have the innate/cognitive capacity to learn two different languages. Regarding the consequences of bilingualism on cognition, they state that it depends on variables such as the environment of acquisition of the languages, the type of bilingualism, and other circumstances surrounding the acquisition of the second language (Paradis et al. 2011, p. 38 & 53). According to Cummins (2000) the cognitive effects on bilingualism, be it positive or negative, depends on the level of bilingualism attained; a lower threshold level of bilingual competence will result in negative effects, conversely, a higher threshold level of bilingualism will have a positive effect on bilingualism (stated in Paradis et al., 2011, p. 52).

Bialystok (2001, p. 207-208) discusses study findings that show that bilinguals excel in tasks that require control on selective attention whereby they have the ability to *inhibit* unimportant information in the performance of the task. Such tasks from Bialystok and Martin (2004) involved four levels of classification each with different levels of conceptualization. Upon completion of each level, participants (bilingual and monolingual children) were given a new set of rules for the next level, and so on. Bialystok and Martin (2004) concluded from their findings that bilinguals have better *inhibitory control* for ignoring perceptual information than monolinguals. Paradis et al. (2011, p. 51) define *executive control functions* as the "activation, selection, inhibition, and organization of information during, for example, problem-solving". Bialystok (2001) and Bialystok and Martin (2004) provide empirical evidence that bilinguals experience a cognitive advantage that Paradis et al. (2011) collectively referred to as executive control functions.

However, other studies have challenged claims of bilingual cognitive advantage over monolinguals (Morton & Harper, 2007; Paap & Greenberg, 2013; Paap, Johnson & Sawi, 2015). Some of the critics were that studies that report evidence of cognitive advantages in bilingualism get publicity due to biases towards bilingual advantage.

Current empirical studies on bilingualism, language impairment, and cognition

This section reviews current empirical studies that look at issues related to cognitive and language performances among bilingual children vis-a-vis monolingual children. It also looks at studies that look at issues related to language assessment of bilingual children.

Gasquoine (2016) reviewed current different studies on effects of bilingualism on executive function, vocabulary, regional brain structure and dementia. The study reviewed studies from January 1999 to 2016 conducted in the United States and Canada. He points out that these studies that look at the effects of bilingualism on cognition and the brain, usually compare bilingualism and monolingualism and consider monolingualism as the norm. After reviewing these studies (Bialystok, 2007; Bialystok, Luk, Peets, & Yang, 2010; Calvo & Bialystok, 2014; Hoff et al., 2012; Portocarrero, Burright, & Donovick, 2007) Gasquoine (2016) found out that bilingual children have low scores (comparative to their monolingual counterparts) on vocabulary tests when assessed in only one of the languages. On explaining why bilinguals have low scores on language tests, Gasquoine (2016) makes references to Green (1998). According to Green (1998), when bilinguals use language, there is "parallel activation" of the languages they speak, which results in "interlanguage interference", so the processing time slows down and the possibility of making errors increases (cited in Gasquoine 2016, p. 989). According to Gasquoine (2016) current studies (such as Baum & Titone, 2014) support this theory that bilinguals have parallel activation of both languages during language processing. Referring to writings by Greenberg, Bellana, & Bialystok (2013), Gasquoine (2016) stated that the ability to activate and process two languages (*"inhibitory mechanism*")

has contributed to the ability to do cognitive tasks that requires inhibition, such as tasks on executive functions (Gasquoine, 2016, p. 989). Conclusions made by Gasquoine (2016) with regards to results from studies on "language-format" cognitive test was that bilinguals score lower than monolinguals and that newer studies do not support findings by older studies that bilingualism has a cognitive advantage over their monolingual counterparts.

Some of the studies Gasquoine (2016) looked at are revisited in the paragraphs that follow. As mentioned earlier on, findings by Bialystok (2001) and Bialystok and Martin (2004) support the notions that bilingualism has some cognitive advantage in the form of executive control functions. Kerrigan, Thomas, Bright and Filippi (2016) and Luo, Craik, Moreno, and Bialystok (2013) report visuospatial advantage among bilinguals over their monolingual counterparts. Kerrigan et al., (2016) reported that bilingual participants scored statistically higher than their monolingual counterparts on background measure of non-verbal reasoning.

A current study, on bilingualism and working memory, is done by Hansen et al. (2016). Their study had 152 participants in grades 2, 3, 5 and 8 (Spanish-English sequential bilinguals, and Spanish monolingual). The study observed a bilingual working memory advantage in younger age groups (2nd and 3rd graders), but there were no significant differences observed among older children (5th and 8th graders). With regard to verbal processing in L1, their findings were in line with other studies showing slower processing time among bilingual participants. Hansen et al., (2016) concluded that sequential bilingualism may "modulate" the development of working memory at a "subcomponent level".

A longitudinal (3 years) study on early bilingualism and working memory by Engel de Abreu (2011) tested the hypothesis that bilinguals may show more efficient working memory abilities than their monolingual counterparts. Test batteries used are similar to our current

study. Test batteries used for language comprehension were The Luxembourgish version of TROG and The Expressive One-Word Picture Vocabulary Test (by Brownell, 2000), which tests grammar and vocabulary respectively in L1. Cognitive tests included, among others, Digit Recall task, where participants were verbally presented sequences of digits which required immediate repetition of the digits in the right order. Other cognitive tasks included counting recall, backward digit recall, et cetera. Data from Engel de Abreu (2011) showed that monolinguals outperformed their bilingual counterpart on language tasks across years, but there were no group differences on working memory. Similar results are found by Paap and Greenberg (2013).

In their study on cognitive advantage associated with bilingualism, Engel de Abreu, Santos, Tourinho, Martin and Bialystok (2012) considered socio-economic status of participants. Scores from Portuguese-Luxembourgish bilingual children from low-income immigrant families (residing in Luxembourg) were matched with monolingual Portuguese children who live in Portugal. Results from visual-spatial tests on memory, abstract reasoning, selective attention, and interference suppression rather gives a positive result on bilingual advantage and support these views.

To our knowledge, there is no empirical study in a Norwegian setting where language abilities and cognitive variables are studied among native monolingual speakers of Norwegian and L2 speakers of the same language. The closest to such a study in Norway is a meta-study done by Lervåg and Melby-Lervåg (2009). Lervåg and Melby-Lervåg (2009) made a meta-analysis of other studies to draw out differences in verbal language, word-coding and reading comprehension between monolinguals and bilinguals. In their study, they considered how big the differences are between the two languages bilingual participants speak. Thus, they made the difference between European languages and non-European languages. Their study showed that bilingual speakers had clearly weaker verbal language proficiency than their monolingual

counterparts had, in the same language. They also noted that there were less verbal language differences between the monolinguals and bilinguals if both the L1 and L2 languages of the bilingual were European than if the L1 was Asian or African and the L2 was European.

Other studies that compared L1 and L2 Norwegian speakers are from preschool. Karlsen, Lyster, and Lervåg (2016) looked at the development of vocabulary among preschoolers with L2 in Urdu. They found no developmental differences among monolinguals and bilinguals, despite weaker vocabulary skills among bilingual participants. Karlsen, Lyster, and Geva (2016) investigated the contribution of cognition, linguistic and contextual factors to narrative production in Norwegian L2 learners (who have Urdu/Punjabi as L1). Their study shows that cognitive, linguistic and contextual factors measured in kindergartens help to explain individual differences in different aspects of L2 story production a year later, although performance in L1 does not.

In summary, our discussions above seem to show that most of these studies on bilinguals are done with participants who balanced bilingualism, in an additive bilingual environment or sequential bilingualism with two equally prestigious languages (or languages with equal social standards). We also saw that most studies on bilingualism conducted in Norway is done with participants at the preschool age, and to our knowledge, none of the studies considers the relationship between cognition, L2, and whether bilingualism gives some cognitive advantages to school children, and what role L2 or bilingualism plays in this relation. None of the Norwegian studies, to our knowledge, sought to answer problems concerning bilingual assessment either. Our study sets out to look at third graders, with diverse L2 background, who have Norwegian as a second language, and who live in a subtractive bilingual environment. This is an interesting group to examine because the bulk of studies done in a Norwegian setting are mostly on monolinguals whereas bilingual studies examine other aspects of bilingualism.

Statement of the problem

From our earlier discussions, we discussed that minority school children mostly lag behind their L1 counterparts in school performances. We also looked at discussions related to using L1 based language assessment tests in clinically assessing L2 children by SLPs, which is often the case in Norway. We also noted that the few Norwegian studies on bilingualism do not address the problems in clinically assessing bilingual children. On the bases of these points raised, we make the following problem statement, which provides the context of this current study and generates the questions which the study aims to answer. "*Assessing bilingual children with only single language tests will over-identify, under-identify, or misidentify them as having SLI and/or dyslexia. There is the need for an evidence based-practice in the assessment of bilingual children based on Norwegian empirical research.*"

Current Study

The aim of the study

The main objective of this study was to investigate the separation between language interference and language impairment in bilingual children using the differential diagnostic method by Morton and Frith. Other objectives were as follows:

1. To find out if subjects in the L2 group exhibit cognitive advantages over those in the L1 group.
2. To find out how information on cognition can contribute to clinical assessment of bilingual children.
3. We aimed to analyze each bilingual participant's language and cognitive profiles in order to find "symptoms" of language impairment and/dyslexia.

Research Questions

1. Do bilingual children exhibit cognitive advantages over their monolingual counterparts?

2. Can bilingual children be identified by Morton and Frith's *differential diagnosis* model (revised by Helland, 2002; see Figure 3) as having SLI/dyslexia?
3. How different are the linguistic abilities (in L1) and cognitive abilities of bilingual children compared with their monolingual peers?
4. Within the bilingual group do subjects exhibit "signs" of language interference or language impairment and/or dyslexia.

Hypothesis

Based on our research questions and discussions above we make the following predictions;

1. The differential diagnostic method by Morton and Frith (revised by Helland, 2002) can identify bilingual children who have SLI and /or dyslexia.
2. The L1 group will outperform the L2 group on single language tests but cognitive tests will reveal no group differences.

Methodology

Research design

Our thesis is part of the ongoing project "Norwegian as a second language (L2) in 3rd grade school children. How can transient between-language problems be separated from problems due to more specific language impairment and/or dyslexia?" The project is managed by associate professor Wenche Andersen Helland, associate professor Frøydis Morken and Professor emerita Turid Helland, all from the Department of Biological and Medical Psychology, University of Bergen. In this study we used control data from the study "Tell a tale - a study of the language of 3rd graders" conducted by the Bergen Logopedic Research group (B.LOG) (Torkildsen, Morken, T. Helland & W.A. Helland, 2015). We have used the same questionnaire and tests as in this study.

The research design of a study spells out the basic strategies that researchers adapt to develop evidence that is accurate and interpretable (Polit & Beck, 2012, p. 180).
In this study, we are using a quantitative design because we are collecting data that is numeric and can be analyzed statistically (Polit & Beck, 2012, p. 14). We have compared two groups of children, one group having Norwegian as L2 and the control group having Norwegian as L1, using the same assessment tools for both groups. This is classified as a qualitative between-subject design (Polit & Beck, p.182).

This study is also a cross-sectional design because the data was collected during a single period of time; describing the phenomena we were investigating at a fixed point in time (Polit & Beck, 2012, p. 184). The participants in the study were not given any training or other interventions before or after the testing. When researchers do not intervene by manipulating the independent variable, the study is non-experimental or observational (Polit & Beck, 2012, p. 223).

An epidemiological study is according to Carter and Lubinsky (2016, p. 145) a type of research documenting the occurrence of disease or injury, determining causes of the disease or injury, or developing mechanisms to control the disease or injury. In our study, we will try to find out if the difficulties of bilingual children may be due to SLI or dyslexic difficulties.

Sampling

The project managers informed school managements about the project and participants were recruited from four of these schools. Written information about the project, a letter of inquiry of participation in the study, and declaration of consent was handed out to parents by the schools. Parents were also given the unpublished questionnaire Risk Index-8 (RI-8), an age-adjusted version of RI-5 (Helland, 2015; Helland, Plante & Hugdahl, 2011), a parental questionnaire identifying the early risk of dyslexia (Helland et.al., 2011). The native language teachers at the schools were asked to translate the questionnaire and other written information

into the native language of the parents. We know that only a few participants received this help, making it difficult to rely on the information given in the questionnaire. The questionnaire was answered and returned by 19 of the 20 parents.

Inclusion criteria for participating was having attended Norwegian school since 1st grade, no identified mental retardation, and no identified neurological disorder. When planning the project, we wanted pupils that teachers and/or parents had concerns about regarding the progress of linguistic and literacy learning. Having this criterion turned out to be a problem getting enough participants, and the project managers decided that having Norwegian as L2 was sufficient to participate in the study. The sample we got was twenty 3rd graders. There were 12 boys and 8 girls, two of the children left-handed. The participants came from 13 different countries and spoke 10 different languages. Four of the participants came from European countries (Slovakia, England, Hungary, and Poland), six from African countries (Somalia, Morocco), seven from Asian countries (Iraq, Syria, Sri Lanka, and Palestine), and one from a South American country (Colombia). For two of the children we only know that they spoke Arabic, not which country they came from. Their parents had a mean of 12,3 years of education, which is under the national average in Norway. According to SSB (2017), 32,9 % of the Norwegian population has higher education (1 year or more of college or university studies). Five of the parents did not give information on education level, so this number is not reliable.

The control group consisted of 42 participants, 16 girls and 26 boys, age 7 and 8, 11 of the children left-handed. They all had Norwegian as L1, and their parents had an education level close to the national average (Torkildsen et.al., 2015).

Data collection

Data collection took place in November and December 2017 and was carried out by four master students in logopedics from the University of Bergen. The project leaders trained

us in administering the different tests. We also trained together in pairs testing each other to make sure that we knew the test procedures well, and to eliminate problems that could occur during testing with the children. The authors of this thesis worked together and tested ten of the participants in two of the schools and our fellow students, Marianne Dahl and Hanna Nordbø, tested the remaining participants in the two other schools. Before we started testing the four of us discussed the details of the test procedures and agreed in what order the tests should be executed to ensure that the test situation would be as similar as possible for all the participants.

We contacted the teachers in charge of the project in advance and agreed on time schedules and details of the practical implementation of the testing. We made sure that we had a room where we could be undisturbed and got timetables for the classes so we could plan the testing regarding breaks and subjects like gym that the children would not miss out on.

All testing was done at school, during school hours, and it took about two hours per child to complete the assessment. The children were alone with the test leaders during testing and we tried to make the environment for the assessment as positive as possible, making sure that they had breaks when needed, and had access to drinks and biscuits to keep the spirit up. The test battery consisted of 11 different tasks that took between five and twenty minutes. The test leaders shared the tests between themselves, one test leader was responsible for the implementation of every second test, and the other test leader observed and helped out with practical issues like timing and recording.

We managed to finish the testing in a day for most of the children, but in one school we were not allowed to use the day-care facilities for school children (SFO), because the school thought it would stigmatize the children. We respected that decision, but it meant that three of the participants were tested over a period of two days.

After the data collection, the four of us scored the different tests according to the manuals, and the results were collated in a common data matrix in the statistical data programme Statistica (StatSoft, 2011).

Test battery used in this study

In the designing and planning of a study, it is important to find assessment tools that are grounded in evidence-based practice, with high validity and reliability. In Norway, the selection of standardized tests with Norwegian norm scores is limited compared to what is available for English speaking countries (Helland, 2012, p. 268). In this study the test-battery we have used has Norwegian norm scores, and most of the tests are translated from English and adapted to Norwegian standards. In principle, all assessment of bilingual children should be done in L1, with diagnostic tools for L1. In this case, we are not using the tests to diagnose children with SLI and/ or dyslexia, but to identify between-language problems from SLI/ dyslexia, using the same test battery for both groups. All the assessment tools we have used have focused on the children's abilities in L2, we have not assessed their first languages.

The test batteries used in this study are based on the causal model by Morton and Frith (1995). It was meant for them to be able to differentially diagnose our participants who might have language impairment. One of the test batteries was in a form of a questionnaire (RI-8, Risk Index for 8-year olds, an unpublished age-modified version of R1-5 (Helland, 2015; Helland et. al., 2011)), which provided important background information on participants from parents. In addition, we used 11 different tests on the subjects basic linguistic and literacy skills in L2 (Norwegian) and neurocognitive factors. In our thesis, we have focused on language, linguistic skills, and neurogenic factors. Because of that, we will not go into details on the tests about written narratives and literacy skills, Single word reading and spelling (STAS**)** (Klinkenberg & Skår, 2001), and Written narratives using keystroke logging.

Risk Index-8 (RI-8): The questionnaire Risk Index-8 is an unpublished age-modified version of the questionnaire Risk Index-5 (Helland, 2015), developed to identify the early risk of dyslexia (Helland et al., 2011). The form is to be answered by parents and/or teachers and gives information on the following subjects: general health, allergies, asthma, left-handedness, motor-skills, language, special needs and heredity (Helland et al., 2011). In this study, it was only parents who filled-out the questionnaire.

British Picture Vocabulary Scale-II (BPVS II): British Picture Vocabulary Scale (BPVS II) (Dunn, Dunn, Whetton, & Burley, 1997) is a standardized test to map out children's receptive vocabulary and detect delays in vocabulary development. The Norwegian version of BPVS II is translated and adapted to Norwegian from the second English version (Lyster, Horn & Rygvold, 2010).

Test for Reception of Grammar, version 2 (TROG-2): Test for Reception of Grammar, version 2 (TROG-2) (Bishop, 2003) is translated and adapted to Norwegian for the age group 4-16 years (Horn & Lyster, 2009). TROG-2 is a receptive language test that assesses understanding of grammatical contrasts marked by inflections, function words and word order. TROG- 2 makes it possible to compare grammatical comprehension with peers of the same age, and to find areas of difficulty (Bishop, 2003).

The model sentence task based on "Ringstedmaterialet": The model sentence task based on "Ringstedmaterialet" (Ege, 1984) is used to measure the expressive language in children, and can give information about phonology, morphology, syntax, and semantics. The test is challenging when it comes to skills in attention, memory, maturity and language awareness (Helland, 2012, p.51). This test has no normed scores. We used a revised version of the material developed for the longitudinal study "Speak up!" (Helland et al., 2011).

Rapid naming (RAN): Rapid naming (RAN) from the Stroop battery (Golden, 1987, Hugdahl, undated version) is a test that measures verbal processing speed given visual

stimuli. Rapid naming (RAN) is a skill composed of multiple factors like attention, perception, concept formation, memory, phonology, semantics, and motor skills. A combination of difficulties with phonological awareness and rapid naming (RAN) is characteristic for people with severe dyslexia difficulties, regardless of language affiliation (Helland, 2012, p. 108).

We scored the tests after standard procedures (Hugdahl, undated version). This test has no normed scores.

Verbal short-term and working memory, digit span: We used the test digit span from the Wechsler test battery (Wechsler, 1991, 2003) to test verbal short-term and working memory. Difficulties with verbal short time memory and working memory are typical for children with SLI or dyslexia. Verbal short time memory is the ability to remember verbal information for a short period of time. Working memory is the capacity of both storing and processing information for a short period. A typical way of testing verbal short time is digit spans forward, and testing working memory is digit spans backward (Helland, 2012, p.73). The test was scored after standardized procedures (Wechsler, 1991, 2003).

***Executive functions using dichotic listening test*:** The Dichotic listening test (Bless, Westerhausen, Kompus, Gudmundsen & Hugdahl, 2014) measures how language sounds are perceived in the brain. We used the app iDichotic, developed by researchers and scientists affiliated with the Bergen fMRI group. The app is based on the consonant-vowel dichotic listening test, often used in the neuropsychological assessment of language laterality (Bless et. al., 2014).

Visuospatial skills using the Rey-Osterrieth Complex Figure Test: Rey-Osterrieth Complex Figure Test (Meyers & Meyers, 1995) is used to measure children's visuospatial skills and visuospatial construction (Helland, 2012, p. 135). Studies show that some people with dyslexia have problems with the dorsal visual network, resulting in problems in copying

or recall visuospatial pictures (Helland, 2012, p.135). The scoring of RFCT gives room for discussions so all four test leaders sat together and scored to make sure the result would be as accurate as possible.

Statistical analyses

We used the computer program Statistica (StatSoft, 2011) to analyze the collected data. Statistical hypothesis testing is how researchers can make objective decisions about the likelihood of their results reflecting chance sample differences or true population differences (Polit & Beck, p. 408). Statistical inference starts with a statement of the null hypothesis and a research hypothesis. The null hypothesis means that the population means are equal; any observed difference is due to random error; the independent variable had no effect. The research hypothesis is that the population means are not equal; the independent variable did have an effect (Cozby, 2009, p. 246). Researchers can make two types of statistical error, rejecting a true null hypothesis (Type I error), or accepting a false null hypothesis (Type II error) (Polit & Beck, p.409). Statistical significance means that the obtained results are not likely to be the result of chance, at a specified level of probability. When statistical results are beyond this limit it is said to be statistically significant. A non-significant result means that the results could reflect chance fluctuations (Polit & Beck, p. 410). The level of significance in all our analyses was set to $p < .05$, which is the minimum acceptable level to avoid type I errors, that the results are based on chance factors (Polit & Beck, p. 409).

In this study, we tested the differences between two independent groups. The test used for this purpose was students independent samples t-test. The t-value is the difference between the group means and the variability within the groups. The group difference is the difference between the obtained means, under the null hypothesis the difference is zero. The value of *t* increases as the difference between the obtained samples means increases, and the level of significance is higher (Cozby, 2009, p. 250).

We used an independent t-tests to find out if there were statistically significant differences between the two groups on scores from both cognitive and language variables. We also measured effect size, by comparing the two groups' means measured in standard deviation units, reported as Cohen`s *d.* The standard interpretation of Cohen`s *d* is .20 – small, .50 – moderate, and .80 – large (Polit & Beck, p.424). To calculate the effect size, we used the Effect Size Calculator for T-test (socscistatistics.com).

We used analyses of variance (ANOVA) for the model sentence task. ANOVA is the parametric procedure for testing differences between means when there are three or more groups. The statistic computed in ANOVA is the *F*-ratio. Variation between groups is contrasted with variation within groups to get an *F*-ratio. When the differences between the groups are large compared to the differences within the group, it is likely that the independent variable has caused the differences (Polit & Beck, 2012, p. 416).

In descriptive statistics, it is important to know if a relationship between variables is relatively weak or strong. A correlation coefficient describes how strongly variables are related to one another. The Pearson product-moment correlation coefficient (Pearson`s *r)* is used when both variables have interval or ratio scale properties. Values of Pearson`s *r* ranges from 0.00 to + - 1.00. A correlation of 0.00 indicates no relationship between variables, the nearer to 1.00(+-) the stronger relationship (Cozby, 2009, p. 230). To measure if there were any relationship between the tests used in the study we performed a correlation analysis. We also did a case by case- study where we analyzed each subject`s score that deviated with *1 SD* or more to the mean of the control group, to identify the children being at risk of developing SLI and/or dyslexia.

Validity and reliability

Validity

Validity means truth, and it is about accurate representation of information (Cosby & Bates, 2009, p. 85). Validity concerns the soundness of the study's evidence, that the findings are unbiased and well-grounded (Polit & Beck, 2012, p.175). There are different types of validity, but it is most common to use the terms internal and external validity.

***Internal validity*:** Internal validity is when the interventions in a study are responsible for the observed effects, rather than uncontrolled extraneous factors (Polit & Beck, 2012, p. 731). In our study, it was mainly the test situation that could weaken the internal validity. As far as possible we made sure that no outside factors would influence the results, causing bias. It was important that we as researchers made sure that only the independent variable rather than anything else could cause empirical relationship, and also that the methods of the study actually measured what they were supposed to do (Polit & Beck, 2012, p. 175).

To avoid weakening the internal validity of the implementation of the assessment tools, we received training from the project managers on procedures of conducting the tests, and we later spent a lot of time practicing the testing on each other. Those of us with children, whose ages were within the required age group for the assessment tools, practiced testing at home with them. When we started the actual testing on participants, we worked in pairs, being able to observe and help each other out if a problem occurred. When we first started testing, the whole procedure took 2,5 hours, in the end, it took less than two hours. We got better at administering the tests and more professional in the test situation during the time we tested, a fact that could mean that the last children we tested got a better result than the first ones.

The fact that we worked in pairs could also be a problem concerning internal validity. We are four individuals with different personalities, so the test situation would be

different depending on who did the testing. We discussed this problem and made guidelines for test procedures regarding test order, and the practical implementation.

Testing situations and conditions during testing were not completely identical from school to school. This was because each school had a different system of practice and culture. This could affect the internal validity, but our joint-training and joint preparations prior to the implementation of the testing was meant to avoid issues as such from weakening the internal validity.

The control group for our study was subjects from Torkildsen et.al. 2015. In the control group, all participants came from the same school environment, were taught by the same team of teachers and followed the same curriculum. The fact that we have tested a less heterogeneous group can be a variable that can affect the results statistically.

To get background information on participants in the study, parents were given the questionnaire RI-8, an unpublished, age adjusted version of RI-5 (Helland, 2015). However, upon evaluation of the questionnaires received, it was determined that their validity was not up to sets standards due to translation problems and other factors. As a result of that we could not use them.

All four students scored the tests according to the manuals and collated the results in a data matrix under the supervision of our supervisor Turid Helland. We worked in pairs scoring the tests to avoid errors. Both modelling sentences and RO are tests where the scoring procedures are subjective with room for discussions. To make sure that the scoring of RO was as accurate as possible, all four of us discussed and scored these tasks together.

***External validity*:** External validity of a study is the extent to which the results can be generalized to other populations or settings (Cozby, 2009, p. 86). The aim of this study, as earlier stated, was to separate transient between-language problems from problems due to

more specific language impairment and/or dyslexia. Our goal was that the results of the study can be generalized to the whole population of 3rd graders with Norwegian as L2.

Our sample for the L2 group was 20 children with Norwegian as L2. They had backgrounds from 13 different countries and had L1s of 11 different languages. Even though the sample size was not large, the variation in nationalities and the different languages participants speak give the sample higher external validity because it gives room for comparison of results within the L1 and provides information on the typology of these languages with relation to their L2. We also have a control group of 42 children, which strengthens the power of the sample.

Reliability

Reliability refers to the accuracy and consistency of information obtained in a study and is associated with methods used to measure variables, or statistical reliability, referring to the probability that the results would hold with a wider group tested (Polit & Beck, 2012, p. 175).

Test reliability is a major criterion in any quantitative study and is the degree of consistency or accuracy with which an instrument measures an attribute. The higher reliability the lower amount of error in obtained scores (Polit & Beck, 2012, p. 348).

Internal consistency reliability is the most widely used reliability approach, and the method used to evaluate it is coefficient alpha (Cronbach`s alpha). The range of values is between .00 and +1.00, higher values reflecting higher internal consistency, the extent to which different subparts of an instrument (test) are reliably measuring the critical attribute (Polit & Beck, 2012, p. 333).

In our study, we have used different tests to measure basic linguistic and literacy skills in Norwegian as L2 and neurocognitive factors. The test batteries used in the study have been

used in various studies at the University of Bergen before which in itself is a sign of high reliability of the tests.

BPVS-II is translated and adapted to Norwegian from the second English version (Lyster, Horn & Rygvold, 2010). This test has Norwegian norm scores and was tested on 884 children at the age of 3-16 years. The results showed good reliability and a high internal consistency (.86 coefficient alpha) (Dunn et.al., 1997).

TROG -2 is translated and adapted to Norwegian and has got Norwegian norm scores. It was tested on 950 Norwegian children between the ages of 4-16. The reliability is considered very high, with an internal consistency of .95 (coefficient alpha) (Bishop, 2003; Lyster & Horn, 2009).

The model sentence task, based on Ringstedmaterialet (Ege, 1984) is not a standardized test, and it has no Norwegian norm scores. The test leaders scored the tests together, discussing and agreeing on the result between them. This could lead to higher reliability of the scores.

To test the working memory, we used the Digit Span task from the Wechsler Intelligence Scale for Children-Third Edition (Wechsler, 1991, 2003). The score was the sum of the forward and backward digit span. The internal consistency for the age group 8 was 0.76 (coefficient alpha) (Wechsler, 2003), showing high reliability.

Rapid naming (RAN) is a well-known test used in psychology testing verbal processing speed using visual stimuli. We used the test colour naming from the Stroop battery (Golden, 1987, Hugdahl, undated version). This test has no Norwegian norm scores. The test has been used in several studies (Helland & Morken, 2015; Helland et al., 2011) and has got high reliability.

Rey-Osterieth Complex Figures Test with the Copy and Recall conditions (Meyers & Meyers, 1995; Spreen & Strauss, 1991) is a well-recognized test to measure visuospatial

skills. The test is standardized and has got Norwegian norm scores. All four students scored the RCFT together, to make sure that the results would be as accurate as possible and not affected by individual subjectivity. This could increase the reliability of the test results.

Dichotic listening test, iDicothic (Bless, et. al., 2014) is a commonly used test in the neuropsychological assessment of language laterality. We used the iDichotic app to measure auditive and executive functions. The testing was easy to administer using an Ipad/ smartphone with headphones, and the results were calculated by the app immediately after the testing was done, making the results more reliable.

Ethical considerations

The Project plan was accepted and approved by the Norwegian Centre for Research Data (NSD). The application was sent in June 2017, and the approval was given in August 2017. In this study, we have executed a variety of linguistic and cognitive tests on 8-year-old children with Norwegian as a second language, L2. Testing on children is challenging, and the fact that this was a particularly vulnerable group made us extra aware of the ethical challenges that this entailed.

According to § 5 of the Act relating to medical and scientific research;

" *Medical and health-related research should be conducted properly. The research shall be based on respect for the human rights and human dignity of the research participants. The welfare and integrity of the participants shall be in front of the interests of science and society. Medical and health research shall cover ethical, medical, health, scientific and privacy issues*. "

In "Children and Consent - What does Norwegian law say?" (Forskningsetikk nr. 4, 2013), the most important legal principles in relation to children and research in Norway are addressed. It states that the principle of Norwegian law is that only persons over the age of 18

have the formal legal capacity, including consent to participate in medical and health research. Children under the age of 12 have no right to decide for themselves about participating in research, but a procedural right to be heard. The Children's Act § 31, second paragraph states that when the child has reached the age of 7, their opinions should be considered before decisions about personal issues are taken. The participation in research will clearly fall under personal issues.

Self-determination is also an important principle of law, which is enshrined in § 33 of the Children's Act. In practice, this is often overruled because it is a public opinion that children are not able to make important decisions on their own behalf and therefore shall not be entitled or obliged to make such decisions. It is therefore up to parents or guardians to give consent that the children can participate in research. It is important to keep in mind that parents' consent should be in the interests of the child and to the best interests of the child. This is anchored in the Children's Act § 30, paragraph 1 and in the Convention on the Rights of the Child, article 3. It must be an overall rule that a child cannot be forced to participate in research against its will. At the same time, it is important that the requirements for consent do not get too strict.

As mentioned earlier, we have performed various tests on 20 children with Norwegian as L2. They have backgrounds from different countries, with different cultural backgrounds, and different first languages. This is an extra vulnerable group and we tried to make the test situation as safe and positive as possible. It was therefore extra important that we mastered the tests procedures in a professional manner and that we were prepared for different reactions from the children. For us it was important to treat them with respect and make sure they had a positive experience when they were being tested.

Upon issuing the RI-8 questionnaire, an unpublished, age modified version of RI-5 (Helland, 2015) to parents, it was important that we ensured that the parents were able to

understand and correctly fill-out the forms. We therefore sought the help of native language teachers at the various schools, as interpreters. In retrospect, we found out that only a few of the parents were offered this help. Furthermore, it is important that parents and teachers are informed about the results of the study. We hoped that this study would be informative, educational and beneficial to teachers, parents and all caregivers who are involved in the daily lives of our participants.

All collected data have been given to the project managers and are kept safe, and the participants were assigned numbers to make sure that they are anonymous. The test leaders having the direct contact with the children have signed an agreement of confidentiality regarding the privacy of the children and parents participating in the study.

Results Excluded from the Article

Descriptive statistics

Descriptive statistics for all measures are reported in Table 1. As mentioned earlier the subjects for this study are solely 3rd graders (in Norway approximately 8-year-olds) and the differences in the mean age between the bilingual group and the monolingual group is since we received only 9 confirmed birthday dates from parents. The number of years of education for 17 of the mothers of subjects in the L2 group was $N = 17$; $M = 12.71$, and for 16 of the fathers in the same group was $N = 16$; $M = 13.53$.

Case study within the L2 Group

We analyzed each score for each measure by each subject in the bilingual group to see how they deviate from the group mean by *1 SD* (see Table 2). The results revealed that three of the children were *1 SD* or more above mean on four of the tests, both linguistic and neurocognitive (compared to their bilingual peers). The analysis also showed that some subjects in the bilingual group stand out with performances in both language and cognitive

tests mostly above *1SD* or more of the mean of the group. Results from the analysis are shown in Table 2.

Table 1

Descriptive statistics for all measures

Variable	Mean L1 *N=42*	Mean L2 *N=20*	Median L1 *N=42*	Median L2 *N=20*	Min L1 *N=42*	Min L2 *N=20*	Max L1 *N=42*	Max L2 *N=20*	Range L1 *N=42*	Range L2 *N=20*	SD L1 *N=42*	SD L2 *N=20*
Age month	98.81	101.44	99.00	102.00	93.00	94.00	104.00	106.00	11.00	12.00	3.34	3.84
DS F	7.10	6.30	7.00	6.00	4.00	4.00	12.00	9.00	8.00	5.00	1.85	1.53
DS B	3.95	3.45	4.00	3.00	1.00	2.00	7.00	6.00	6.00	4.00	1.34	1.28
DS Total	11.05	9.75	11.00	10.00	6.00	7.00	18.00	14.00	12.00	7.00	2.67	1.77
RAN, sec	48.76	59.53	44.00	58.39	28.00	36.45	103.00	104.71	75.00	68.26	16.13	17.12
RO1	16.33	22.48	16.00	22.25	6.50	2.50	26.00	34.00	19.50	31.50	4.62	8.26
RO2	7.58	14.93	7.25	16.00	0.00	3.50	20.00	31.00	20.00	27.50	4.12	7.69
BPVS	92.02	64.55	91.00	66.00	61.00	32.00	115.00	85.00	54.00	53.00	13.90	15.85
TROG	102.05	78.70	99.00	81.00	65.00	55.00	122.00	103.00	57.00	48.00	12.66	13.27
Mod morph	16.60	13.15	17.00	14.00	12.00	7.00	20.00	18.00	8.00	11.00	2.13	3.51
Mod synt	16.26	12.85	16.00	13.00	9.00	6.00	24.00	20.00	15.00	14..00	2.90	3.95
Mod Sem	16.14	13.70	16.00	14.00	12.00	6.00	19.00	18.00	7.00	12.00	2.10	3.29
Mod SUM	49.00	39.70	48.00	41.50	39.00	22.00	58.00	54.00	19.00	32.00	5.84	9.18
DL NF Re	11.60	14.55	11.00	14.00	3.00	8.00	21.00	22.00	18.00	14.00	3.46	3.47
DL NF Le	10.48	9.35	10.00	9.00	0.00	5.00	17.00	18.00	17.00	13.00	3.08	3.12
DL FR Re	14.56	15.40	14.00	15.00	8.00	9.00	22.00	24.00	14.00	15.00	3.24	4.17
DL FR Le	9.00	9.20	9.00	9.50	3.00	2.00	13.00	16.00	10.00	14.00	2.15	3.38
DL FL Re	13.41	13.84	14.00	14.00	5.00	8.00	24.00	24.00	19.00	16.00	3.66	4.65
DL FL Le	9.87	8.53	10,00	9.00	2.00	2.00	21.00	13.00	1900	11.00	2.98	2.89

Note: F = Frequency (of Mode); Min = minimum; Max = Maximum; SD = Standard deviation; D F = digit span forward; DS B = digit span backwards; DS Total: Digit span total score; RAN = Stroop test; RO1 = Rey-osterrieth complex figure; TROG = Test for Reception of Grammar, version 2; Mod: model sentence task (morphology, syntax, semantics); DL NF Re = dichotic non-forced right; DL NF Le = dichotic non-forced left; DL FR Re = dichotic forced-right right; DL FR Le = dichotic forced-right left; DL FL Le = dichotic forced-left left

Table 2

Case By Case Study for L2 Group Analyzed from the Mean of the Group.
X = >1SD below mean; * = > 1 SD above mean; Empty slot indicate that subjects are within normal range

No.	ID	BPVS	TROG	MOD SUM	Digit span forward	Digit span back-ward	RAN	RO1	RO2	DL NF Re	DL NF Le
1	101	X		X			X		X		
2	102		*							X	
3	103	*	*	*				*	*		
4	104							*		X	
5	105	*	*		*					X	*
6	106	X				X	*	X	X		
7	107	*						X	X		
8	201		X				*	*	*		
9	202									*	X
10	203					X	X				
11	204		X	X		X			X		
12	205			*	X			*			
13	206				X	*		*	*		
14	301			X	*					*	X
15	302	X	X	X	*	X					
16	303	*							X		
17	401					*	*				
18	402				*			X			
19	403			*	*						*
20	404	*			*	*				*	X

Note: BPVS Mean = 64.55; TROG mean = 78.70; MOD SUM mean = 39.70; Digit span forward mean = 6.30; Digit span backward mean = 3.45; RAN mean = 59.53; RO1 mean = 22.48; RO2 mean = 14.93; DL NF Re mean = 14.55; DL NF Le mean = 9.35

References

Anderson, N. B., & Shames, G. H. (2014). *Human Communication Disorders An introduction (6th ed.)*. UK: Pearson Education Limited

Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.). (2015). *Assessing multilingual children: disentangling bilingualism from language impairment.* Retrieved from https://ebookcentral-proquest-com.pva.uib.no

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208. http://dx.doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–40. doi:10.1016/j.cub.2009.12.014

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*, 525–531. doi:10.1017/S1366728909990423

Bialystok, E. & Martin, M. (2004), Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science, 7:* 325-339. doi:10.1111/j.1467-7687.2004.00351.x

Bialystok, E. (2001). *Bilingualism in development : language, literacy, and cognition.* Retrieved from https://ebookcentral-proquest-com.pva.uib.no

Bishop, D.V. M. (2003). The Test for Reception of Grammar, version 2 (TROG-2). London: Pearson Assessments.

Bjørndal, A., & Hofoss, D. (2010). *Statistikk for helse- og sosialfagene* (2.utg.). Oslo: Gyldendal Norsk Forlag.

Bless, J. J., Westerhausen, R., Kompus, K., Gudmundsen, M. & Hugdahl, K. (2014). Self-supervised, mobile-application based cognitive training of auditory attention: A behavioral and fMRI evaluation. *Internet Interventions, 1*(3), 102-110. doi:http://dx.doi.org/10.1016/j.invent.2014.06.001

Bloom, L. & Lahey, M. (1978), Language Development and Language Disorders. New York: John Wiley & Sons. *Columbia University Academic Commons*. Downloaded from https://doi.org/10.7916/D8QZ2GQ5

Brownell, R. (2000). *Expressive one word picture vocabulary test.* London: Psychological Corporation.

Carter, R. E., & Lubinsky, J. (2016). *Rehabilitation Research. Principles and Applications.* (5. ed.) Missoury: Elsevier.

Cosby, P. C., & Bates, S. C. (2012). *Methods in behavioral research* (11.utg.). International edition: McGraw-Hill.

de Jong, Jan. (2008). Bilingualism and language impairment. In Martin Ball, Michael R. Perkin, Nicole Müller & Sara Howard (eds.), *Handbook of Clinical Linguistics, 261-274*. Oxford: Blackwell Publishing. Retrieved from ProQuest eBook Central database.

Dunn, L. M., Dunn, D. M., Whetton, C. & Berley, J. (1997). *The British Picture Vocabulary Scale II*. Windsor: GL Assessment.

Ege, B. (1984). *Ringstedmaterialet. En procedure til undersøkelse af børns realiserbare sprog*. Herning: Special-pedagogisk forlag A/S.

Egeberg, E., (2016). *Minoritetsspråk og flerspråklighet. En håndbok i utredning og vurdering*. Oslo, Cappelen Damm Akademisk.

Engel de Abreu P. M. E., Cruz-Santos A., Tourinho C. J., Martin R., Bialystok E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychological Science, 23* 1364–1371. DOI: 10.1177/0956797612443836

Engel de Abreu, P. M. J., (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory, 19:5*, 529-537, DOI: 10.1080/09658211.2011.590504

Fossheim, H., Hølen, J., Ingierd, H. (2013). *Barn i forskning- etiske dimensjoner*. De nasjonale forskningsetiske komiteene. Downloaded from https://www.etikkom.no/globalassets/documents/bladet-forskningsetikk/alle-utgaver/130631_forskningsetikk_nr4_150dpi.pdf

Gasquoine, P. G. (2016). Effects of bilingualism on vocabulary, executive functions, age of dementia onset, and regional brain structure. *Neuropsychology*, *30(8),* 988-997. http://dx.doi.org/10.1037/neu0000294

Golden, C. J. (1987). *Manual for the Stroop Color and Word Test. Chicago*: Stoelting Co.

Hoff, E. (2014). Language development. 5th ed. Wadsworth Cengage Learning. US.

Hansen, L. B., Macizo, P., Duñabeitia, J. A., Saldaña, D., Carreiras, M., Fuentes, L. J. and Bajo, M. T. (2016), Emergent Bilingualism and Working Memory Development in School Aged Children. In *Language Learning, 66*: 51–75. doi:10.1111/lang.12170

Helland, T. (2012). *Språk og dysleksi*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.

Helland, T. (2002). *Neuro-Cognitive Functions in Dyslexia. Variations According to Language Comprehension and Mathematics Skills.* Unpublished Doctoral thesis, Department of Special Education, Faculty of Education, University of Oslo.

Helland, T. (2015) RI-5. *Hva er Risiko Indeks 5? Spørsmål og svar.* Retrieved from https://ri5.infovestforlag.no/nedlastinger/dokumenter/FAQ_BM.pdf

Helland, T. & Morken, F. (2015). Neurocognitive development and predictors of L1 and L2 literacy skills in dyslexia: A longitudinal study of children 5-11 years old. *Dyslexia, 22*(1), 3-26. Doi: 10.1002/dys.1515

Helland, T., Plante, E. & Hugdahl, K. (2011). Predicting Dyslexia at Age 11 from a Risk Index Questionnaire at Age 5. *Dyslexia, 17*(3), 207-226. doi: 10.1002/dys.432

Hugdahl, K. (undated version). *Stroop Test.* Bergen: University of Bergen.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and aging, 28*(1), 28-34. doi:10.1037/a0030875

Karlsen, J., Lyster, S.A.H. &, Geva, E. (2016). Cognitive, linguistic and contextual factors in Norwegian second Language learners' narrative Production. *Applied Psycholinguistics*. Doi: 10.1017/S014271641500051X

Karlsen, J., Lyster, S.A.H. & Lervåg, A.O. (2016). Vocabulary development in Norwegian L1 and L2 learners in the kindergarten–school transition. *Journal of Child Language.* ISSN 0305-0009. doi: 10.1017/S0305000916000106

Kerrigan, L., Thomas, M. S. C., Bright, P., & Filippi, R. (2016). Evidence of an advantage in visuo-spatial memory for bilingual compared to monolingual speakers. *Bilingualism: Language And Cognition, 1–11*. doi:10.1017/S1366728915000917

Klinkenberg, J. E. & Skaar, E. (2001) STAS. Standardisert test i avkoding og staving. Hønefoss: Pedagogisk psykologisk tjeneste.

Kohnert, K. (2013) *Language disorders in bilingual children and adults*. 2. San Diego, CA: Plural. Retrieved from ProQuest eBook Central database.

Larsen, T. A. (2013). Barn og samtykke- hva sier norsk rett? *Forskningsetikk*, *4,* 14-15. Retrieved from https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/barn-i-forskning-web.pdf

Lervåg, A. & Melby-Lervåg, M. (2009). Muntlig språk, ordavkoding og leseforståelse hos tospråklige: En sammenfatning av empiriske studier. Norsk pedagogisk tidsskrift, 2009 (4), 264-277. Retrieved on the14.februar 2018, from https://www.idunn.no/npt/2009/04/art06

Lyster, S.-A. H. & Horn, E. (2009). *Trog-2. Norsk versjon.* Manual, Pearson Assessment.

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, Second Edition (BPVS II). *Spesialpedagogikk 09*, 35-43.

Meyers, J. E., & Meyers, K. R., (1995). *Rey Complex Figure Test and Recognition Trial.* Odessa, Fl.:Psychological Assessment Resources, Inc.

Morton, J., & Frith, U. (1995). Causal modelling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds*.), Developmental psychopathology. Theory and methods, Vol. 1*. (pp. 357−390). New York: Wiley. Retrieved from http://www.icn.ucl.ac.uk/dev_group/ufrith/documents/Morton%20and%20Frith,%20Causal%20modeling%20%20A%20structural%20approach%20to%20developmental%-20psychopathology%20copy.pdf

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10(6*), 719–726. doi:10.1111/j.1467-7687.2007.00623.x

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66,* 232–58. doi:10.1016/j.cogpsych.2012.12.002

Paap, K.R., Johnson, H. A. & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined

circumstances. *Cortex, 69*, 265-278. Downloaded from https://doi.org/10.1016/j.cortex.2015.04.014

Paradis, J., Genesee, F. & Crago, M. (2011). *Dual language development & disorders: A handbook on bilingualism and second language learning (2nd ed.)*. Baltimore, MD: Paul H. Brookes.

Polit, D. F., & Beck, C. T. (2012). *Nursing Research: Generating and Assessing Evidence for Nursing Practice*. Philadelphia: Lippincott, Williams & Wilkins.

Rambøll Management (2016). *Evaluering av særskilt språkopplæring og innføringstilbu*d. Oslo. Retrieved from https://www.udir.no/globalassets/filer/tall-og-forskning/forskningsrapporter/evaluering-av-sarskilt-sprakopplaring-2016.pdf

Sandnes, T. (2017) red. *Innvandrere i Norge 2017*. Oslo/Kongsvinger: Statistisk sentralbyrå. Retrieved from https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/332154?_ts=SA155_web.pdf

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests (pp.341-363).* New York and Oxford: Oxford University Press

StatSoft, I. (2011) STATISTICA (data analyses software system) (Version 12). Retrieved from www.statsoft.com

Statistisk sentralbyrå (2017). *Befolkningens utdanningsnivå*. Retrieved from https://www.ssb.no/utniv/

Statistisk sentralbyrå (2017b). *Hvordan går det med innvandrere og deres barn i skolen?* Retrieved from https://www.ssb.no/utdanning/artikler-og-publikasjoner/hvordan-gar-det-med-innvandrere-og-deres-barn-i-skolen

Social Science Statistics. *Effect Size Calculator for T-test*. Retrieved from http://www.socscistatistics.com/effectsize/Default3.aspx

Torkildsen, J., Morken, F., Helland, W. A., & Helland, T. (2015). The dynamics of narrative writing in primary grade children: writing process factors predict story quality. [journal article]. *READING AND WRITING*, 1-26. doi:10.1007/s11145-015-9618-4

Ward, J. (2010). *The Students Guide to Cognitive Neuroscience*. New York: Psychology Press.

Wechsler, D. (1991). *The Wechsler intelligence scale for children – Third edition.* San Antonio: The Psychological Corporation.

Wechsler, D. (2003). *The Wechsler intelligence scale for children – Third edition. Norwegian version.* Stockholm: Assessio Norge AS.

Wilhelmsen, M., Holth, B.A., Kleven, Ø. & Risberg, T. (2013). Minoritetsspråk i Norge: En kartlegging av eksisterende datakilder og drøfting av ulike fremgangsmåter for statistikk om språk. Oslo/Kongsvinger: *Statistisk sentralbyrå*. Retrieved from http://www.ssb.no/utdanning/artikler-og-publikasjoner/_attachment/100940?_ts=13d3a8c3cf0

Øzerk, K. (2016). *Tospråklig oppvekst og læring.* Oslo: Cappelen Akademiske Forlag.

The interface between language abilities and cognitive abilities, how much does it contribute to the assessment of bilingual children. A study on Norwegian L2 8-year-olds.

By

Lilian Haugereid and Anne Dalland, Coldwell

Master's programme in health sciences – Logopedics

The Faculty of Psychology

Department of Biological and Medical Psychology

University of Bergen

Spring 2018

Sammendrag

Hovedformålet med denne studien var å kunne skille mellom språkvansker som skyldes en forbigående mellomspråklig vanske, og vansker som skyldes SLI og /eller dysleksi hos tospråklige 8 -åringer med Norsk som andrespråk. Utvalget i studien var 20 tospråklige 3.-klassinger fra en større norsk by. Vi hadde en kontrollgruppe på 42 enspråklige barn, fra en tidligere studie. Testbatteriet var basert på både språklige (L2) ferdigheter og nevrokognitive faktorer, i henhold til de tre nivåene (symptomatisk, kognitivt og biologisk) hentet fra den differensialdiagnostiske modellen til Morton og Frith. Vår hypotese var at tospråklige barn med språkvansker kan identifiseres av Morton og Frith sin modell for differensialdiagnostisering. Vi ønsket også å se om tospråklige barn har noen nevrokognitive fordeler i forhold til den enspråklige gruppen. Vår hypotese var at enspråklige vil ha bedre resultat på de språklige testene i forhold til L2-gruppen, men det var ikke forventet noen forskjeller på gruppene med hensyn til nevrokognitive evner.

Som ventet, viste funnene dårligere resultat på språktestene for L2 -gruppen i forhold til den enspråklige gruppen. De kognitive testene viste at de tospråklige hadde bedre resultat på de testene som omhandlet visuelle og auditive evner sammenlignet med L1-gruppen. Resultatene fra studien viser at det er mulig å bruke den differensialdiagnostiske modellen til å identifisere tospråklige barn som er i risikosone for å ha språkvansker og/eller dysleksi og å skille mellom dem og de som har en mellomspråklig vanske.

Nøkkelord: Tospråklighet, Mellomspråklig vanske, Språkvansker, Dysleksi, Differensialdiagnose, Nevrokognitive fordeler.

Abstract

The main aim of this study was to separate transient between language problems from problems caused by SLI and/or dyslexia among bilingual 8-year olds having Norwegian as their second language (L2). The sample in the study was 20 bilingual-3rd grade school children from a larger Norwegian city. We had a control group of 42 monolingual children, from a previous study. Assessment tools were based on both linguistic (L2) skills and neurocognitive factors, according to the three levels (symptomatic, cognitive and biological) of the differential diagnosis model by Morton and Frith. We hypothesized that bilingual children with language impairment can be identified by Morton and Frith's differential diagnostic model. We thus aimed to explore whether the differential diagnosing model, can serve as a method in defining language impairment in bilingual children. We also wished to see if bilingual participants exhibit some cognitive advantages over their monolingual counterparts. We hypothesized that monolinguals will outperform bilinguals on language abilities, but no group differences were expected with respect to neurocognitive abilities.

As predicted our findings showed poor bilingual performances in language tests compared to monolingual performances. The neurocognitive tests showed bilingual strengths in visual and auditory processing compared to the L1 group. Findings from our study clinically imply that the differential diagnostic model can be used to identify bilingual children at risk of having language impairment and/ or dyslexia, from children with between language problems.

Keywords: Bilingualism, Between language problems, Language impairment, Dyslexia, Differential diagnosis, Neurocognitive advantage.

Introduction

Language is a fundamental aspect of our social being and an important tool for academic achievements, as well as general life achievements. Learning outcomes in schools are strongly linked to a child's competence in the language of instruction in school (Lervåg & Aukrust, 2010). In Norway, Norwegian is the language of instruction in regular classes. However, besides Norwegian, which is the majority language, statistics from 2013 records over 300 other languages (Wilhelmsen, Holth, Kleven & Risberg, 2013). Statistics Norway's study on examination results and completion of secondary school by bilingual students report lower performances relative to monolingual students (those with Norwegian as their first language - L1) (SSB 2017b). Norwegian studies conducted on both kindergarten children and school children point out that children with minority language lag behind their monolingual counterparts with respect to vocabulary, reading skills, and other language aspects (Lervåg & Melby-Lervåg, 2009; Lervåg & Aukrust, 2010; Karlsen, Lyster & Lervåg, 2016).

The notion that bilingualism has some negative impact on children has long been proved to be wrong (see Paradis, Genesee and Crago, 2011; Øzerk, 2016; and Egeberg, 2016). However, there are varying empirical findings on whether bilingualism provides cognitive advantages or not. Studies such as Bialystok (2001), Bialystok and Martin (2004), Engel de Abreu, Santos, Tourinho, Martin and Bialystok (2012), and Hansen et al., (2016) report findings that support bilingual cognitive advantage. Results from these findings revealed bilingual advantages in executive function and working memory, among others. Meanwhile, studies by Morton and Harper, (2007); Paap and Greenberg, 2013; Paap, Johnson, and Sawi (2015) do not find supporting evidence for this notion. Friesen, Latman, Calvo, Bialystok (2015) report on evidence for better control of visual attention in bilinguals, however McVeigh, Wylie, and Mulhern (2017) did not find bilingual advantages in either visual pattern recall nor in working memory.

When the language(s) of a bilingual child is delayed or becomes a cause of worry to either parents, teachers or other caregivers, there is always the question of whether the delay is due to bilingualism or language impairment. One of the commonest language impairments among preschool children and school children is Specific Language Impairment - SLI (more recently referred to by Bishop et al. (2017) as Developmental Language Disorder (DLD). A child has SLI if the child has serious problems verbally expressing herself, and or problems comprehending what others verbally express to her, in the absence of other developmental disorders (Helland, 2012; Hulme & Snowling, 2009). SLI in sequential bilinguals is referred to by Ebert and Kohnert (2016, p. 301) as impairment in the "acquisition and use of first and subsequent languages due to some inefficiency in a child's internal language learning system." According to Bjerkan, Monsrud, and Thurmann-Moe (2013) SLI is often identified a little later in bilingual children than in monolingual children since the problem can easily be perceived as a weak development of the second language (L2), rather than symptoms of SLI.

There are concerns about the use of single language test assessment tools for testing or assessing language abilities on bilinguals. Speech and language pathologists (SLPs) risk over-identifying, under-identifying or miss-identifying developmental bilinguals when they use single language tests in assessing bilingual children (Kohnert, 2013, p. 146). De Jong (2008), referring to Gutiérrez-Clellen (1996), states that the reasons why there is skepticism towards the use of single language assessment tools in assessing for language ability among bilinguals, is that these tools are not able to identify those without disorders (lacks specificity); and they do not have the ability to identify those with disorders (lacks sensitivity).

Regarding language processing in bilingualism, Desmet and Duyck (2007) point out that most classical models on language processing are mainly for monolingualism, and that, theoretical models proposed for bilingual language processing involve all linguistic components; lexical representation, orthography, phonology, semantics, syntax, auditory word

processing, etcetera. In their investigation, they reviewed studies that investigated language processing of bilinguals. Desmet and Duyck (2007) write that at the lexical level, bilinguals have more than one lexical representation to express the same meaning. Paradis et. al., (2011, p. 67) also states that a majority of researchers agree that bilingual children have dual language system at birth. With regards to language perception in bilingualism, results from Burns, Yoshida, Hill, and Walker (2007) suggest that infants exposed to two languages have two separate perceptual systems for the two languages (stated in Paradis et al., 2011, p. 63). Furthermore, unlike lexical representations, semantic representations do not differ across languages (Desmet & Duyck, 2007).

Concerning phonological processes, a study on an English - Norwegian bilingual 2-year old (Johnson & Lancaster, 1998) examined phonological segments in the two languages the child was acquiring, and the results from the production and distribution of sounds showed that the child used separate sound inventories for his two languages (stated in Paradis et al., 2011, p. 64). To add to that, Desmet and Duyck (2007) referring to studies by Brysbaert, Van Dyck and Van De Poel (1999) report that "bilinguals are faster to recognize words from their second language if these words are primed by non-words that sound like that word (pseudohomophones) if they are pronounced as in the native language" (Demet & Duyck, 2007, p. 171). With respect to syntax, Demet & Duyck (2007) write about research that compare syntactic processing of bilinguals with syntactic processing of monolinguals: "the underlying idea is that if bilinguals process their first language differently from monolinguals of that language, then it must be that exposure to a second language influenced the processing of their native language." (Desmet & Duyck, 2007, p. 182). In conclusion, as discussed in chapter three, four and five of Paradis, et al. (2011), evidence from speech perception of bilingual children support the notion that they have two language systems, which includes phonology production, vocabulary building, and morphosyntax. They do, however, have the

ability to choose which language to use based on the context. Meanwhile, bilingual children are not "two monolinguals in one person", Paradis et al. (2011) warn. Speech and Language Pathologists (SLPs), parents, caregivers and those involved in the development and care of bilinguals can expect similar developmental milestones seen in monolinguals, in bilingual development as well (Paradis, et al., 2011, p. 67). Meanwhile, research shows poor L2 abilities in bilingual children compared with their monolingual counterparts (Gasquoine, 2016; Bialystok, Luk, Peets & Yang, 2010; Bialystok, Barac, Blaye & Poulin-Dubois, 2010)

Concerning clinical assessment of bilingual children, it is widely accepted that assessment should include the two languages the child is exposed to (De Jong 2008; Kohnert, 2013; Egeberg 2016). In practice, in a clinical setting, assessing the bilingual child in both languages can be challenging to the SLPs since they most often do not know the child's L1 (Kohnert, 2013). This observation is true for SLPs in Norway. However, Boerma and Blom (2017) write that due to many linguistic diversities, assessing bilingual children in both languages cannot be realized. "Time restrictions, insufficient financial resources, and the lack of (culturally) appropriate instruments, bilingual speech-language pathologists and skilled interpreters are just a number of obstacles to overcome" (Boerma & Blom, 2017, p. 66). This statement is true for Norway, where this current study was situated.

Some empirical studies that aim to find means of assessing bilingual children, study either the L1 and/or L2 of the subjects. Armon-Lotem and Meir (2016) used test scores of both languages as well as additional information from parental/teacher concerns; Boerma and Blom (2017) rely on L2 testing exclusively; Boerma, Chiat, Lesemana, Timmermeister, Wijnen, and Blom (2015) used the Language Impairment Testing in Multilingual Settings (LITMUS) Tools (described in Armon-Lotem et al., (2015) in their study of identifying bilingual children with either SLI or language interference. Thordardottir and Brandeker

(2013) report that non-word repetition task and sentence repetition task can distinguish between children with SLI and those without irrespective of bilingualism.

Norwegian empirical studies in bilingualism often concentrate on other aspects of bilingualism rather than clinical assessment issues on this target group. Recent studies like Karlsen, Lyster, and Lervåg (2016) look at vocabulary development of Urdu-Norwegian pre-schoolers. Karlsen, Lyster, and Geva (2016) also investigate narrative production in Urdu/Punjabi - Norwegian pre-schoolers. In their study, they use similar measuring tools as it is in our study (TROG 2 - Test for Reception of Grammar by Bishop, 2009, and BPVS 2 - British Picture Vocabulary Scale II by Dunn, Dunn, Whetton, & Burley, 1997). In their assessments, Karlsen, Lyster, and Geva (2016) considered the cognitive ability of their participants while comparing it to their L1 and L2 proficiency. In their suggestions to future research in the area, they noted that the development of appropriate L1 and L2 tests is a perennial challenge that researchers face in this field of study. Rodina (2016) is also a Norwegian study that investigates narrative development in Norwegian-Russian simultaneous bilingual pre-schoolers. It concludes that their ability to compose and understand narratives is equally developed in the two languages. In their longitudinal study Grøver, Lawrence, and Rydland (2018) also looked at the role of L1 in vocabulary development of bilinguals from pre-school to fifth grade. They found out that bilingual children with more developed L1 vocabulary skills who were exposed to teacher-led talk and peer-play talk with a high density of tokens had more developed L2 vocabulary at age five.

Meanwhile, Ryen and Simonsen (2015) refer to an assessment tool (Cross-linguistic Lexical Tasks - CLT) which has been tested on Polish-Norwegian preschool children (with typical development) in the age 3.5-6 years. These children have been compared to monolingual children of the same age in Poland and Norway. In Armon-Lotem, de Jong and Meir (2015) there are discussions on using narrative abilities of bilinguals, parental-

questionnaires, executive functions (among others) in assessing bilingual children. In our study, we take the revised causal model of Morton and Frith, done by Helland (2002) as the basis of our analysis. Helland (2012) discusses four levels (biological level, symptomatic level, cognitive level and environmental level) for investigating and understanding of SLI / dyslexia by Morton and Frith. Information concerning biology includes brain laterality (also in handedness) and other health information that occurred during pregnancy, during and after birth (including gender, hearing, sight, gene, heredity and so on) are shown by research to be important for language and thus relevant during assessment of SLI and dyslexia (Helland, 2012, p. 125). Information from the symptomatic level concerns behavioral information that can be observed or tested as symptoms of SLI/dyslexia (Helland, 2012, p. 123). At the cognitive level, neurocognitive and neurolinguistic information are investigated. The neurocognitive aspect investigates concerns on attention, short-term memory/working memory (where auditory and visual abilities are investigated) and executive functions. The neurolinguistic level looks at the language processing abilities (where both language comprehension and language production are investigated), phonological awareness and rapid naming abilities. According to Helland (2012), environmental information includes information about the type of language a child is exposed to at home, school and other influences from his/her surroundings.

We also based our analysis on theories from the working memory by Baddeley and Hitch (1974) (see also Baddeley, (2003) and Baddeley, (2010). According to the theory, short-term memory is the "temporary storage of small amounts of material over brief periods of time" (Baddeley, 2010). Findings prove that children with SLI have poor short-term memory (Baddeley, 2010). Ebert and Kohnert (2016) write that although language is the main difficulty for children with SLI, it is not the only difficulty they face. They explain this by stating that SLI also comes with “inefficiencies in processing non-linguistic information”

which affect processing speed, working memory, and attention (Ebert & Kohnert, 2016). Ebert and Kohnert, 2016, supported this claim by referring to empirical studies with findings that support this notion; children with SLI demonstrate deficit in working memory skills which is beyond "verbal stimuli to non-linguistic targets such as spatial locations "(Vugs et al. 2013); they demonstrate poorer sustained selective attention abilities, for non-linguistic stimuli (Ebert & Kohnert 2011), (mentioned in Ebert & Kohnert, 2016). Helland (2012, p. 72) also points out that research points to the fact that children with SLI have limited capacity to process linguistic information, and that they have specific difficulties associated with phonological memory and processing; and they also have difficulty with auditory discrimination. These difficulties can be analyzed within Baddeley and Hitch's working memory model (Helland, 2012, p. 72). The causal model by Morton and Frith can help to make a differential diagnosis as a core in the assessment work on persons with language impairment (Helland, 2012, p. 124). With reference to how it is illustrated by Helland (2002), we consider both the neurocognition and neurolinguistic aspects in our analysis. Within the area of neurolinguistics, the comprehension, processing, and production of language are aspects that are important when diagnosing language impairment.

Current Study

The main aim of our study was to investigate the separation between language problems associated with SLI and/or dyslexia, and those due to language interference in bilingual children. Based on our discussions above, we hypothesized that bilingual children with language impairment can be identified by Morton and Frith's differential diagnostic model. We thus aimed to explore whether the *differential diagnosing* model can serve as a method in defining language impairment in bilingual children. We also wished to see if bilingual participants exhibit some cognitive advantages over their monolingual counterparts, and if they do, what explanations/contributions these cognitive abilities provide for the

investigation on language-interference and language-impairment divide in bilingualism. We also aimed to investigate the linguistic and cognitive abilities of bilingual children while comparing them to monolingual children. Research has shown that monolingual children will outperform their bilingual counterparts with respect to language abilities. We, therefore, hypothesized that monolinguals will outperform bilinguals on language abilities, but no group differences were expected with respect to cognitive abilities. Since our main aim was to investigate the separation between language interference and language impairment, we aimed to analyze subjects in the L2 group individually to compare their language and cognitive profiles to the profile of a typical SLI and/or dyslectic child as discussed above.

Methods

This study is part of the larger project "Norwegian as a second language (L2) in 3rd-grade school children. How can transient between-language problems be separated from problems due to more specific language impairment and/or dyslexia?" The project is managed by associate professor Wenche Andersen Helland, associate professor Frøydis Morken and Professor emerita Turid Helland, all from the Department of Biological and Medical Psychology, University of Bergen. Control data are from the study "Tell a tale - a study of the language of 3rd graders" conducted by the Bergen Logopedic Research group (B.LOG), (Torkildsen, Morken, Helland & Helland, 2015). In our study, we used the same types of test batteries and parental questionnaire that were used in Torkildsen, et.al., (2015).

The Norwegian Centre for Research Data (NSD) gave approval for the project in August 2017. Declaration of consent to participate in the study was signed by the guardians of the children.

Participants

Invitation to take part in the project was sent to schools in a larger Norwegian city. The project managers had meetings with the school managements informing about the project,

and participants were recruited from four of these schools. Written information about the project, declaration of consent, and a questionnaire on background information of the children, was handed out to parents by the schools. The native language teachers at the schools were asked to translate the written information into the native languages of the parents if needed. In retrospect we found out that this was not done, making it impossible to fully rely on information given in the questionnaire.

Inclusion criteria for participating was having attended Norwegian school since 1st grade, no identified mental retardation, and no identified neurological disorder. When planning the project, we wanted children participating that teachers and/or parents had concerns about regarding the progress of linguistic and literacy learning. Having this as a criterion turned out to be a problem getting enough participants to the study, and the project managers decided that having Norwegian as L2 was sufficient to participate, not knowing which of the children teachers/parents were concerned about. The sample we got was twenty 3rd. graders, with Norwegian as L2, as participants in the study. There were 12 boys and 8 girls, two of the children left-handed. The participants came from 13 countries, from four continents and spoke 10 languages (L1). We received information on the number of years of education from a total of 17 mothers, (*N*= 17; *Mean* = 12.71, SD = 4.93); and a total of 16 fathers' (*N*= 16; *Mean*= 13.53, SD = 3,78) for the bilingual group.

The control group consisted of 42 3rd graders, 16 girls and 26 boys, 11 of the children left-handed. The inclusion criteria were the same as in the current study. They all had Norwegian (bokmål) as L1, and their parents had an education level close to the national average. Approximately 46 % of participating mothers and 49 % of participating fathers had higher education (minimum one year of college or university studies) compared to 54 % of women and 39 % of men in the relevant age group of the Norwegian population (Statistics Norway, 2013) according to Torkildsen et.al., (2015).

Procedure

Data collection took place in November and December 2017 and was carried out by four master students in logopedics who were trained in administering the different tests.

All testing was done at school, during school hours, and it took about two hours per child to complete the assessment. The children were alone with the test leaders during testing and the test leaders tried to make the environment for the assessment as positive as possible, making sure that they had breaks when needed, and had access to drinks and biscuits to keep the motivation up.

The test battery consisted of 11 different tasks that took between five and twenty minutes to complete. Test administration was shared among the test leaders. Each test leader was responsible for the administering of every second test. When not in charge of administering, the test leader's responsibility was to observe and help with practical things like setting up the time and recording.

The four master students scored the different tests according to the manuals, under the guidance of our supervisor, and the results were collated in a common data matrix in the statistical data programme Statistica (StatSoft, 2011), and is the basis for the analysis of the collected data.

Assessment tools

For background information, we tentatively used the questionnaire. For individual testing, we used 11 different tests on the subjects' basic linguistic and literacy skills in L2 (Norwegian) and neurocognitive factors. We have focused on language, linguistic skills, and relevant cognitive factors.

The questionnaire used is an age-modified version of the questionnaire Risk Index-5 (RI-5) (Helland, 2015), developed to identify preschool risk of dyslexia (Helland, Plante, & Hugdahl, 2011). The original RI-5 form is to be answered by parents and/or teachers and

gives information on the following subjects: general health, allergies, asthma, left-handedness, motor-skills, language, special needs and heredity (Helland et al., 2011). In this study, it was only parents who answered the questionnaire. The questionnaire was answered and returned by 19 of the 20 parents. As mentioned before we assumed that parents would be given help translating the questionnaire to their native language. The results showed that only a few received this help, and the project leaders concluded that the overall quality of the questionnaires was not good enough to draw any useful conclusions.

British Picture Vocabulary Scale (BPVS II) (Dunn et al., 1997) is a standardized test to map out children's receptive vocabulary and detect delays in vocabulary development. The Norwegian version of BPVS II is translated and adapted to Norwegian from the second English version (Lyster, Horn & Rygvold, 2010). The test consists of 12 different sets, with 12 pictures in each set. Each set has an age span, starting at 2,5 -3 years up to 16-21 years. The test is started on the set equivalent to the child's age. The test leader says a word and the child points at the picture matching the word. There is no need for spoken language by the child. Each correct answer gives one point, and the maximum score is 160 points. Normed score for the age group 7:0-7:11 is 87.32 and 95:52 for the age group 8:0-8:11(Lyster et. al. 2010).

Test for Reception of Grammar, version 2 (TROG-2) (Bishop, 2003) is translated and adapted to Norwegian for the age group 4-16 years (Horn & Lyster, 2009). TROG-2 is a receptive language test that assesses understanding of grammatical contrasts marked by inflections, function words and word order. TROG- 2 makes it possible to compare grammatical comprehension with peers of the same age, and to find areas of difficulty (Bishop, 2003). The test consists of 80 four-choice items. A simple vocabulary of nouns, verbs, and adjectives is used. For each grammatical contrast, there is a block of four items. The block is passed if all items are answered correctly. The test is ended when five

consecutive blocks have been failed. The test leader reads a sentence and the child points to the correct picture, no spoken language needed. Normed score for the age group 7:0-7:11 is 14.86, and 15.93 for the age group 8:0-9:11 (Horn & Lyster, 2009).

The model sentence task based on "Ringstedmaterialet" (Ege, 1984) is used to measure comprehension and the expressive language in children and can give information about phonology, morphology, syntax, and semantics. The test is challenging when it comes to skills in attention, memory, maturity and language awareness (Helland, 2012, p. 51). In this study, we did not focus on phonology, because of the age of the children, and the fact that they had attended Norwegian school from first grade. We did not hear any deviant pronunciation, except foreign accent. Phonological difficulties would most likely be a between-language problem. Phonology was also excluded in the control group. The model sentence task consists of 20 model sentences with corresponding pictures. The test administrator displays a picture to the child and says the model sentence; the child gets a different but analogue picture and is going to construct a corresponding sentence. The sentences get more complex and have increasing difficulty throughout the test. Giving one point for correct use of morphology, syntax and semantics scores the sentences, with a total of three points possible for each sentence. This test has no normed scores. We have used a revised version of the material developed for the longitudinal study "Speak up!" (Helland et al., 2011).

Rapid naming (RAN) from the Stroop battery (Golden, 1987, Hugdahl, undated version) is a test that measures verbal processing speed given visual stimuli. The test we used in this study was a part of the Stroop-test called "Colour naming". The test consists of a form with 48 dots in different colors. The child shall name the colors as quickly as possible. The test leader is timing and takes notes of errors and corrections. We scored the tests after

standard procedures (Hugdahl, undated version). This test has no normed scores, only the time scores in seconds used.

We used the test digit span from the Wechsler test battery (Wechsler, 1991, 2003) to test verbal short-term and working memory. The test consists of 15 sets of two tasks in each set. Eight sets are numbers said out by the test leader with a second interval, and to be repeated forward and seven sets are numbers repeated backward with increasing difficulty. The test leader note errors and the testing is discontinued when both tasks in one set have failed. The test was scored after standardized procedures (Wechsler 1991, 2003).

The Dichotic listening (DL) test (Bless, Westerhausen, Kompus, Gudmundsen & Hugdahl, 2014) measures how language sounds are perceived in the brain. We used the app iDichotic, developed by researchers and scientists affiliated with the Bergen fMRI group. The children used an iPad/ smartphone with headphones, doing the test. The test contains six pairs of CV syllables presented simultaneously on both ears, forming 30 different combinations. The first part is the Non-Forced (NF), where the children tap on the image of the sound they hear best, the second part is Forced Left, and Forced Right (FR/FL) where they focus attention on and report on the syllables heard at one ear at the time. The syllables used are presented 30 times in each part of the test. The results were scored in numbers of correct answers on both ears both Non-Forced and Forced.

Visuospatial skills. Rey-Osterrieth Complex Figure Test (Meyers & Meyers, 1995). RCFT is a standardized test showing a complex picture containing 18 different figures. The test is divided in two, RO-copy, where they are copying the picture on a piece of paper, and RO-recall, where they are drawing the picture recalling it from their memory 25 minutes after they did the RO-copy. Both tasks must be done in five minutes. The test was administered and scored according to the manual, with a maximum score of 36 for the age group.

Statistical analyses

The data in our analysis were analyzed in accordance with the properties of the collected data and our research questions discussed above. To compare significant differences between the means of bilinguals and monolinguals, a t-test for independent samples was used, with the design Test scores by Group (2: L1, L2). To assess possible variances in speech production (Model sentences) a repeated measure (ANOVA) was used with the design Tasks (3: morphology, syntax, and semantics) by Group (2: L1, L2). LSD follow-up test was used to assess significant effects. To assess the relationship between the language scores and the cognitive scores one-list correlational analysis was used. An alpha level of *.05* was used for statistical significance. To find out the magnitude, or size of the effect, Cohen's *d* effect size was calculated. The values for small, moderate, and large Cohen's *d* are 0.2*, 0.5,* and *0.8* respectively. Finally, a case by case analyses is done for the L2 group to find out those with scores over and above *1 SD* for all variables.

Results

Correlation Analysis

Correlation analyses between all the test batteries used in this study (summary on Table 1) revealed a strong positive correlation between receptive vocabulary (BPVS) and receptive grammar (TROG), and also between BPVS and Model sentences. Meanwhile, a significantly moderate correlation was shown between TROG and model sentences. A very strong positive correlation was revealed between RO1 and RO2; and a significantly weakly negative relationship was recorded between dichotic listening test, non-forced right (DL NF RE) and the two language test batteries, BPVS and TROG. Significantly weak negative correlation between RAN and digit span, backward was revealed. Dichotic listening test, non-forced left, and non-forced right had significantly weakly positive correlations. Significantly weakly positive correlations were revealed between digit span, forward and backward and

also with the language test batteries, BPVS, TROG and Model sentences. Likewise, significantly weakly negative correlations were shown between RAN and BPVS, TROG and Model sentences. Details are shown in Table 1 below.

Table 1

Correlations Between Variables. Marked correlations are significant at p < .050. N = 62.

Variable	BPVS	TROG	Mod SUM	Digit span for-ward	Digit span back-ward	RAN	RO1	RO2	DL NF Re	DL NF Le
BPVS										
TROG	.64***									
Mod SUM	.72***	.60***								
Digit span forward	.31*	.30*	.32*							
Digit span backward	.28*	.32*	.36**	.25*						
RAN	-.31*	-.21	-.35**	-.24	-.27*					
RO1	-.11	-.30*	-.05	.04	-.01	.01				
RO2	-.20	-.25	-.05	.00	.11	.01	.80***			
DL NF Re	-.30*	-.28*	-.20	-.09	.19	.07	.11	.21		
DL NF Le	.12	.06	.12	.06	.06	-.12	.01	-.10	-.25*	

Note: SD = Standard deviation; *p< .05* p< .01** p<.001****; BPVS = British Picture Vocabulary Scale II; TROG = Test of Receptive Grammar; Mod Sum = speech production (Model sentences); RAN = Rapid naming; RO1 and RO2: Rey-Osterrieth Complex Figure Test (copy and recall); DL NF Re = Dichotic listening non-forced right; DL NF Le = Dichotic listening non-forced left.

Language variables

As can be seen in Table 2 The monolingual group scored significantly higher in all language test batteries, compared to their bilingual counterparts (see Table 2). An independent-sample t-test indicated that scores for receptive vocabulary (BPVS) were significantly higher for the control group than for the bilingual group. With regards to receptive grammar (TROG), the analysis revealed that subjects scored significantly higher in the L1 group than in the L2 group. Likewise, significant group differences were revealed in subjects' ability to make model sentences. The L1 group scored significantly higher than did the L2 group.

Table 2

Independent sample T-tests between monolinguals (control group) and bilinguals (L2 group) for language variables

Variable	Mean (L2) N= 20	*SD* (L2)	Mean (Control) N= 42	*SD* (Control)	t-value	p	Cohen`s d
BPVS	64.55	15.85	92.02	13.90	-6.952	0.001	1.84
TROG	78.70	13.27	102.05	12.66	-6.683	0.001	1.80
Mod SUM	39.70	9.18	49.00	5.84	-4.841	0.001	1.20

Note: SD = Standard deviation*; p< .05* p< .01** p<.001***;* BPVS = British Picture Vocabulary Scale *II; TROG* = Test of Receptive Grammar; Mod Sum = speech production (Model sentences)

Cognitive Variables

Independent t-test analysis of cognitive variables revealed varying results. Statistically significant group differences are revealed on the variables; RAN (rapid naming), RO1 and RO2 and dichotic listening test, non-forced right ear (DL NF Re) (see Table 3). Monolingual participants used statistically significant lesser time span on RAN than did their bilingual counterparts However, bilingual participants scored statistically higher on RO1 and RO2 than did the monolingual participants. The DL test revealed that participants in the L2 group had significantly right ear scores over their monolingual counterparts ($p < .003, d = 0.85$). Non-significant group differences were revealed in the analysis of digit span (forward and backward). Other statistical details are presented in Table 3.

Table 3

Independent Sample T-tests between monolinguals (control group) and bilinguals (L2 group) for Cognitive Variables

Variable	Mean (L2) N=20	*SD* (L2)	Mean (control) N=42	SD (control)	t-value	p	Cohen`s d
Digit span forward	6.30	1,53	7.10	1.85	-1.672	0.100	0.47
Digit span backwards	3.45	1,28	3.95	1.34	-1.399	0.167	0.38
RAN, sec	59.53	17,12	48.76	16.13	2.409	0.019	0.64
RO1	22.48	8,26	16.33	4.62	3.756	0.001	0.91
RO2	14.93	7.69	7.58	4.12	4.663	0.001	1.19
DL NF Re	14.55	3.47	11.60	3.46	3.137	0.003	0.85
DL NF Le	9.35	3.12	10.48	3.08	-1.341	0.185	0.36
DL FR Re	15.40	4.17	14.56	3.24	0.849	0.399	0.22
DL FR Le	9.20	3.38	9.00	2.15	0.277	0.783	0.07
DL FL Re	13.84	4.65	13.41	3.66	0.387	0.701	0.10
DL FL Le	8.53	2.89	9.87	2.98	-1.627	0.109	0.45

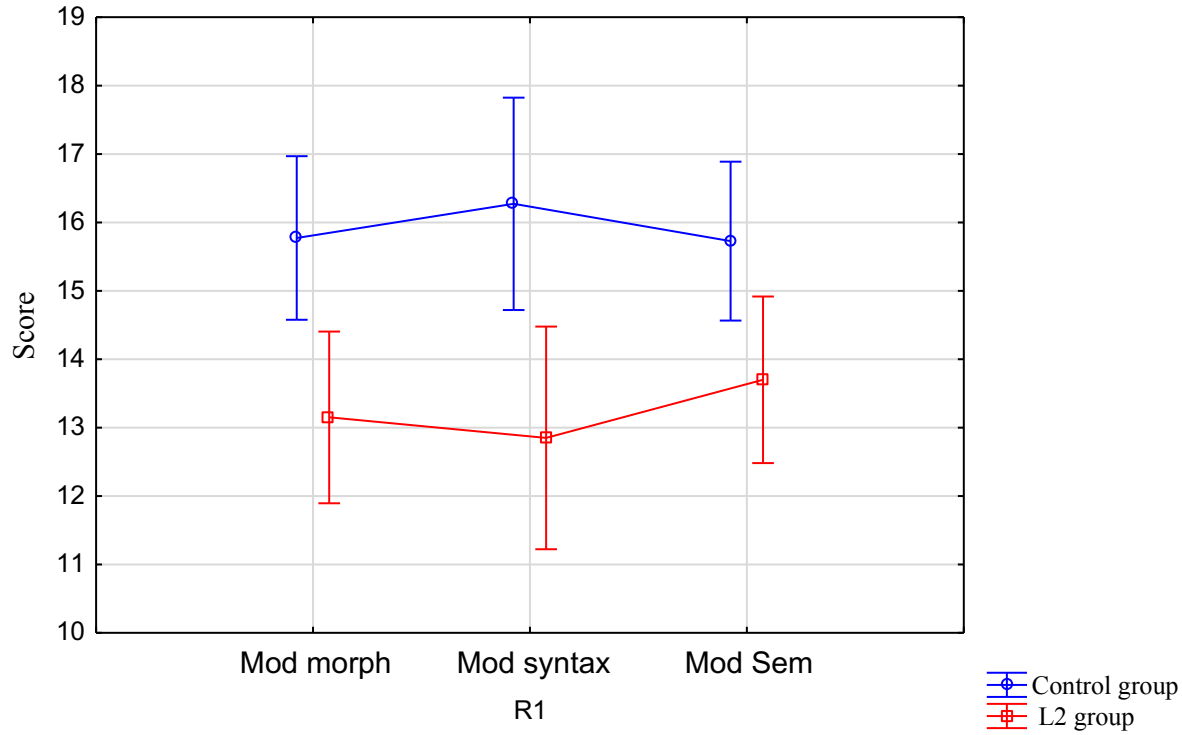
Note: SD = Standard deviation; *; p< .05* p< .01** p<.001***;* RAN = Rapid naming; RO1 and RO2: Rey-Osterrieth Complex Figure Test (copy and recall); DL NF Re = Dichotic listening non-forced right; DL NF Le = Dichotic listening non-forced left; DL FR Re = Dichotic listening forced right; DL FR Le = Dichotic listening forced-right left ear; DL FL Re = Dichotic listening forced-left right ear; DL FL Le = Dichotic listening forced-left left ear

ANOVA Results

The analysis showed no interaction effects on repeated measurement for morphology, syntax and semantic (See Figure 1). However, it revealed a significant effect by group: *F(1, 60) = 23.43, p <.001*. An LSD post hoc test showed that the group effect was due to higher scores in the control group compared to the L2 group ($p < .0001$). These effects were further assessed by t-test which revealed statistically significant group differences for the three linguistic components of syntax, morphology, and semantics ($p < .001$). Statistical details are shown in Figure 1.

Figure 1

Analysis of Variance (ANOVA) for Model sentences



Case study on L2 group

A case by case profile on subjects in the L2 group was performed. For each variable, we analyzed each subject's score that deviated by *1 SD* or more to mean of the control group. We found that four of the children were *1 SD* or more under mean on four or more of the tests, both linguistic and neurocognitive. Three of the children were *1 SD* or more above mean on four of the tests, both linguistic and neurocognitive. The profiles are shown in Table 4.

Table 4

*Case by Case Study on L2 Group for all Variable. X >1 SD Below Mean; * > 1 SD Above Mean of the Control group*

No.	ID	BPVS	TROG	MOD SUM	Digit span forward	Digit span back-ward	RAN	RO1	RO2	DL NF Re	DL NF Le
1	101	X	X	X	X		X				
2	102	X						*	*		
3	103				X			*	*		
4	104	X	X	X			X	*	*		
5	105							*	*	X	*
6	106	X	X	X		X		X			X
7	107		X	X						*	
8	201	X	X	X				*	*		
9	202	X	X		X					*	X
10	203	X	X	X	X	X	X	*	*	*	
11	204	X	X	X		X		*			
12	205		X		X		X	*	*		
13	206	X	X		X	*		*	*		
14	301	X	X	X	*		X	*	*	*	X
15	302	X	X	X		X	X				
16	303		X		X						
17	401	X	X	X	X	*		*	*	*	
18	402	X	X	X							
19	403	X	X				X	*			*
20	404		X			*		*	*	*	X

*Note: X = below SD; * = above SD; empty slot = within normal*

Discussions

The main aim of this study was to separate transient between language problems from problems caused by SLI and/or dyslexia among bilingual 8-year olds. We thus hypothesized that the differential diagnosis model can help identify and diagnose bilinguals with either SLI and/or dyslexia. The study also aimed to investigate the linguistic and cognitive abilities of monolinguals and bilinguals while checking if the L2 group exhibited some cognitive advantages. We therefore hypothesized that the L1 group would outperform the L2 group with respect to language tests, but that scores from cognitive tests would be about the same for subjects in both groups. In order to find out if subjects in the bilingual group have either language interference or language impairment and/or dyslexia, we aimed to investigate the

language and cognitive profiles of subjects in the L2 group, individually; We expected to find cases that could be regarded as instances of language interference as well as those that would raise suspicions as being either SLI and/dyslexia. We finally aimed to explore the cognitive tests scores that may indicate bilingual advantages.

As expected, our findings are in line with studies (Gasquoine, 2016; Bialystok, Luk, Peets & Yang, 2010; Bialystok, Barac, Blaye & Poulin-Dubois, 2010) that show poor bilingual performances in language tests relative to monolingual performances. Our results reflect and support theoretical claims and empirical findings that bilinguals have a smaller vocabulary than do monolinguals. However, we want to state here that, as pointed out by Kohnert (2013), Paradis et al. (2011), Egeberg (2016) and other writers on bilingualism, bilinguals have a shared vocabulary across the languages they speak. Thus, lack of vocabulary for certain concepts in L2 does not necessarily mean lack of vocabulary for those concepts in L1. Nevertheless, if the bilingual child has a very small vocabulary in the language of instruction in the school, which is Norwegian in the case of the participants in our study, then study outcomes will be directly affected, as reported by Lervåg and Aukrust (2010). This analogy is true for bilinguals' receptive grammar abilities (measured by TROG 2) and their speech abilities (measured by Model sentences). The low scores in these two tests also reflect their lack of vocabulary in the L2, as shown by low scores in BPVS.

Our findings from cognitive tests for bilingual subjects do not show clear-cut bilingual advantages as reported by Bialystok (2001), Bialystok and Martin (2004) and Hansen et al., (2016). However, our findings from the tests RO1 and RO2, that test visuospatial skills and visuospatial construction (Helland, 2012, p. 135), revealed bilingual advantages. The results from RO indicate that the L2 group exhibits better visual perception, organization, and memory which is similar to findings by Friesen et al., (2015) but contrary to findings by McVeigh et al., (2017). Considering the differences in sample sizes for the two groups (L1

group, N = 42, L2, N = 20) and the high mean score differences, and the fact that the language of instruction under testing was Norwegian (L2), these could serve as proof that our bilingual participants exhibit advantages in visuospatial skills and visuospatial construction over the monolingual participants. We, therefore, speculate that because the language skills in bilinguals are impaired, they compensate them with visual skills and therefore pay more attention to visual cues, and as a result, they have more sharpened visual short-term memory.

Findings from RAN showed that the L1 group used significantly less processing time than did the L2 group. The process of rapidly naming is composed of skills such as attention, memory, phonology, semantics, and motor skills (Helland, 2012, p. 108). Our findings then imply that the monolingual subjects were better at this linguistic and cognitive skills. According to Helland (2012, p. 108), difficulties with rapid naming can be explained based on the fact that there are deviations in the magnocellular system that results in slower processing of visual stimuli. As discussed above regarding findings from RO1 and RO2, the L2 group exhibited stronger visual skills than the L1 group. In spite of that, they used longer processing time on the RAN test. This result can be discussed on the basis that since RAN tests attention, perception, concept formation, memory, phonology, semantics, and motor-skills (Helland, 2012), then the test requires both linguistic and cognitive skills. Moreover, the test-language was Norwegian (L2 for the bilingual participants), the low scores in this particular test could be attributed to the language of the test. So, we speculate here that the monolingual group had an automatic advantage over their bilingual counterparts with respect to the fact that the language of instruction under testing and the language for the test itself, was Norwegian. We conclude therefore that the low RAN scores were not caused by attention deficits.

Another interesting result from our analyses concerns the dichotic listening test. The results for non-forced right showed a higher right ear score for the L2 group. Even though there were no significant group differences from the remaining DL tests, we observed from t-

test results that the L2 group had higher scores for all DL tests except DL forced-left, left ear (DL FL Le). However, the statistically non-significant scores showed that there were no group differences with respect to DL forced right or forced left condition which are measures of selective attention (Torkildsen et al., 2015). One may speculate that the relatively high scores on DL NF Re are a result of L2-children's need to pay special attention to what they hear, i.e. a result of efforts to perceive and understand. Similarly, visual attentiveness as seen in the RO-scores can be seen as they paying special attention to visual cues. In both cases, one can interpret them as their means of compensating for their linguistic impairment.

The main objective of this study was to separate transient between-language problems from problems caused by SLI and/or dyslexia among bilingual children. We hypothesized that the differential diagnostic method by Morton and Frith (revised by Helland, 2002) can identify bilingual children within the risk group of having SLI and/or dyslexia. The model suggests the three levels of assessment; symptomatic, cognitive and biological. At the symptomatic level, we observed during testing that subjects exhibited varying language abilities (in L2). We also observed that they had some language difficulties, but it was difficult to determine whether these difficulties were due to language impairment or due to language interference. Besides foreign accents, we did not observe any phonological difficulties from the verbal inputs from participants. The case by case analyses (see Table 4) conducted from results from each participant in the L2 group concentrated on the neurocognitive and neurolinguistic aspects of the differential diagnostic model. Findings from the case study showed that four of the participants scored *1 SD* below the mean of the control group on 4 or more of the test batteries, both language, and cognitive tests. They were subject`s number 101, 106, 204 and 302 (details on Table 4). As mentioned earlier, we cannot make conclusions about their language abilities solely based on results from single language tests (which is Norwegian), however, we can base our assessment on the underlying factors

from the cognitive test scores (as Frith and Morton's model suggests). A low score on Digit span, forward and backward suggests poor STM and WM which is common for SLI children according to empirical evidence (Baddeley, 2010); poor performance on RAN reflects lack of linguistic and cognitive elements, such as attention, perception, concept formation, memory, phonology, semantics, and motor-skills that is found to be impaired in children with SLI/dyslexia. So, if these four bilingual participants have low scores in these cognitive tests and corresponding low scores in language tests, we can argue here that they are identified by the model as having "symptoms" of SLI/dyslexia; or as being in the risk zone of developing SLI/dyslexia. Further clinical assessments for language impairment for these four subjects can then be suggested. We will suggest that the remaining aspect of the differential diagnostic model, that is, a more reliable biological and environmental information (through for example the RI-8 filled by both parents and teachers) should be collected on these four subjects. These will provide important information on the divide between SLI/or dyslexia and language interference; and will give a more concise SLI/or dyslexia diagnosis as suggested by Morton and Frith's model. We also suggest an L1-based non-word repetition task and sentence repetition tasks (Thordardottir & Brandeker, 2013) or a quasi-universal repetition task (Boerma & Blom, 2017) which are reported to identify bilingual children with SLI. We also recommend the Children's Communication Checklist - Second edition, Norwegian version (CCC-2) by Helland and Hollund-Møllerhaug, 2012).

We would like to further discuss here that 3 of the subjects in the L2 group scored *1SD* above the mean of the control group even though their scores in the language tests were mostly below *1SD* of the mean of their monolingual counterparts. They were subject numbers 301 (Digit span-forward, R01, R02 and DL NF Re), 401 and 404 (Digit span-backward, R01, R02 and DL NF Re). We therefore speculate here that they may not have language impairment and/ or dyslexia, rather language interference. Meanwhile, two of the subjects L2

group (103 & 105) scored within normal range on language tests compared to the mean of the L1 group. An explanation for this result may be that the typology of their L1 may not be too distant from Norwegian (L2). Subjects 103 and 105 have, respectively, Slovak and English as their L1 (which are both European languages).

Limitations of the Study

One of the main limitations of this study is the small sample size. A larger sample size could yield more significant relationships from the data, and could be more representative of the population. Another issue one may consider as a limitation of this study would be the fact that our participants have diverse level of proficiency in their L1 and L2. This could have some effects on the outcome of our findings. However, there are issues that poses recruitment problems when conducting studies with bilingual children. These include the vulnerability of the study participants, especially the L2 group (children with minority background); the complexity and diversity of the bilingual child's language background; language and interpretation issues with parents, and so on. These issues could pose recruitment problems and make it difficult to create strict inclusion criteria for a study such as ours.

Clinical Implications

Findings from our study clinically imply that the differential diagnostic model can be used to identify bilingual children suspected of having language impairment and/ or dyslexia. The findings that showed bilingual strengths in visual and auditory processing provide clinicians, teachers, caregivers and other professions involved in the bilingual child's educational planning, daily life, and teaching, important information about their needs, which can promote the use of visual and auditory aids in teaching.

Further Research

Further research is needed to investigate the bilingual child's L1 and L2 with the differential diagnosis method. Such study can decide whether IQ-testing should be included as one of the assessment tools. A longitudinal study, with a larger sample size, using the similar approaches as in this current study will provide insights on changes over time in linguistic and cognitive abilities of bilingual children compared with monolingual children.

References

Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.). (2015). *Assessing multilingual children: disentangling bilingualism from language impairment.* Retrieved from https://ebookcentral-proquest-com.pva.uib.no

Armon-Lotem, S., and Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew. *Int. J. Lang. Commun. Disord. 51*, 715–731. doi:10.1111/1460-6984.12242

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208. http://dx.doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–40. doi:10.1016/j.cub.2009.12.014

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bialystok, E. (2001). *Bilingualism in development: language, literacy, and cognition.* Retrieved from https://ebookcentral-proquest-com.pva.uib.no

Bialystok, E., Barac, Blaye, A. & Poulin-Dubois, D. (2010). Word Mapping and Executive Functioning in Young Monolingual and Bilingual Children, *Journal Of Cognition And Development, 11:4*, 485-508, DOI:10.1080/15248372.2010.516420

Bialystok, E. and Martin, M. M. (2004), Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science, 7:* 325-339. doi:10.1111/j.1467-7687.2004.00351.x

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., ... House, A. (2017), Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry, 58*: 1068-1080. doi:10.1111/jcpp.12721

Bishop, D. V. M. (2003). *The Test for Reception of Grammar, version 2 (TROG-2).* London: Pearson Assessments.

Bjerkan, K.M., Monsrud, M.-B., Thurmann-Moe, A.-C. (2013). *Ordforråd hos flerspråklige barn.* - Pedagogiske og spesialpedagogiske utfordringer. Oslo: Gyldendal Norsk forlag.

Bless, J. J., Westerhausen, R., Kompus, K., Gudmundsen, M. & Hugdahl, K. (2014). Self-supervised, mobile-application based cognitive training of auditory attention: A behavioral and fMRI evaluation. *Internet Interventions, 1*(3), 102-110. doi:http://dx.doi.org/10.1016/j.invent.2014.06.001

Boerma, T., and Blom, E. (2017). Assessment of bilingual children: what if testing both languages is not possible? *Journal of Communication Disorders. 66*, 65–76. doi:10.1016/j.jcomdis.2017.04.001

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., and Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *JSLHR 58*, 1747–1760. doi:10.1044/2015_JSLHR-L-15-0058

Cosby, P. C., & Bates, S. C. (2012). *Methods in behavioral research* (11.utg.). International edition: McGraw-Hill.

de Jong, Jan. (2008). Bilingualism and language impairment. In Martin Ball, Michael R. Perkin, Nicole Müller & Sara Howard (eds.), *Handbook of Clinical Linguistics, 261-*

274. Oxford: Blackwell Publishing. Retrieved from ProQuest eBook Central database.

Desmet, T., & Duyck, W. (2007). Bilingual language processing. *Linguistics and Language Compass, 1 (3)*, 444-458. Downloaded on 17.05.2018 from http://users.ugent.be/~wduyck/articles/DesmetDuyckInPress.pdf

Dunn, L. M., Dunn, D. M., Whetton, C. & Berley, J. (1997). *The British Picture Vocabulary Scale II*. Windsor: GL Assessment.

Ebert, K. D., and Kohnert, K. (2016). Language learning impairment in sequential bilingual children. *Language Teachin*g. 49, 301–338. doi: 10.1017/S0261444816000070

Ege, B. (1984). *Ringstedmaterialet. En procedure til undersøkelse af børns realiserbare sprog*. Herning: Special-pedagogisk forlag A/S.

Egeberg, E., (2016). *Minoritetsspråk og flerspråklighet. En håndbok i utredning og vurdering.* Oslo, Cappelen Damm Akademisk.

Engel de Abreu P. M. E., Cruz-Santos A., Tourinho C. J., Martin R., Bialystok E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychological Science, 23* 1364–1371. DOI: 10.1177/0956797612443836

Friesen, D. C., Latman, V., Calvo, A., & Bialystok, E. (2015). Attention during visual search: The benefit of bilingualism. *International Journal of Bilingualism 19*, 693-70. doi: 10.1177/1367006914534331

Gasquoine, P. G. (2016). Effects of bilingualism on vocabulary, executive functions, age of dementia onset, and regional brain structure. *Neuropsychology*, *30(8),* 988-997. http://dx.doi.org/10.1037/neu0000294

Golden, C. J. (1987). Manual for the Stroop Color and Word Test. Chicago: Stoelting Co.

Grøver, V., Lawrence, J., & Rydland, V. (2018). Bilingual Preschool Children's Second-Language Vocabulary Development: The Role of First-language Vocabulary Skills and Second-language Talk Input. *International Journal of Bilingualism*. ISSN 1367-0069. doi: 10.1177/1367006916666389

Hansen, L. B., Macizo, P., Duñabeitia, J. A., Saldaña, D., Carreiras, M., Fuentes, L. J. and Bajo, M. T. (2016), Emergent Bilingualism and Working Memory Development in School Aged Children. In *Language Learning, 66*: 51–75. doi:10.1111/lang.12170

Helland, T. (2012). *Språk og dysleksi*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.

Helland, T. (2002). *Neuro-Cognitive Functions in Dyslexia. Variations According to Language Comprehension and Mathematics Skills.* Unpublished Doctoral thesis, Department of Special Education, Faculty of Education, University of Oslo.

Helland, T. (2015) RI-5. *Hva er Risiko Indeks 5? Spørsmål og svar.* Retrieved from https://ri5.infovestforlag.no/nedlastinger/dokumenter/FAQ_BM.pdf

Helland, T., Plante, E. & Hugdahl, K. (2011). Predicting Dyslexia at Age 11 from a Risk Index Questionnaire at Age 5. *Dyslexia, 17*(3), 207-226. doi: 10.1002/dys.432

Helland, W.A., & Hollund-Møllerhaug, L. (2012). Children's Communication Checklist Second Edition. Et nytt instrument til kartlegging av språk og kommunikasjon hos barn og unge. *Norsk tidsskrift for logopedi, 2*, 61.

Hugdahl, K. (undated version). Stroop Test. Bergen: University of Bergen.

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Chichester, UK: Wiley-Blackwell.

Karlsen, J., Lyster, S.A.H. & Lervåg, A.O. (2016). Vocabulary development in Norwegian L1 and L2 learners in the kindergarten–school transition. *Journal of Child Language.* ISSN 0305-0009. doi: 10.1017/S0305000916000106

Karlsen, J., Lyster, S.A.H. &, Geva, E. (2016). Cognitive, linguistic and contextual factors in Norwegian second Language learners' narrative Production. *Applied Psycholinguistics*. Doi: 10.1017/S014271641500051X

Kristoffersen, K. E., & Simonsen, H. G. (2012, January 30). *Språkkartlegging – viktig og nødvendig*. Retrieved from https://www.hf.uio.no/iln/tjenester/kunnskap/sprak/sprakvansker/kartlegging/hvorfor-kartlegge.html

Kohnert, K. (2013) *Language disorders in bilingual children and adults*. 2. San Diego, CA: Plural. Retrieved from ProQuest eBook Central database.

Kohnert, K., Windsor, J., and Ebert, K. D. (2009). Primary or "specific" language impairment and children learning a second language. *Brain Lang*. 109, 101–111.doi: 10.1016/j.bandl.2008.01.009

Lervåg, A. og Aukrust, V.G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. In *Journal of Child Psychology and Psychiatry,* 51 (5), 612-620.doi.org/10.1111/j.1469-7610.2009.02185.x

Lervåg, A. & Melby-Lervåg, M. (2009). Muntlig språk, ordavkoding og leseforståelse hos tospråklige: En sammenfatning av empiriske studier. Norsk pedagogisk tidsskrift, 2009 (4), 264-277. Retrieved on the 14.februar 2018, from https://www.idunn.no/npt/2009/04/art06

Lyster, S.-A. H. & Horn, E. (2009). *Trog-2. Norsk versjon.* Manual, Pearson Assessment.

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, Second Edition (BPVS II). *Spesialpedagogikk 09*, 35-43.

McVeigh, C., Wylie, J., & Mulhern, G. (2017): Verbal and visuospatial working memory in immersion-educated bilingual children. *International Journal of Bilingual Education and Bilingualism*, DOI: 10.1080/13670050.2016.1271769

Meyers, J. E., & Meyers, K. R., (1995). *Rey Complex Figure Test and Recognition Trial.* Odessa, Fl.: Psychological Assessment Resources, Inc.

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Wiley series on personality processes. Developmental psychopathology, Vol. 1. Theory and methods* (pp. 357-390). Oxford, England: John Wiley & Sons. Downloaded from http://psycnet.apa.org/record/1995-97696-013

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10(6*), 719–726. doi:10.1111/j.1467-7687.2007.00623.x

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66,* 232–58. doi:10.1016/j.cogpsych.2012.12.002

Paap, K.R., Johnson, H. A. & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278. Downloaded from https://doi.org/10.1016/j.cortex.2015.04.014

Paradis, J., Genesee, F. & Crago, M. (2011). *Dual language development & disorders: A handbook on bilingualism and second language learning (2nd ed.)*. Baltimore, MD: Paul H. Brookes.

Rodina, Y. (2016). Narrative abilities of preschool bilingual Norwegian-Russian children. *International Journal of Bilingualism*. doi:10.1177/1367006916643528

Ryen, E. & Simonsen, H. G. (2015). Tidlig flerspråklighet – myter og realiteter. In *NOA Norsk som andrespråk 1*-2/2015, s. 195-217. Retrieved 18.04.2018 from http://ojs.novus.no/index.php/NOA/issue/archive

Statistisk sentralbyrå (2017b). *Hvordan går det med innvandrere og deres barn i skolen?* Retrieved from https://www.ssb.no/utdanning/artikler-og-publikasjoner/hvordan-gar-det-med-innvandrere-og-deres-barn-i-skolen

StatSoft, I. (2011) STATISTICA (data analyses software system) (Version 12). Retrieved from www.statsoft.com

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46*(1), 1–16. Retrieved from https://doi.org/10.1016/j.jcomdis.2012.08.002

Torkildsen, J., Morken, F., Helland, W. A., & Helland, T. (2015). The dynamics of narrative writing in primary grade children: writing process factors predict story quality. [journal article]. *READING AND WRITING*, 1-26. doi:10.1007/s11145-015-9618-4

Ward, J. (2010). *The Students Guide to Cognitive Neuroscience*. New York: Psychology Press.

Wechsler, D. (1991). *The Wechsler intelligence scale for children – Third edition.* San Antonio: The Psychological Corporation.

Wechsler, D. (2003). *The Wechsler intelligence scale for children – Third edition. Norwegian version.* Stockholm: Assessio Norge AS.

Wilhelmsen, M., Holth, B. A., Kleven, Ø. & Risberg, T. (2013). Minoritetsspråk i Norge: En kartlegging av eksisterende datakilder og drøfting av ulike fremgangsmåter for statistikk om språk. Oslo/Kongsvinger: Statistisk sentralbyrå. Retrieved

from http://www.ssb.no/utdanning/artikler-og-publikasjoner/_attachment/100940?_ts=13d3a8c3cf0

Øzerk, K. (2016). *Tospråklig oppvekst og læring*. Oslo: Cappelen Akademiske Forlag.

Appendices

Appendix 1: Application letter to NSD

1 Prosjektittel

Norwegian as second language (L2) in 3rd grade school children. How can transient between- language problems be separated from problems due to more specific language impairment and /or dyslexia?

2 Applicant

Project manager

Wenche A. Helland 1.amanuensis

Telephone: +47 90133397

3 Project field

Logopedics

4 Principal project objective and sub goals

It is often observed that children speaking Norwegian as their second language (L2) struggle with learning to read and write. For teachers it may be difficult to distinguish a transient between-languages problem from a problem due to specific language impairment (SLI) and/or dyslexia. However, how these problems are understood in school is essential to mediation. By definition both SLI and dyslexia are constitutional in origin, and research across languages suggests that their benchmark neurocognitive assets and deficits are universal. Thus, assessing both observable linguistic problems and underlying neurocognitive factors will pinpoint how the L2 problems can be understood. To our knowledge this is not further investigated in children with Norwegian as L2. Rather, focus has been at the symptomatic level, analyzing types of error in a cross-linguistic perspective.

Thus, the main aim of the present study is to separate transient between-language problems from problems due to more specific language impairment and /or dyslexia. We seek to achieve this goal by applying a linguistic and neuro-cognitive bottom-up approach assessing the impact on emergent literacy skills, i.e. in the period of schooling when learning to read and write is a main objective.

A. Background developmental information:
   - Parent and teacher evaluation of communicative skills and risk factors for developmental dyslexia
B. L2 (Norwegian) basic linguistic and literacy skills:
   - L2 comprehension as measured by tests of vocabulary and grammar
   - L2 production as measured by model sentences
   - L2 single word reading and spelling abilities
C. L2 written narrative skills Neurocognitive factors
   - Rapid naming (RAN)
   - Short term memory and working memory
   - Executive functions
   - Visuo-spatial skills

1 Project plan

Project plan summary

Verbal and literacy skills are central to children's academic performance throughout the school years, shown by the emphasis on reading and writing skills in the school curriculums and on national tests. Yet very few studies have assessed what background variables influence L2 children's linguistic competence. In the present project, we seek to assess the relation between a number of selected background variables on verbal language, reading and written narrative skills during the emergent literacy stage (grades 1-3). The study is unique in its focus on dynamic measures of the L2 language skills as well as in its neurocognitive approach. We expect the study to yield novel insights into the relation between verbal language competence and literacy in children with Norwegian as L2 competence.

Project plan (maximum 1000 words):

Motivation.

The school curriculum in first language (Norwegian) focuses on oral and written discourse and communication. More specifically, the students are expected to produce, analyze and evaluate narratives from the early school years. It is argued that the National tests given to 5th, 8th and 9th grades have had an impact on the content of what is being taught, and these tests focus on students' abilities to find information in texts, interpret and comprehend texts and reflect upon and evaluate their content and structure. As narrative skills are an important factor in educational success, we seek to assess the background variables that contribute to the development of these skills.

Based on previous research, we will assess the interplay of linguistic and benchmark neurocognitive skills in children with Norwegian as L2 in comparison to children with Norwegian as L1. Especially we will search for L2 (Norwegian) learners where parents and teachers question if learning problems may be caused by typical between-language problems or specific language problems and/or dyslexia. We will also obtain a broader spectrum of language assessments than what is used in most previous studies on narrative competence, by obtaining measures of formal language skills as well as wider communication abilities, such as pragmatic comprehension. Interestingly, previous research indicates that narrative skills are more resistant to recovery than other language abilities. For example, Fey et al. (2004) found that children who had recovered from early language impairment in the sense that they tested within the normal range on standardized language tests, still had significant problems in producing narratives. These findings suggest that narrative ability is a highly complex skill requiring the optimal co-functioning of several basic cognitive abilities, and that language competence is just one factor which influences the outcome.

Method.

Participants will be approximately twenty 3rd graders (age 8) having Norwegian as L2.

Inclusion criteria are

1) attended Norwegian schools since 1st grade (age 6)
2) no identified mental retardation
3) no identified neurological disorders

Control data will be anonymized data from our former study on typical Norwegian 3rd graders (see 2012/909/REK vest, 2012/909; Narrativ kompetanse.) published in international peer reviewed journal (Torkildsen, Morken, T. Helland, & W. A. Helland, 2015) using the same questionnaires and tests as in the present study.

Explicitly we will search children where teachers and/or parents are concerned of the progress of linguistic and literacy learning. References will also be to Norwegian L1 typical and

dyslexic developmental data from our research group (T. Helland & Morken, 2015; T. Helland, Plante, & Hugdahl, 2011)
Assessment tools.
The assessments will be carried out during school hours and take approximately two hours for each child.

A. Background developmental information to be filled out by parents and teachers:
- Children's Communication Checklist (CCC-2) (Bishop, 2003), a questionnaire designed to identify children with communication impairments (W. A. Helland, Biringer, T. Helland, & Heimann, 2009)
- Risk Index-8 (RI-8), an age adjusted version of RI-5, a parental questionnaire identifying early at-risk (ri5.infovestforlag.no/) and (T Helland et al., 2011)

See also (T. Helland, Jones, & W. A. Helland, 2017) for the use of the two questionnaires in combination. These questionnaires, as well as all written information, will be translated into the parents native language by the native language teachers.

B. L2 (Norwegian) basic linguistic and literacy skills:
- The British Picture Vocabulary Scale II (BPVS) (Dunn, Dunn, & Styles, 2003), Norwegian version
- The Test of Reception Of Grammar -2 (TROG-2)(Bishop, 2009) Norwegian version
- The model sentence task based on Ringstedmaterialet (Ege, 1984)
- Single word reading and spelling (Klinkenberg & Skaar, 2001)
- Written narratives using key stroke logging to measure the ability to construct a written text from a four picture story with the comic strip character "Der kleine Herr Jakob" (SCHUBI, undated). First, children will be asked to sort the four cards in the comic strip in chronological order. After completing this procedure, the examiner will instruct the children to write a story using all four picture stimuli. The stories will be tracked using the TextPilot—Research edition. Bergen: Include A/S, 2012.

C. Neurocognitive factors
- Rapid naming (RAN) from the Stroop battery (Golden, 1987; Hugdahl, undated version)
- Verbal short term and working memory, digit span from the Wechsler test battery (Wechsler & Undheim, 1974)
- Executive functions, using dichotic listening test (Bless, Westerhausen, Kompus, Gudmundsen, & Hugdahl, 2014)
- Visuo-spatial skills using the Rey-Osterieth Complex Figures Test with the Copy (RO copy) and Recall (RO recall) conditions (Meyers & Meyers, 1995; Spreen & Strauss, 1991)

1 Timetable/milestones for main activities/tasks

June 2017:	NSD Application.
September 2017:	Recruiting participants
October 2017:	Data collection
November 2017:	December 2017: Data analyses.
May 2018:	Submission of Master theses. Four master students in logopedics at the University of Bergen will write master thesis on data from the study.
2018-2020:	Paper submissions (tentative themes: Verbal and neurocognitive skills in L2 children at-risk of developmental dyslexia; Attention and linguistic skills in children with Norwegian as L2; Separating typical and deviant literacy development in L2 learners in Norwegian schools; Narrative writing in L2 learners in Norwegian schools)

2 Active partners

- Wenche Andersen Helland, Associate professor, Department of Biological and Medical Psychology, University of Bergen
- Frøydis Morken, Associate professor, Department of Biological and Medical Psychology, University of Bergen
- Turid Helland, Professor emerita, Department of Biological and Medical Psychology, University of Bergen
- Four master students in logopedics at the University of Bergen who will write their master thesis on data from the study

References

Bishop, D. V. M. (2009). Test for Reception of Grammar, revised, Norwegian version.

Bless, J. J., Westerhausen, R., Kompus, K., Gudmundsen, M., & Hugdahl, K. (2014). Self-supervised, mobile-application based cognitive training of auditory attention: A behavioral and fMRI evaluation. *Internet Interventions, 1*(3), 102-110. doi:http://dx.doi.org/10.1016/j.invent.2014.06.001

Dunn, L. M., Dunn, D. M., & Styles, B. (2003). British Picture Vocabulary Scale: 3rd Edition.

Ege, B. (1984). Ringstedmaterialet. En procedure til undersøkelse af børns realiserbare sprog. Herning: Special-pedagogisk forlag A/S.

Golden, C. J. (1987). Manual for the Stroop Color and Word Test. Chicago: Stoelting Company.

Helland, T., Jones, L. Ø., & Helland, W. (2017). Detecting Preschool Language Impairment and Risk of Developmental Dyslexia. *Journal of Research in Childhood Education*, 1-17. doi:10.1080/02568543.2016.1274928

Helland, T., & Morken, F. (2015). Neurocognitive Development and Predictors of L1 and L2 Literacy Skills in Dyslexia: A Longitudinal Study of Children 5–11 Years Old. *Dyslexia*, n/a-n/a. doi:10.1002/dys.1515

Helland, T., Plante, E., & Hugdahl, K. (2011). Predicting Dyslexia at Age 11 from a Risk Index Questionnaire at Age 5. *Dyslexia, 17*(3), 207-226. doi:10.1002/dys.432

Helland, W. A., Biringer, E. A., Helland, T., & Heimann, M. (2009). The usability of a Norwegian adaptation of the Children's Communication Checklist Second Edition (CCC-2) in differentiating between language impaired and non-language impaired 6- to 12-year-olds. [Populærvitenskapelig foredrag]. *Scandinavian Journal of Psychology, 50*(3), 287-292. doi:http://dx.doi.org10.1111/j.1467-9450.2009.00718.x

Hugdahl, K. (undated version). Stroop Test. Bergen: University of Bergen.

Klinkenberg, J. E., & Skaar, E. (2001). STAS. Standardisert test i avkoding og staving [Standardised test in decoding and spelling]. Hønefoss: Pedagogisk-psykologisk tjeneste.

Meyers, J. E., & Meyers, K., R. (1995). *Rey Complex Figure Test and Recognition Trial*. Odessa, Fl.: Psychological Assessment Resources, Inc.

SCHUBI. (undated). Collaboration Story box SCHUBI Sir James. *Schule und Zukunft - mit SCHUBI Lernmedien*. Schaffhausen: SCHUBI Lernmedien AG.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests (pp. 341-363)*. New York and Oxford: Oxford University Press.

Torkildsen, J., Morken, F., Helland, W. A., & Helland, T. (2015). The dynamics of narrative writing in primary grade children: writing process factors predict story quality. [journal article]. *READING AND WRITING*, 1-26. doi:10.1007/s11145-015-9618-4

Wechsler, D., & Undheim, J. (1974). *Wechsler Intelligence Scale for Children — Revised* (Vol. null).

Appendix 2: Information letter to Parents and Consent Form

Forespørsel om deltakelse i forskningsprosjektet

Norsk som andrespråk (L2) hos tredjeklassinger. Hvordan skille mellom en typisk mellomspråklig vanske, spesifikk språkvanske og/eller dysleksi?

Bakgrunn og formål
Dette er et spørsmål om å delta i en forskningsstudie for å få *vite mer om norsk som andrespråk hos barn i 3. klasse*. Vi ønsker å få mer kunnskap om hvordan vi skal skille mellom en vanlig vanske som mange barn med et annet morsmål kan ha, og en vanske som kan forklares ut fra en mer grunnleggende språkvanske/og eller dysleksi. Vi ønsker å starte med de barna som nettopp har lært seg å lese og skrive, og da er 3. klasse et godt utgangspunkt. Dere som mottar denne forespørselen har barn som har norsk som sitt andrespråk og har gått på norsk skole fra og med første klasse.

Hva innebærer deltakelse i studien
Studien innebærer at hvert enkelt barn gjennomfører en rekke oppgaver/tester i løpet av ca. to skoletimer, og at foresatte svarer på et spørreskjema. Testingen vil bli gjennomført av fire mastergradsstudenter i logopedi ved Universitetet i Bergen under veiledning av prosjektlederne. Testene er lagt opp slik at en undersøker de ulike byggesteinene i språket; hukommelse for språklige elementer, språkforståelse, muntlig og skriftlig fortelling til bilder. Erfaringsmessig er de utvalgte oppgavene og testene lystbetonte, og skulle derfor ikke føre til noen form for prestasjonsangst hos barna. Tidsbruk for hver enkelt test er fra 5 til 20 minutter, og det blir lagt opp til pauser etter barnas behov. Dataene som kommer fram kan ikke karakteriseres som spesielt sensitive. Foresatte blir bedt om å fylle ut et spørreskjema som fokuserer på tidlige risikofaktorer for å utvikle dysleksi. Skolens morsmålslærere vil bistå med oversettelse av informasjon og spørreskjema til foresattes språk. Vi ønsker å gjennomføre testingen i oktober 2017 i nært samarbeid med de involverte lærerne og med tilpasning til barnas timeplan.

Hva skjer med informasjonen om deg?
Alle personopplysninger vil bli behandlet konfidensielt. Testresultatene og informasjonen som registreres om eleven skal kun brukes slik som beskrevet i hensikten med studien. Alle opplysningene vil bli behandlet uten navn, fødselsnummer eller andre direkte identifiserende opplysninger. En kode knytter barnet til opplysningene om han/henne gjennom en navneliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til navnelisten og som kan finne tilbake til det aktuelle barnet. Det vil ikke være mulig å identifisere det enkelte barn, klasse eller skole i resultatene av studien når disse publiseres.
Studenter og prosjektledelse er underlagt taushetsplikt med hensyn til møtet med elevene.
Prosjektet skal etter planen avsluttes 31.12. 2020 og datamaterialet anonymiseres ved prosjektslutt.

Frivillig deltakelse
Deltakelse er selvfølgelig frivillig og er basert på informert samtykke underskrevet av foresatte. Dersom du ønsker at barnet ditt skal delta, undertegner du den vedlagte samtykkeerklæringen. Om du sier ja til å delta nå, kan du likevel senere trekke tilbake ditt samtykke senere uten å oppgi grunn for dette.
Ledergruppen for prosjektet er 1.amanuensis Wenche A. Helland, 1.amanuensis Frøydis Morken og Professor em Turid Helland, alle ved Institutt for biologisk og medisinsk psykologi, Universitetet i Bergen. Det vil i tillegg være fire masterstudenter knyttet til prosjektet.

Dersom du har spørsmål til studien, ta kontakt med prosjektleder Wenche A. Helland; telefon 90133397

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien «Norsk som andrespråk (L2) hos tredjeklassinger. Hvordan skille mellom en typisk mellomspråklig vanske, spesifikk språkvanske og/eller dysleksi?»

Jeg har mottatt informasjon om studien og ønsker at mitt barn

……. skal

delta. (barnets navn)

(Signert av prosjektdeltakers foresatte, dato)

Appendix 2: Approval of the study by NSB

Wenche Andersen Helland
Jonas Lies vei 91
5009 BERGEN

Vår dato: 09.08.2017 Vår ref: 54777 / 3 / HIT Deres dato: Deres ref:

Tilbakemelding på melding om behandling av personopplysninger

Vi viser til melding om behandling av personopplysninger, mottatt 19.06.2017. Meldingen gjelder prosjektet:

54777	*Norsk som andrespråk (L2) hos tredjeklassinger.Hvordan skille mellom en typisk mellomspråklig vanske, spesifikk språkvanske og/eller dysleksi?*
Behandlingsansvarlig	*Universitetet i Bergen, ved institusjonens øverste leder*
Daglig ansvarlig	*Wenche Andersen Helland*

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database.

Personvernombudet vil ved prosjektets avslutning, 31.12.2020, rette en henvendelse angående status for behandlingen av personopplysninger.

Dersom noe er uklart ta gjerne kontakt over telefon.

Vennlig hilsen

Marianne Høgetveit Myhren

Hildur Thorarensen

Kontaktperson: Hildur Thorarensen tlf: 55 58 26 54 / hildur.thorarensen@nsd.no
Vedlegg: Prosjektvurdering

Personvernombudet for forskning

Prosjektvurdering - Kommentar

Prosjektnr: 54777

FORMÅL
Barn som har norsk som andrespråk (L2) strever ofte med å lære seg å lese og skrive. Det kan være vanskelig å skille et forbigående mellomspråkproblem fra et problem som har sin årsak i medfødte språklige vansker.
Forsking på tvers av ulike språk indikerer at grunnleggende nevrokognitive styrker og svakheter er universelle. Hovedmålet med studien er derfor å skille forbigående mellomspråklige problem fra mer grunnleggende språklige vansker og/eller dysleksi. Denne kunnskapen vil være av stor betydning for hvordan det kan legges til rette for å hjelpe barna med å tilegne seg gode lese og skriveferdigheter de første skoleårene, ferdigheter som er av avgjørende betydning for barnas videre skolefaglige og sosiale utvikling.

INFORMASJON OG SAMTYKKE
Utvalget informeres skriftlig og muntlig om prosjektet og samtykker til deltakelse.
Informasjonsskrivet er i utgangspunktet godt utformet, men det må tydeliggjøres hvilke opplysninger lærer skal svare på i spørreskjema om barnet, slik at taushetsplikten ikke er til hinder for dette.

Merk at når barn skal delta aktivt, er deltagelsen alltid frivillig for barnet, selv om de foresatte samtykker. Barnet bør få alderstilpasset informasjon om prosjektet, og det må sørges for at de forstår at deltakelse er frivillig og at de når som helst kan trekke seg dersom de ønsker det.

SENSITIVE OPPLYSNINGER
Det behandles sensitive personopplysninger om helseforhold.

INFORMASJONSSIKKERHET
Personvernombudet legger til grunn at forsker etterfølger Universitetet i Bergen sine interne rutiner for datasikkerhet.

PROSJEKTSLUTT
Forventet prosjektslutt er 31.12.2020. Ifølge prosjektmeldingen skal innsamlede opplysninger da anonymiseres. Anonymisering innebærer å bearbeide datamaterialet slik at ingen enkeltpersoner kan gjenkjennes. Det gjøres ved å:
- slette direkte personopplysninger (som navn/koblingsnøkkel)
- slette/omskrive indirekte personopplysninger (identifiserende sammenstilling av bakgrunnsopplysninger som f.eks. bosted/arbeidssted, alder og kjønn)