UNIVERZITA KARLOVA

Filozofická fakulta

Ústav anglického jazyka a didaktiky

DIPLOMOVÁ PRÁCE

Bc. Tatiana Lebedeva

Nonword Repetition Task for Detecting Developmental Language Disorder in English-Russian Bilingual Children

Test opakování pseudoslov pro odhalení vývojové poruchy řeči u anglicko-ruských bilingvních dětí

Praha 2024

Vedoucí práce: doc. Luca Cilibrasi, Ph.D.

Acknowledgements

I would like to thank my supervisor, Luca Cilibrasi, Ph.D., whose guidance, support, and expertise were invaluable throughout the entire process of completing this master's thesis. I am grateful for the time invested and his openness to all creative ideas that lie in the foundation of the present work.

I extend my gratitude to all participants and their parents for their contributions shaping the outcomes of this study. To my friends and colleagues who helped to record and assess the stimuli and other materials – this work would be incomplete without it.

Lastly, I want to thank my family for the support and good advice along the way.

Prohlášení autora

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu, a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

Author's declaration

Hereby I declare that I have written this master's thesis on my own, that I have properly cited all sources and literature used, and that this thesis has not been used as part of any other university program in order to obtain another or identical degree.

V Praze dne …………. 2024 …………………………………

Tatiana Lebedeva

Abstract

The present thesis is concerned with designing and evaluating a novel language-specific nonword repetition task (NWR) capable of identifying developmental language disorder (DLD) in English-Russian bilingual children. The proposed NWR task manipulates three variables, item length, morphological complexity, and phonological complexity that reflect respective language processing mechanisms, namely, phonological working memory strain, phonological processing, and the knowledge of grammatical rules and long-term memory. The main question of the study was whether the proposed task could show adequate performance results for a typically developing bilingual child when matched with standardized language ability tests (TROG2, YARC, CNRep). To evaluate the efficacy of the task, a case study with one participant was conducted. The participant was a bilingual child (7;5) with simultaneous acquisition of English and Russian in an English-dominant environment. The findings showed that the proposed task passed the evaluation procedure and yielded expected patterns when matched against standardized tests both in terms of the patterns of difficulty and language dominance. As a result, the proposed NWR task demonstrated the potential for distinguishing DLD from typical development in bilingual children speaking English and Russian. The limitations of this study include the need for quantitative analysis with a larger sample of participants to determine the task's precision and reliability.

Keywords: bilingualism, bilingual acquisition, child bilingualism, developmental language disorder, language acquisition, phonological acquisition, simultaneous bilingualism, nonword repetition

Abstrakt

Předkládaná diplomová práce se zabývá návrhem a vyhodnocením nového testu opakování pseudoslov (Eng. nonword repetition task = NWR) schopného detekovat vývojovou jazykovou poruchu řeči (Eng. developmental language disorder = DLD) u anglicko-ruských bilingvních dětí. Navrhovaný test opakování pseudoslov pracuje se třemi proměnnými: délkou položky, morfologickou složitostí a fonologickou složitostí, které odrážejí příslušné mechanismy zpracování jazyka, a to fonologickou pracovní paměť, fonologické zpracování, znalost gramatických pravidel a dlouhodobou paměť. Hlavní otázkou studie bylo, zda navrhovaný test opakování pseudoslov může prokázat adekvátní výkonnostní výsledky u typicky se vyvíjejícího bilingvního dítěte při porovnání se standardizovanými testy jazykových schopností (TROG2, YARC, CNRep). Pro vyhodnocení efektivity úkolu byla provedena případová studie s jedním účastníkem. Účastníkem bylo bilingvní dítě (7;5) se simultánním osvojením angličtiny a ruštiny (Eng. simultaneously bilingual) v prostředí s převahou angličtiny. Zjištění ukázala, že navrhovaný test opakování pseudoslov prošel evaluační procedurou a přinesl očekávané vzorce při porovnání se standardizovanými testy, a to jak z hlediska vzorců obtížnosti, tak z hlediska jazykové dominance. Výsledkem je, že navrhovaný test prokázal potenciál odlišit DLD od typického jazykového vývoje u bilingvních dětí mluvících anglicky a rusky. Mezi omezení této studie patří potřeba kvantitativní analýzy s větším vzorkem účastníků, aby bylo možné určit přesnost a spolehlivost testu.

Klíčová slova: bilingvismus, bilingvní osvojování jazyka, dětský bilingvismus, osvojování jazyka, osvojování fonologie, opakování pseudoslov, simultánní bilingvismus, vývojová porucha řeči

Table of Contents

List of Tables

List of Abbreviations

CPM = Colored Progressive Matrices

DLD = developmental language disorder

L1 = first language

L2 = second language

LTM = long-term memory

MNW = morphological nonwords

ND = neighborhood density

NNW = non-morphological nonwords

NWR = nonword repetition

PP = phonotactic probability

PWM = phonological working memory

SRep = sentence repetition

SLI = specific language impairment

TD = typically developing

TROG = Test for Reception of Grammar

YARC = York Assessment of Reading for Comprehension

1. Introduction

In the field of child language acquisition, the phenomenon of bilingualism has gathered considerable attention for its visible impact on linguistic development. The evidence that a child growing up in a bilingual context uses two languages differently from a child growing up monolingual or learning a second language later in life is obvious to the parents, caregivers, and scholars. The juxtaposition of two linguistic systems within a single mind, a developing mind, presents a dynamic context wherein language acquisition unfolds following specific patterns. One such well-known observation is that the two languages in the mind of a bilingual child influence one another and create a system that is often unbalanced and gravitates towards one of the languages (Grosjean, 2013a). Because of this interplay, bilingual children often tend to make more grammatical errors, have larger gaps in the vocabulary knowledge and appear to be less proficient in the two languages than a monolingual peer. It is well-known among the scholars that to discriminate or disentangle typical bilingual development, which always presupposes certain linguistic difficulties, from a developmental disorder in bilinguals can be very challenging (see for example Armon-Lotem et al., 2015a).

Therefore, assessment of language disorders in bilingual children is often a central element in the discourse of bilingual language acquisition. A developmental language disorder (DLD), also known as the specific language impairment (SLI) in some literature is a frequent ailment affecting 7-10% of preschool children (Armon-Lotem & Jong, 2015). There are numerous clinical and research instruments already existing and being developed that target language assessment in the English-speaking population, such as the nonword repetition tasks CNRep (Gathercole et al., 1994) and NRT (Dollaghan & Campbell, 1998) developed specifically for identifying DLD; there are also numerous tools that sample particular language areas such as vocabulary, e.g., British Picture Vocabulary Scales (BPVS), grammar, e.g., Test for Reception of Grammar (TROG), or Verb Agreement & Tense Test (VATT) that could aid in identifying this type of language delay in children. Many of these and similar tools, however, have often been designed with monolingual development in mind, meaning that the cut-off points and normative data might not reflect the reality of bilingual development. This is why in the past few years, such research initiatives as COST Action IS0804 (2019) have been working on creating assessment instruments for children growing up with more than one language – to provide the parents, teachers, and speech and language therapists with reliable tools to assess language disorders.

The present thesis attempts to aid this research by filling one of the gaps and present a test that could be potentially used for the assessment of bilingual children in the English-Russian language pair. The type of test presented here is known as the nonword repetition task (NWR), frequently used as part of the DLD assessment. The task presents a list of stimuli, nonwords, that a child is asked to repeat during the assessment. A special detail about the present task is that it was developed specifically for bilingual children aged 7 or older and growing up in the English-Russian context. The relevance of such an instrument is provided by the fact that the Russian clinical tradition did not have standardized instruments for such assessments for a very long time and relied on qualitative measures, such as observation by a specialist (Lopukhina A. et al., 2019). It is only recently that comprehensive Russian-language assessment tools supported by quantitative analysis started appearing, such as RuCLAB (Lopukhina A. et al., 2019; *RuCLAB (Russian Child Language Assessment Battery)*, 2023). There are, therefore, certain grey areas in the study of bilingual development with Russian as one of the languages in the pair.

The objectives of the present thesis are twofold: to design a language-specific English-Russian NWR task and to perform the evaluation of the proposed task with a case study. As it comes to task design, the present NWR task aims to provide an instrument that could not only distinguish the presence of DLD in a bilingual child, but also determine the cognitive mechanisms that are likely to underly a particular DLD case. The proposed task is designed with the knowledge that such factors as deficiency in the phonological processing, phonological working memory (phonological loop), and morphological processing are very likely to be affected in DLD. The stimuli in the proposed task are manipulated in terms of length, phonological complexity, and morphological complexity, corresponding to the three areas of language processing. Therefore, the stimuli in the present task were developed in such a way so as to differentiate which of the three areas are affected in a given child, should a case of DLD be determined. The construction of the test is described in chapter 3.1. *The construction of bilingual English-Russian NWR task,* and the lists of stimuli are presented in Table 3 and Table 4 in the chapter and in the appendix (1).

In the second part of this thesis, the proposed NWR task is tested with a typically developing bilingual child who performs the task and whose results are then studied qualitatively. Additionally, the child is asked to participate in standardized language ability assessment (CNRep, TROG-2, YARC) and a cognitive development assessment (Raven's CPM), the results of which are then matched to the result on the proposed NWR task in order to validate the child's performance on the proposed NWR task. The aim of the second part of the work is to answer the question of whether the given NWR task can show adequate performance results for a TD bilingual child when evaluated against standardized language ability tests in terms of patterns of difficulty. Even though the proposed experimental setup with a case study is not a definitive answer to the test's clinical validity, it serves as the first step on the way of its evaluation. If the present instrument passes the proposed evaluation procedure, it could be then offered for further studies with a normative sample. The case study is described in chapter 4, *Experiment*, and the evaluation of results follows in chapter 5, *Discussion*. The limitations of the present study are discussed in sub-chapter 5.1. following the discussion.

2. Theoretical overview

2.1. Language acquisition in monolingual and bilingual children

Language acquisition begins long before a child learns to speak. By the time the first words are uttered, infants will have learned to perform many complex tasks. Within the first months of life, they learn to distinguish speech sounds from other noises, they exercise and then lose an ability to differentiate between virtually all human languages, until eventually, they attune to their native tongue (e.g., Bosch & Sebastián-Gallés, 1997; Eimas et al., 1971). Considering general language acquisition processes, bilingual-to-be infants exposed to two languages from birth pass through the same phases as their monolingual peers, such as babbling, one-word, and two-word stages (Yip, 2013). The following chapter will provide an overview of selected language acquisition processes and discuss crucial linguistic milestones of monolingual and bilingual development.

2.1.1. Early phonological development in monolingual and bilingual children

As found by Mehler and colleagues (1988), infants as young as four days old and up to two months of age can discriminate between an unfamiliar and the native language. It has also been demonstrated that infants are able to tell apart between two unfamiliar languages, for instance, infants born into a French environment could detect a change when the languages switched to English and Japanese (Nazzi et al., 1998). Discussing newborns' remarkable sensitivity to languages, Guasti (2002) explains that in order to perform this task, infants should be able to build mental representations of the languages and compare them, relying on specific linguistic properties "that capture infants' interest and that can be extracted from utterances in very little time and with limited exposure" (Guasti, 2002, p. 30).

In this regard, numerous studies have converged in finding that infants build acoustic representations relying on the rhythmical properties of individual languages – an idea that lies at the foundation of the rhythm-based language discrimination hypothesis (Mehler et al., 1996; Nazzi et al., 1998). Broadly construed, this hypothesis suggests that infants up to two months of age can differentiate between stress-timed (e.g., English, Russian), syllable-timed (e.g., French), and mora-timed (e.g., Japanese) language classes, unless the two languages in a given set belong to the same class. To illustrate, while successfully telling apart English (stressed-timed) and Japanese (mora-timed), four-day-old infants are not able to distinguish two stress-timed languages such as English and Dutch (Nazzi et al., 1998).

Besides the capacity to discriminate native and non-native languages belonging to different rhythmical groups soon after birth, infants are also sensitive to phonemic contrasts within one given language. This means infants can register changes in voicing (e.g., /ba/ and /pa/ in Eimas et al., 1971), place (e.g., /ba/ and /da/ in Eimas, 1974), and manner of articulation (e.g., /ba/ - /ma/ in Eimas & Miller, 1980). Similarly, studies have shown that infants are sensitive to vowel contrasts; for instance, Trehub (1973) reported on infants' ability to distinguish between /a/ and /i/, /i/ and /u/ vowel pairs. Furthermore, it has been shown that infants do not need to have prior experience with the language to be able to discriminate the contrasts. For example, Streeter (1976) investigated the discrimination of changes in the voicing of /p/ and /b/ in the infants acquiring Kikuyu, where such a contrast is irrelevant, and found that infants could successfully register the contrast even though it was unfamiliar for them.

The processes of language and phoneme discrimination described above are general for all infants, regardless of the number of languages they are exposed to. The bilingual context, however, is characterized by some differences in language development. First of all, if an infant is growing up in a bilingual environment, he or she needs to acquire two separate language systems. Evidence suggests that from the very beginning, bilingual-to-be infants learn to distinguish between their native tongues even if they belong to the same rhythmic group. For example, Bosch and Sebastián-Gallés (2001) have shown that four-month-old infants learning Spanish and Catalan register different languages even though these languages are both syllable-timed. The authors have suggested that bilingual acquisition can thus propel an early ability to tell the difference between familiar languages – a skill that monolingual infants have not been shown to possess.

Moreover, bilingual children might have a different approach to distinguishing between the native and foreign languages. Bosch and Sebastián-Gallés (1997) examined the differences of bilingual perception in a study with a visual orientation procedure. In this procedure, an infant is seated facing a monitor and two speakers, which are used to present visual and auditory stimuli to the infant. When the auditory stimuli are played in the speakers, the infant turns his or her head or looks towards one of the speakers, and the eye movements are tracked and then analyzed. This study has demonstrated that monolingual children responded more rapidly to the native language stimulus, while bilingual children responded to the foreign language faster. The authors mention that it is hard to interpret this particular behavioral pattern and speculate about potential causes (see experiment 4 in the source for a review).

To draw a conclusion, the present overview of studies shows that bilingual and monolingual early language development follows the same path in general aspects. Thus, all infants before six months of age and regardless of the number of languages acquired learn to discriminate between native and foreign tongues relying on rhythmical properties of languages. They can also distinguish between individual consonant and vowel sounds even in the language they have never been exposed to. However, it has been noted that in the early months, bilingual development differs from monolingual in certain aspects, and, as infants grow older, this difference becomes even more visible.

2.1.2. Phonological development after six months of age

It has been established that mere days after birth, infants can register the difference between any languages they are presented with. However, already by the second month of life, this ability becomes less pronounced. There is evidence that around this age, infants start discriminating only between a native and a foreign language, in contrast to their earlier ability to tell apart different foreign languages with different rhythmical patterns (Christophe & Morton, 1998). For example, two-month-old infants growing up in an English environment (stressed-timed language) cannot discriminate French from Japanese but can tell apart Dutch (also stress-timed) from Japanese (Christophe & Morton, 1998). To simplify, data suggest that two-month-old infants begin perceiving languages either as native-like or non-native-like. This is the first step to a process known as perceptual narrowing, which is best observed in infants after the 6-8 months of life.

In language acquisition, perceptual narrowing or reorganization refers to an increasing sensitivity to the sound of native speech followed by a decline in sensitivity to non-native phonetic contrasts (Werker & Tees, 1984). This effect is accountable for the finding that having surpassed the sixth month, babies get worse at discriminating between non-native phonemes; thus, for example, Japanese babies lose the ability to distinguish /l/ and /r/. Further evidence for this is described in a study by Werker & Tees (1984), where English infants in two age groups were presented with a set of non-English contrasts, such as Hindi /t̪a/ - /ta/ (unvoiced, unaspirated retroflex/dental contrast). The second experiment in this study showed that by 10-12 months, the infants' ability to discriminate non-native contrasts significantly declined.

The progression of this effect has been described by Werker & Tees (1984) thusly: the ability to distinguish all contrasts present in 6-and 8-month-old infants becomes much weaker at 8-10 months and disappears by 10-12 months. The authors explain that while 6–8-month-old monolingual infants are listening to various phonological aspects of any language to pick up as many contrasts as possible, by 12 months, they learn to discriminate only those phones that will help them map sound onto the meaningful units of the native language, which is expected to facilitate word learning (Werker & Tees, 1984). Therefore, after 6-8 months of age, monolingual babies can no longer distinguish phonemes in non-native languages.

Conceivably, the bilingual acquisition has been found to affect perceptual narrowing and phoneme discrimination. For example, 10–11.5-month-old bilinguals learning English and Mandarin Chinese have been shown to discriminate a foreign (Hindi) contrast when this ability has already been lost in monolinguals (Singh et al., 2017). In addition, neuroimaging studies, e.g., by Petitto and colleagues (2012) and Ramírez et al. (2017), showed that older bilingual babies retained sensitivity to phonetic contrasts of foreign languages when monolingual babies have lost the phoneme discrimination ability. It is thus assumed that early bilingual exposure attenuates perceptual narrowing, allowing infants to perceive more contrasts than their monolingual counterparts can. Consequently, Petitto et al. (2012) propose that early exposure to more than one language is advantageous to supporting "language analyses" in infants, including phonological processing (2012, p. 140).

Another important development that occurs in monolingual and bilingual children after the sixth month of age and follows perceptual narrowing is the onset of productive abilities, namely, babbling. In a word, babbling means producing simple meaningless syllable sequences, such as *bababa* (Jusczyk, 1997). The maturation of productive skills passes through some of the stages of phonological development: when infants start to babble, their repertoire of sounds is not language-specific. In fact, some features of early-stage babbling appear to be universal, such as the preference for stop consonants and nasals, and the higher frequency of /a/ and /æ/ rather than /i/ and /u/ (Guasti, 2002). Then, around 8-10 months, infants begin relying on their linguistic experience: studies suggest that infants born and living in various speech communities tend to produce the sounds that are more frequent in their target language. For instance, a longitudinal study with infants from French, Swedish, English, and Japanese backgrounds showed similarities between the adult production of consonants and infants' babbling (Boysson-Bardies et al., 1992). Bilingual babbling, too, follows the phonology of the native languages. For example, Maneva & Genesee (2002) showed that 10–15-month-old infants produced different babbling patterns depending on the language of the speaker who interacted with them.

To conclude the chapter on infant phonological development in monolingual and bilingual contexts, it is necessary to stress that all children, regardless of the languages they speak, follow the same crucial stages in the first months of life, such as sensitivity to rhythmic properties of languages, universal sensitivity to phonemic contrasts, perceptual narrowing, and the onset of babbling with reference to the native sound system. The bilingual context adds a few pieces to this picture: it has been shown to alter phonemic discrimination and broaden the scope of sounds infants remain sensitive to even after they surpass 6-8 months, which means that bilingual acquisition may influence an early ability to tell more linguistic contrasts apart, such as contrasts between familial languages. The following chapters will take the discussion on bilingualism further to observe what other cognitive effects bilingualism contributes to in the later years of life.

2.2. Bilingualism as a concept

Today, bilingualism is a widespread phenomenon: Grosjean (2013a, p. 6) mentions that more than 50% of the inhabitants of Europe report speaking a second language "well enough to have a conversation in it." Curiously, a few decades back, speaking a language "well enough" would not be adequate to count as a bilingual competence for authors like Bloomfield (1984), who construed bilingualism as a native-like mastery of two languages. Other views, however, reinvented the concept of bilingualism and made it less rigid, allowing for such notions as, for example, *passive bilingualism* (a scenario when a speaker understands a language but lacks written or oral proficiency) rather than treating it as a binary system based on possessing or lacking a native-like capacity. This chapter discusses the concept of bilingualism, shows how it is treated in the field today and offers a classification of the types of bilingualism by language competence, age, and the onset of exposure.

2.2.1. Bilingualism defined

To begin with, in modern practice, bilingualism is often defined as the use of two languages in everyday life (Baker, 2006; Grosjean et al., 2013). However, the notion of bilingualism carries a host of uncertainties associated with its definition and classification. The classic interpretation by Leonard Bloomfield (1984) treating bilingualism as "the native-like control of two or more languages" is no longer sufficient as it introduces ambiguity and raises more questions. For instance, how should those whose first language is considerably more advanced than the second language be treated? Should be those who use their second language only in specific contexts and do not require a native-like proficiency called bilingual? To offer more clarity, a modern approach has moved away from the idea of native-like proficiency and regards bilingualism in the context of continua involving the age of onset, the context of usage, and degrees of proficiency. As an illustration, the definition of bilingualism offered by Grosjean (2013b) and Yip (2013) does not presume equal skills in both languages; instead, it embraces the reality where the ability to speak and write, or *productive competence*, and the ability to read or understand, or *receptive competence* may be developed unevenly (Baker, 2006). According to this approach, people who read and understand a language but do not speak or write in it will be classified as so-called passive bilinguals rather than monolinguals.

At the same time, a common interpretation of bilingualism assumes that the first language (L1) and the second language (L2) can be used in different, not necessarily overlapping, domains, such as school and home. Thus, according to the so-called complementarity principle (Grosjean, 1997), bilinguals usually acquire and use their languages for a variety of purposes in different domains of life. In this regard, the fact that a speaker's language skills in L2 may be lacking in some of the domains will not exclude him or her from the bilingual continuum. This view also presumes that one of the languages may be wider in scope, i.e., having a larger vocabulary and grammatic inventory as well as overall fluency or a combination of these. It is customary to term such a language as the dominant language (Grosjean, 1997; Yip, 2013).

Although it is common for bilinguals to have a dominant and a weaker language, there might be speakers with near equal mastery of the two languages. To quote Baker (2006, p. 9): "someone who is approximately equally fluent in two languages across various contexts may be termed an *equilingal* or *ambilingual*, or, more commonly, a balanced bilingual." Conversely, a speaker who uses one language more often than the other, has a larger vocabulary in one of the languages or has acquired more complex grammatical structures in one of the languages is considered an unbalanced bilingual (Yip, 2013). Even though balanced bilingualism is a very rare phenomenon (Grosjean, 1997), it might be practical to delineate balanced from unbalanced bilinguals as these two groups tend to demonstrate different performance results in studies focused on various psycholinguistic phenomena. For example, unbalanced bilinguals tend to be influenced to a varying extent by language interference in reading or speech production (Yip, 2013).

To sum up this section, bilingualism is a complex notion that is best defined within a framework of related functional concepts referring to domain-specificity, competence, proficiency, and the balance of skills in the speaker. As has been shown, a modern take on bilingualism allows to regard it broadly and include scenarios where one's language proficiency cannot be described "native-like". To proceed further, bilingualism needs to be described with respect to other dimensions, such as the age of onset and the order of acquisition.

2.2.2. Types of bilingualism

Two crucial parameters that characterize bilingualism are the age of acquisition and the order in which the child is exposed to the two languages. On the base level, bilingualism is divided into *simultaneous* and *sequential* (Armon-Lotem et al., 2015a) or *successive* (Grosjean, 2013a). Traditionally, the type of bilingualism whereby children are exposed to two languages either at birth or in early childhood is known as simultaneous bilingualism or Bilingual First Language Acquisition (BFLA) (Yip, 2013). However, the cut-off points for the period known as "early childhood" are treated differently by different scholars. While some suggest the age of five to be the end of the period when language acquisition can be classified as simultaneous, (e.g., Yip, 2013), others scholars set a stricter limit at the age of three (e.g., Montrul, 2009). These landmarks are determined based on the cognitive abilities of children at the respective periods of time: while the age of three-four is the approximate age when basic syntactic knowledge has already developed (Montrul, 2009), by the age of five, TD children should have acquired the knowledge of pragmatics and are able to participate in various communicative contexts (Baird, 2008).

Sequential bilinguals are those who first acquire one language, then learn a second language after the foundations for first language acquisition had been laid out (Montrul, 2009). Sequential language acquisition may be further divided into early and late stages, relative to the age of onset and the presence or absence of explicit instruction in a language. Thus, Montrul (2009) describes *early sequential* bilingualism as a condition that occurs around three-four years of age when the individual has acquired basic grammar of the first language, and *late child L2 acquisition* as a condition occurring in the elementary school years when children are receiving explicit instruction in one or two of the languages. There is yet another type of successive bilingualism, namely *late or adult second-language learning,* which takes place after the onset of puberty when an individual comes into L2 learning with fully developed grammatical and phonological systems of the first language (Montrul, 2009).

It is customary to divide the bilingual continuum into the said stages because when compared, these groups may display different language processing. To illustrate, a study by Guillelmon and Grosjean (2001) of gender marking processes in bilinguals revealed that early English-French bilinguals (age of onset of bilingualism 5.4 years old) differed from late sequential English-French bilinguals (age of onset of bilingualism 24.8 years old) in their treatment of gender marking cues in French. In the experiment, participants were asked to listen to short noun phrases made up of a determiner, an adjective, and a noun, and they were asked to repeat the noun. The reaction times of early bilinguals and late bilinguals were measured. While early bilinguals were affected by wrong gender marking determiners, late bilinguals were insensitive to either gender-congruent or incongruent determiners.

The authors of the experiment concluded that the gender processing mechanism, which may serve as a cue to speed up word recognition and which was acquired by early bilinguals, was not acquired by late bilinguals at all. Consequently, the age and order of acquisition may influence not only one's usage of these languages but also learning of the new ones. Therefore, the distinction between several acquisition stages does not only stem logically from the language acquisition reality, but it is also functional in that it might explain different outcomes of language acquisition and learning throughout one's life.

2.3. Cognition and bilingualism

The following chapter is conceived as a bridge between the general outline of language processing in bilinguals and the language deficit discussed further on. It introduces the core concepts of phonological processing in monolinguals and bilinguals as well as briefly discussing their roles in language acquisition and use. Finally, it provides an overview of the advantages of bilingual development from cognitive and functional perspectives.

2.3.1. Phonological processing and memory

One of the fundamental mechanisms in language acquisition and use is memory. It is customary to distinguish between long-term and short-term memory. Long-term memory (LTM) "stores information for long periods of time (and perhaps permanently)" (Groot, 2013, p. 171). In other words, long-term memory is a vast storage that contains the knowledge and information gathered throughout one's life. Most importantly for the present discussion, long-term memory acts as a lexicon repository that stores lexical representations necessary for learning and using a language. In contrast, short-term memory can be defined as a faculty of a human mind that stores a limited amount of information that is readily accessible temporarily (Cowan, 2008). The two types of memory differ in a fundamental aspect: only short-term memory is characterized by quick temporal decay and capacity limits (Cowan, 2008).

There is another type of memory distinguished, known as working memory. It is related to short-term memory as it "retrieves information from long-term memory and holds (and manipulates) it for the duration it is needed to perform some mental operation" (Groot 2013, p. 171). As summarized by Cowan (2008), it became prominent in the field when Baddeley and Hitch (1974) demonstrated that temporary memory cannot rely on a single module for all its functions. Thus, according to the Baddeley and Hitch model (1974; revised 2000), the working memory consists of the phonological loop (a subsystem concerned with verbal and acoustic information), the visuospatial sketchpad (provides its visual equivalent), the central executive (which controls the two systems), and the episodic buffer. An essential aspect of this model is that it views verbal-phonological and visual and spatial representations as two separate kinds of storage.

Since this thesis is concerned with phonological processing, it is useful to further narrow down on the working memory functions involved in it. The phonological working memory, also known as the phonological loop (e.g., Casalini et al., 2007; Montgomery, 1995), acts as a storage of immediately presented new phonological information. As summarized by Baddeley (2000, p. 419), "it is assumed to comprise a temporary phonological store in which auditory memory traces decay over a period of a few seconds, unless revived by articulatory rehearsal." The phonological loop has been found crucial in learning new phonological forms, including those of a foreign language. For instance, a study involving a participant with a phonological memory deficit showed that she was unable to learn any foreign words but could learn associations for words in a native language as successfully as unimpaired controls (A. Baddeley et al., 1988). The study has demonstrated that the factors that impair the performance of the phonological loop may disrupt foreign language learning while leaving native language learning intact. Another important point to add to the present description is that working memory and long-term memory are not disconnected: long-term memory acts as a lexical repository, which supports phonological working memory. Therefore, working memory relies on the knowledge stored in the mental lexicon in various tasks, such as reading, speech production or comprehension (de Groot, 2013).

2.3.2. Lexicon activation in bilinguals

When it comes to bilingual versus monolingual language processing, one of the most intriguing and widely studied questions concerns language activation in the bilingual lexicon. Broadly construed, the mental lexicon is a repository of words, their forms, and meanings, which is part of the long-term memory storage (de Groot, 2013). To be more specific, it is customary to differentiate at least three systems within the mental lexicon, namely a concept system that stores word meanings, a lemma system that stores grammatical information about particular words, and a phonological system that stores phonemes and sound sequences for words (Stille et al., 2020). Various models of the lexicon and speech production treat these systems differently, for example, Caramazza (1997) describes three independent networks in the Independent network model (namely the lexical-semantic network, lexical-syntactic network, and a subnetwork containing grammatical categories). Therefore, the lexicon combines the information from all linguistic levels, such as phonology, orthography, syntax, and morphology.

One of the main points of interest regarding the mental lexicon is whether speech processing in bilinguals is *selective* (i.e., only one language is activated) or *non-selective* (i.e., two languages are active and intervene on any occasion during speech processing). To start with, Grosjean (2013b) suggests that bilinguals are capable of processing speech in bilingual and monolingual modes, which constitute a continuum rather than being strictly delineated. Thus, *the bilingual processing mode* is activated when the input is coming in two languages or contains code-switching and borrowings. In such a case, bilingual speakers would have both lexicons activated, albeit to a different degree: one of the languages would be the most active (known as the base language), while the other one would be less active (Grosjean, 2013b).

In contrast, *monolingual processing* occurs when the input is coming in one language. Nevertheless, even in the case of monolingual input, both languages could become activated in a bilingual speaker. For example, studies such as by Spivey and Marian (1999) using an eye-tracking technique with Russian-English bilinguals showed that even when the task is administered in one language only, both lexicons can be consistently activated. Another study by Ju and Luce (2004) on Spanish-English bilinguals showed that the occurrence of subtle phonetic input, such as a single phoneme in another language, could be enough to activate both lexicons.

On the other hand, artificial experimental conditions designed to be strictly monolingual (i.e., eliminating any associations with the other language and controlling for the language in which the task was administered) led the participants towards selective processing in one language only (see for example, Marian & Spivey, 2003). It thus follows that speech processing in bilinguals can be both selective and non-selective, depending on such factors as the communicative context and place, and the familiarity of the interlocutor with one or two of the languages. Grosjean (2013b) also adds that speech processing might be affected by language proficiency: if the speaker is processing the dominant language, then the weaker language might be deactivated, but when the weaker language is being processed, the dominant one is likely to interfere with the processing.

2.3.3. The Interplay of bilingualism and cognitive development in childhood

Historically, bilingualism (coupled with multilingualism) was regarded as a disadvantaged or at least undesired route of language development (e.g., Barke, 1933; Jones & Stewart, 1951; Yoshioka, 1929). Acquiring two languages in childhood was believed to cause a delay in language acquisition, lead to a state where the two languages would be confused in a speaker's mind, or be a hardship 'devoid of apparent advantage' for young children (Yoshioka, 1929, p. 479). Even though these and similar misconceptions have been long refuted and it has been established that both bilingual and monolingual children reach expected developmental milestones within the same age spans (Grosjean, 2012), recent findings show that bilingualism does come with the features that could be construed negatively: for example, bilingual speakers tend to have a smaller vocabulary than monolingual speakers of each language (Bialystok & Luk, 2012), show slower lexicon access in speech production tasks (Ivanova & Costa, 2008) and are less accurate in speech production (Roberts et al., 2002). However, a sizable number of studies have reported on the benefits of growing up bilingual, as well as of using two languages daily as an adult, both from behavioral and cognitive perspectives.

Two areas where bilingual language skills show a significant advantage over monolingualism are the executive function and the theory of mind. To start with the former, the executive function includes the processes of attention, selection, inhibition, monitoring, and flexibility (Bialystok & Barac, 2013). Specifically, research has focused on the relationship between bilingualism and inhibitory control, which is the ability to dismiss distractions while performing a task (Bialystok & Senman, 2004). In this area, the effect of bilingualism is visible already in pre-verbal infants: Kovacs and Mehler (2009), in an eye-tracking study with early simultaneous bilingual infants of 7 months of age, found that they outperformed matched monolinguals on cognitive control abilities.

Next, bilingual children tend to perform better or faster when confronted with a task where they need to choose between conflicting representations of the object without being distracted by deceitful perceptual cues (Bialystok & Senman, 2004). The task does not need to be necessarily verbal: for instance, Bialystok and Senman (2004) found that 4- to 5-year-old bilingual children performed better than monolinguals of the same age on a card sorting task.

The bilingualism effect shown here tends to be explained by the fact that bilinguals need to inhibit one language while switching to the other language in speech production.

Furthermore, simultaneous acquisition of two languages in early childhood might propel the development of the ability to take another person's perspective earlier (Bialystok & Senman, 2004). This has to do with the theory of mind (ToM), an ability to consider the mental states and ascribe beliefs to others that are different from one's own (Premack & Woodruff, 1978), which develops in children at about four years of age (Bialystok & Senman, 2004). The positive effect of bilingualism on the ToM advancement has been shown for various sets of languages: Goetz (2003) assessed ToM and perspective-taking on English monolingual, Chinese monolingual, and Chinese-English bilingual children and showed that bilinguals performed significantly better than the monolingual children at the first testing time. Then, Farhadian and colleagues (2010) tested Kurdish-Persian bilingual and Persian monolingual preschool children on false-belief tasks and found that bilinguals had more advanced ToM compared to monolingual children of the same age. This chapter therefore concludes that there is no cognitive disadvantage that is associated with early bilingual development, and conversely, there is certain cognitive advantage of growing up with two languages.

2.4. Developmental language disorder

As noted by Norbury and colleagues (2016), "approximately two children in every class of 30 pupils will experience language disorder severe enough to hinder academic progress." When children display considerable language difficulties that are not caused by intellectual disability, neurological damage, or impaired hearing, the scientific community uses the term developmental language disorder, or DLD (Leonard, 2014a). This chapter introduces DLD, discusses theoretical issues associated with it, provides an overview of symptoms across languages, and finally, introduces the most common assessment methods.

2.4.1 Terminological complexity

In general terms, developmental language disorder is a deficit in linguistic development and abilities of children that is not associated with any other developmental disorder (C. Norbury et al., 2008). DLD, however, is not the only title utilized in the field for this condition. Historically, authors have relied on abundant terminology to describe a language delay in children. The terms such as developmental aphasia or dysphasia (Weiner, 1969), developmental language impairment (Wolfus, 1980), and delayed language (Weiner, 1974) were used, among others. Norbury and colleagues in the introduction to *Understanding Developmental Language Disorders in Children* (2008) mention that a lack of agreement in the field is a "consequence of continuing uncertainties as to how best to conceptualize children's problems". In essence, diagnosticians and researchers have been varying in opinion on whether the said language difficulties should be considered independently from or as part of impairments beyond the language domain.

Stemming from this dilemma, the discrepancy between general cognitive development and verbal abilities has been the primary factor for establishing appropriate terminology. Thus, a language impairment may be regarded as *specific,* or disconnected from nonverbal functioning. In contrast, it may be *non-specific* when language and nonverbal abilities are both low (C. Norbury et al., 2008). Therefore, the term *specific language impairment* (SLI) has been the most widely-used term for describing language problems in children who show a deficit in language ability in the absence of significant limitations in nonverbal IQ or neurological impairment (C. Norbury et al., 2008). However, as Norbury and colleagues (2016) mention, it is often problematic to establish boundaries between the SLI and the non-specific impairment (NLI) because for a *specific* language impairment to be diagnosed, the nonverbal IQ scores have to be within the norm. However, many of the children requiring speech-language therapy tend to exhibit some nonverbal variation, such as impaired auditory memory, working memory, and executive function skills (Baird, 2008; C. Norbury et al., 2008). This means that if the nonverbal IQ criterion was implemented, such children could be denied clinical services as they would not fit the *specific* category (see C. F. Norbury et al., 2016 for discussion). In addition, as noted by Norbury et al. (2016), the 5th revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), which uses a generic term *language disorder*, no longer includes the nonverbal IQ criterion (American Psychiatric Association & American Psychiatric Association, 2013).

In this light, a number of authors and official statements (e.g., ICAN/RCSLT, 2018; National Institute of Deafness and Other Communication Disorders (NIDCD), 2021) prefer the term DLD, standing for *developmental language disorder.* While DLD is used for a language deficit that occurs without any other major disability, it is a more encompassing term that does not exclude the possibility that affected children might have a slight deficit outside of the language domain e.g., poor motor skill (Hill, 2001) or procedural learning impairment (Tomblin et al., 2007) that may or may not be directly associated with a language disorder. Thus, the notion of DLD is inclusive towards the children who might be difficult to place in either specific or non-specific impairment categories (C. F. Norbury et al., 2016).

Since the present work is concerned with the English-Russian bilingual context, it might be relevant to mention how the concept of DLD accepted in Europe and the United States could be mapped conceptually and terminologically on to the Russian context. There is a clinical tradition in Russia for DLD diagnosis, although, conceivably, the terms SLI and DLD are not employed. Instead, a similar developmental deficit is referred to as the *General Underdevelopment of Speech (GUS)* (in Russian: общее недоразвитие речи, ОНР) (Tomas et al., 2019). The so-called underdevelopment is defined as a complex developmental disorder of speech and language in children with normal intellectual abilities and hearing, which affects a broad spectrum of language competencies (Levina, 1968).

In the Russian tradition, GUS is subject to a psycho-pedagogical classification that recognizes four levels of speech development in children, where 1 is the lowest level (no speech). The assumption is that both TD and children with an impairment go through the four stages; however, whereas unimpaired children overcome language difficulties with age, children with GUS lag behind unless they receive assistance (Tomas et al., 2019). To make another note with regard to DLD in the Russian clinical tradition, standardized instruments for DLD assessment have not been available until very recently. DLD assessment has mostly relied on qualitative measures, such as specialist observation (Lopukhina A. et al., 2019). It is only in the recent years that comprehensive Russian-language assessment tools supported by quantitative analysis and normative data started appearing, such as RuCLAB (Lopukhina A. et al., 2019; *RuCLAB (Russian Child Language Assessment Battery)*, 2023).

To sum up, it is evident that language disorders can be conceptualized in a number of ways depending on the clinical traditions as well as on the constituents or "symptoms" that an impairment is believed to include. For the present work, the term DLD has been chosen to refer to a delay without obvious links to other disorders; however, the term SLI may occur in citations of other authors. Also, the present work does not focus on the Russian tradition of the assessment of language disorders as much as on the Russian language in the context of bilingualism, so where translation was necessary, DLD was used instead of the term accepted in the Russian literature on the subject.

2.4.2. DLD: Behavioral and linguistic effects

DLD is one of the most common developmental disorders, with 7.6 % of 4–5-year-old children reported for England in 2016 (C. F. Norbury et al., 2016). Children diagnosed with DLD acquire language and reach linguistic milestones later than typically developing (TD) children. Furthermore, DLD may affect emotional health and result in higher anxiety and depression risk (Conti-Ramsden & Botting, 2008) as well as in the inability to form good quality friendships (Durkin & Conti-Ramsden, 2007). In addition, the difficulties related to poor command of speech and language may persist in adult life and result in the lack of success in higher education or employment (Johnson et al., 2010). Ultimately, the scope of DLD is significantly larger than to only encompass language skills: affected children present a vulnerable group for learning disability and unsuccessful social adaptation both in school years and later in life.

When it comes to language abilities, DLD affects many linguistic processing levels to a varying degree (Leonard, 2014a). Studies show that one of the major difficulties for children with language deficits is morphosyntax. To illustrate, they may have difficulties with the productive application of both inflectional and derivational morphological inventory (e.g., Carlisle, 1988; Moats & Smith, 1992). Children with DLD learning various languages struggle with grammatical morphemes, such as case and tense markers as well as clitics (e.g., English: Rice & Wexler, 1996; van der Lely & Ullman, 2001; Czech: Smolík & Vávrů, 2014; Russian: Tribushinina & Dubinkina, 2012), auxiliary verbs and subject-verb agreement (Leonard, 1995). Furthermore, children with DLD show lexical knowledge and processing issues, or simply put, they utter first words later, and it is difficult for them to learn and understand new words (Trauner et al., 2000). They also exhibit a less deep knowledge of word meanings as well as a more limited word stock (McGregor et al., 2013).

To add a note, speakers of structurally different languages tend to exhibit symptoms that depend on a particular linguistic system. Therefore, the typology of the native language may predetermine the areas that could be problematic for children. For example, in Romance languages, the verbs are inflected, and the inflection system tends to be highly transparent. As reviewed by Leonard (2014b), children with DLD learning Italian or Spanish do not exhibit serious deficits in the tense and agreement inflections like English children do; however, Romance speakers show difficulties with the use of unstressed direct object pronouns that precede the verb instead.

2.4.3. DLD assessment tools

To assess DLD, a comprehensive evaluation including case history and the assessment of nonverbal cognition and hearing is conducted. In addition, language production and comprehension skills are assessed using both spontaneous measures and formal tests (Thordardottir, 2015). At present, a variety of tests is used for the assessment of morphosyntactic skills, such as picture description tasks testing subject-verb agreement, case-marking, or the use of clitics (Armon-Lotem et al., 2015a). In recent years, researchers have been paying attention to sentence repetition (SRep), and nonword repetition tasks (NWR) as these proved themselves especially accurate in detecting DLD in monolingual children (Armon-Lotem et al., 2015a). Sentence repetition involves listening to sentences and repeating them verbatim. This simple task is a powerful assessment tool because in order to repeat a sentence, participants have to be able to analyze it in terms of multiple levels of representation, such as lexical representations, syntax, morpho-syntax, and semantics. Then, the participants have to rely on their long-term memory storage in order to retrieve the representations needed to repeat the stimulus (see Sentence Repetition in Armon-Lotem et al., 2015b). Thus, sentence repetition successfully reveals morphosyntactic difficulties, which are the first key to DLD diagnosis.

However, SRep coupled with other instruments relying on morphosyntax might not be the best tool for assessing bilingual children, especially younger ones. Haman and colleagues explain that "vocabulary size and processing speed can become confounding variables when diagnosticians attempt to disentangle bilingualism from SLI at the lexical level" (Haman et al., 2015). In addition, Thordardottir (2015) showed that the amount of exposure to each language affects grammatical and vocabulary development. Because bilingual children tend to be exposed to one of the languages to a greater extent, language skills could be distributed unequally across the two languages. Thus, tasks like sentence repetition could put a bilingual child at a disadvantage by requiring them to use forms not yet acquired or not yet applied systematically in one of the languages.

In contrast, nonword repetition tasks rely on morphosyntax to a lesser degree. For NWR, children are asked to repeat nonsense words modeled in accordance with the phonology of their native language. In recent years, NWR has recommended itself as a reliable tool for assessing monolinguals speaking a number of languages, for example, English: (e.g., Bishop et al., 1996; Weismer et al., 2000), Italian (Casalini et al., 2007), Spanish (Girbau & Schwartz, 2007), Russian (Kavitskaya et al., 2011), and Hebrew (Armon-Lotem & Meir, 2016). Thus, a host of studies across typologically different languages show a similar pattern of monolingual children with DLD performing significantly poorer than children with typical development when assessed with a NWR task.

However, to establish NWR as a sensitive[1] assessment instrument for bilinguals, more research is needed. While some studies reported inconsistent accuracy (e.g., Kohnert et al., 2006), others report reliable results (e.g., Armon-Lotem & Meir, 2016). Nevertheless, if administered properly (considering the specifics of target languages and bilingual norms), NWR tasks offer a massive advantage over other tasks in testing bilingual children because the NWR task is not affected by the amount of language exposure and transfer as much as other tasks. In other words, the task is mostly insensitive to the size of the lexicon (but see 2.5.4 *Wordlikeness, Prosody, and Morphological Complexity* below for discussion) or familiarity with specific morphosyntactic constructions, so TD bilingual children who are unfamiliar with certain vocabulary or grammatic representations can be expected to show better results than bilingual children with DLD (Thordardottir, 2015).

2.5. The nonword repetition task

Research has shown that spoken word recognition heavily depends on the number and nature of words activated in the memory (e.g., Rispens et al., 2015). Specifically, the frequency of a word's occurrence in a language as well as the frequency of the occurrence of a combination of sounds affect the speed and accuracy of word recognition in TD children versus children with DLD. Thus, knowing that children's inability to repeat nonwords could signal a language impairment, it is necessary to describe the task in more detail. A host of factors may contribute to the task's being more or less accurate in assessing the phonological working

[1] Armon-Lotem and Meir (2016, p. 2) define *sensitivity* as 'the ability of a test to classify correctly an individual with a disorder (i.e., the percentage of children with SLI).'

memory, morphological processing, and other domains. The following chapter continues the overview of the NWR task and discusses important considerations on its use in bilingual children.

2.5.1. NWR task design

While the developmental, linguistic, and educational effects of DLD are sufficiently clear, the precise mechanism of the impairment is less so. Gathercole and Baddeley (1990) proposed that the underlying causes of DLD involve a deficit in the phonological working memory or the phonological loop. Thus, originally, NWR tasks were conceived as measurements of the phonological working memory capacity (Gathercole & Baddeley, 1989). However, many studies since then (e.g., Archibald & Gathercole, 2006) have shown that apart from the working memory, the performance on NWR tasks depends on language-specific variables, such as the phonological complexity of the items or their similarity to actual words.

Overall, the language-specificity of NWR means that the task requires more language-processing mechanisms to be active. A holistic look on NWR reveals that these processes include the reception of the acoustic signal, the transformation of the signal into a phonological representation, and its consequent storage in the working memory, aided by other skills involved in planning and executing the response (summarized in Graf Estes et al., 2007). Therefore, any of the skills involved in receiving, processing, encoding, and producing a word form may affect performance.

Thus, if a child has difficulty repeating a nonword, the problem may have arisen in the working memory as well in the lexical knowledge or output processing (Archibald & Gathercole, 2006). A carefully constructed NWR task can thus tax specific processing mechanisms, revealing more issues than just a phonological deficit. Among the most common design variables are item length, segmental complexity, wordlikeness, morphological structure, phonotactic probability, and prosody. Combined, these variables can be used to make NWR tasks that rely more on the lexical knowledge or that exclude it altogether, tasks that strain phonological processing and output processing, and tasks that can be said to rely predominantly on the phonological working memory. These variables are presented below, along with the studies showing how their interplay probes into various language processing skills.

2.5.2. Item length

The first design factor, item length, is linked with one of the most common interpretations of poor NWR performance in children with DLD, namely the limitations in the phonological working memory. Graf Estes and colleagues (2007) showed in a meta-analysis that even one-syllable items could differentiate between DLD and TD groups. The explanation lies in phonological processing: to repeat a nonword, one needs to store an unfamiliar phonological sequence in the working memory (Baddeley et al., 1998). Thus, longer items require more processing than shorter items and tend to be harder to repeat, which means that if children with DLD have a deficit in the phonological working memory capacity, they will consistently fail to repeat longer nonwords.

Length effects have been tested on children learning various languages, such as English (e.g., Archibald & Gathercole, 2006), Hebrew and Russian (Armon-Lotem & Chiat, 2012). Both TD and children with DLD tend to perform poorer on items with three or more syllables, and the DLD group tends to show significantly poorer results. Interestingly, children speaking different languages tend to show disparate results on nonwords longer than five syllables. For instance, Armon-Lotem & Chiat (2012) found that five-syllable items were extremely difficult for TD Russian-Hebrew bilingual children. However, Gathercole et al. (Gathercole et al., 1994) observed the opposite tendency: children found them easier to repeat than four-syllable items. The authors concluded that the effect could be caused by the greater degree of familiarity of children with the morphological constituents of such nonwords.

2.5.3. Articulatory complexity

The next variable that proved essential in NWR is articulatory complexity, which can be altered by adding a consonant cluster or late-acquired phonemes (Chiat, 2015). A common finding is that nonwords containing a consonant cluster are harder to repeat both for children with DLD and their typically developing peers (Graf Estes et al., 2007). Moreover, with some languages, articulatory complexity might be affected even more by the placement of the cluster word-initially, word-medially, or word-finally. For instance, Armon-Lotem & Chiat (2012) found that for TD bilingual Russian-Hebrew-speaking children, segmental complexity could be achieved by adding a cluster word-medially, even though many studies focus on the initial or final cluster.

Another common finding is that although children with DLD are not expected to have an articulatory deficit (i.e., a clinically identified disorder) (see Chiat, 2015)., they tend to perform significantly more poorly on nonwords with a cluster, compared to their TD peers (e.g., Archibald & Gathercole, 2006). While there is no definitive answer as to why this happens, an explanation might be offered, in part, by poor "speech motor output skills." Archibald & Gathercole (2006), referencing Goffman (2004), summarize that children with DLD "have difficulty producing well-organized and stable rhythmic speech motor movements, which may affect their ability to repeat nonwords."

Furthermore, a number of studies have shown the relationship between phonological working memory and oral motor processing. For example, Reuterskiöld and Grigos (2015) observe, citing a study by Bishop and colleagues (1990), that the retention of unfamiliar lexical forms for oral repetition is supported by overt or covert articulation. This strategy might not be available to individuals with speech impairment, which, in turn, results in difficulty remembering nonwords. Moreover, in their study Reuterskiöld and Grigos (2015) showed that additional task demands (such as length) put a strain on articulatory control.

2.5.4. Wordlikeness, prosody, and morphological complexity

Wordlikeness is defined as a subjective rating of a nonword's resemblance to actual words of a given language done by a native speaker (Chiat, 2015). Nevertheless, wordlikeness depends on objective constituents, such as morphological complexity and prosody. The more nonwords draw on existing morphology and typical prosodic patterns, the higher the degree of wordlikeness, and vice-versa (Chiat, 2015). Studies show that both children with typical development and DLD perform better on highly wordlike items compared to low-wordlike ones (e.g., Gathercole, 1995). To better study this phenomenon, one could look separately at the parameters of prosody and morphological complexity.

First, the prosodic structure concerns the organization of stressed and unstressed syllables in a nonword. If the prosodic structure follows the patterns of a particular language, it renders nonwords more wordlike, and vice versa. Chiat (2015) discusses the effects of prosodic structure, citing studies in English (Chiat & Roy, 2007) and Swedish (Sahlen, Christina Reuterskiold-Wagn, 1999) and notes that items with the atypical syllable structure are likely to be more difficult to repeat. Conceivably, typical prosodic patterns tend to vary across languages. For example, Kavitskaya et al. (2011) found that the atypical absence of an onset renders syllable structure more difficult for Russian-speaking TD children and their peers with DLD. Thus, the VC syllable pattern is more difficult than CVC, and VCC is more difficult than CVCC for Russian-speaking children. It is thus essential for the NWR task design to consider typical and atypical syllable patterns and stress placement in the target language as these variables could render the task either unnecessarily difficult or easy.

Another constituent of wordlikeness is morphological complexity. Regarding morphology, there are two main considerations for nonword design: the first is whether existing morphemes of the target language should be used in the design, and the second is whether the morphemes included should be inflectional, derivational, or both. Several studies have shown that the presence of inflectional and derivational morphology affects children’s performance. For example, Caramazza and colleagues (1988) found that nonwords created using existing inflectional morphology in Italian required more processing time, even though processing occurred in the absence of meaning. Then, Cilibrasi et al. (2019), testing sensitivity to inflectional morphemes in English, corroborated these findings. Caramazza et al. conclude that this result indicates that lexical representations are morphologically decomposed, i.e., that individual morphemes can be stored separately. Although these studies do not feature a repetition task, they are relevant to the present purposes in showing that a single morpheme can be recognized and retrieved from long-term memory during operations with nonwords, which has implications for NWR.

Furthermore, studies investigated the involvement of derivational morphology with nonword repetition. For example, Casalini et al. (2007) used derivational morphemes for a repetition task in Italian to construct two nonword sets, where one contained “morphological” nonwords made up of existing roots and existing affixes in illegal combinations, and the other one did not rely on existing morphology. The results showed that children affected by DLD relied on verbal representations similarly to their TD peers: nonwords containing existing morphology were easier for both groups to repeat.

It follows from the above that the inclusion of realistic prosody and morphological complexity can produce highly wordlike items that could affect children's performance. Specifically, nonwords relying on derivational morphemes could render the task easier for both TD and DLD groups. However, whether this is the desired effect is debatable. For example, Gathercole (1995) suggested that low-wordlike nonwords are a more accurate measure of PWM compared to highly wordlike items, because when children repeat nonwords that draw on existing roots and affixes, they have an opportunity to rely on verbal representations stored in the long-term memory and children's PWM would be supported by the lexical knowledge. On the other hand, as studies (Archibald & Gathercole, 2006; Casalini et al., 2007) have shown, highly wordlike items are able to initiate and measure LTM facilitation or the lexical effect, which could reveal a lexical processing deficit (Casalini et al., 2007).

To sum up, there is no straightforward answer as to the desired degree of wordlikeness in a NWR task, and many of today's measures operate with it differently. For instance, the CNRep (Gathercole et al., 1994) and the NRT (Dollaghan & Campbell, 1998) which have been shown to produce the largest effect sizes and which have been used by several research groups, vary in wordlikeness: while the former uses both wordlike and non-wordlike items, the latter includes only low-wordlike items (Archibald & Gathercole, 2006). Furthermore, the factors discussed above are skewed towards monolingual children: when bilingualism is involved, NWR design variables might show different effects, which will be discussed further in the chapter dedicated to the performance expectations.

2.5.5. Phonotactic probability

Phonotactic probability (PP) refers to "the frequency of occurrence across a language that a segment (phoneme) or a sequence of segments occurs in a given position within a word" (Metsala & Chisholm, 2010, p. 491). In other words, high-probability sequences of phonemes are those sequences that can be found in many real words (Munson et al., 2005). Research has demonstrated that children repeat high-probability phoneme sequences better than the low-probability ones. For example, Edwards et al. (2004) found that 3- to 8-year-old children demonstrated greater accuracy and fluency repeating high-frequency versus low-frequency two-phoneme sequences within multisyllabic words. Moreover, research has shown that the PP effects decline with age: older TD children have fewer problems with low PP stimuli compared to younger children performing a NWR task (Edwards et al., 2004; Rispens et al., 2015).

The decline in the PP effect in older children is associated with the growing size of the lexicon. To elaborate, Rispens et al. (2015) explain that the PP effect declines with age and is less visible in 7–8-year-olds due to the lexical facilitation effect: as the vocabulary becomes more developed, lexical associations grow stronger and facilitate spoken language processing in school-age children. Naturally, if PP is linked to the vocabulary size, then the effect of PP would be greater in children who have a smaller vocabulary, such as children affected by DLD. Munson et al. (2005) have shown this by testing the PP effects on 3 groups of children, namely children with DLD, age-matched peers, and vocabulary-matched controls. As a result, the size of vocabulary proved to be the strongest predictor of the difference of repetition accuracy between sequences with high and low PP.

These findings can be applied to the matters of bilingualism as well: since bilingual children tend to have a less developed lexicon in the weaker language, the PP effects should be particularly visible when bilingual children are tested in the non-dominant language. This idea is confirmed in Messer et al. (2010) who administered a nonword recall task to the groups of Turkish-Dutch bilinguals and their monolingual peers. Conceivably, bilingual children showed a greater PP effect in their dominant language compared with their monolingual Dutch peers. These findings, coupled with the findings on the PP effect in children affected by DLD, show that in experimental tasks, such as the nonword recall or nonword repetition, the results of the test may be affected by the phonotactic probability. Since the investigation into the PP effects is not one of the objectives of the present work, it is essential for the purposes of the present nonword repetition task to maximally reduce the variation in phonotactic probability. Therefore, in order to reduce the variation in PP values for the stimuli, the decision was made to design items with equally high phonotactic probability. The calculations for this measure are presented in chapter *Phonotactic Probability Calculations in Russian and English*.

2.5.6. Neighborhood density

Previous research has shown that the phonological similarity of the stimulus to other words may significantly influence spoken word recognition and production. To operationalize phonological similarity, the term phonological similarity neighborhoods is utilized and is further divided into neighborhood density and neighborhood frequency. Neighborhood density (ND) is conventionally interpreted as “the number of words differing from the target word by a one phoneme substitution, addition, or deletion” (Metsala & Chisholm, 2010), whereas neighborhood frequency “refers to the frequencies of occurrence of the neighbors” (Goldinger et al., 1989, p. 502). With regard to ND, an item may be either from sparse neighborhood (also referred to as low density), meaning that there are few other phonetically similar words – or it can be of high density (referred to as dense neighborhood), meaning that there are many similar sounding items.

Items with high ND that cause increased phonological overlap have been found to either facilitate or inhibit language processing depending on the nature of the task. Marian et al. (2008) note that in monolingual studies, phonological similarity may depend on whether the task engages lexical or sub-lexical processing. The authors summarize it as “while competition effects during auditory word recognition have been localized to the lexical level, facilitation effects during recognition and production have been localized to a pre-lexical phonological level” (Marian et al., 2008, p. 142). To illustrate the former, the authors refer to a study by Slowiaczek & Hamburger (1992) that investigated the ND effect on a shadowing task. During the task, inhibition was observed for stimuli that were preceded by high density primes, such as *blast-black*, but not for nonwords. The effect suggests that the competition between similar-sounding words was observed only on a lexical level.

Therefore, to generalize the ND effect in monolingual studies, a common finding is that in lexical-level tasks, recognizing high-density words takes more time than recognizing low-density words (Rispens et al., 2015). This phenomenon is conventionally explained by the competition of targets with other similar sounding words during lexical activation. Similar to the phonotactic probability effects described in previous chapter, ND effects are closely tied to vocabulary size and tend to fluctuate with a child’s age. To exemplify, kindergarten-age children are affected by the ND effect to a lesser degree than are 7-8-year-old children whose vocabularies are more developed (Rispens et al., 2015).

Conceivably, bilingual children are sensitive to phonological similarity effects as well. In bilinguals, phonological similarity has been found to either inhibit or facilitate language processing, depending on such factors as language dominance and proficiency level (Marian et al., 2008). A comprehensive study by Marian et al. (2008) provides an overview of within-language and cross-linguistic ND effects on the performance of bilinguals. The authors observe that phonological similarity influences word retrieval in both dominant- and non-dominant-language naming tasks. Further they note that the "efficiency of retrieval" is particularly sensitive to word similarity in non-dominant contexts, which means that ND effects occur asymmetrically (Marian et al., 2008, p. 165). In addition, bilinguals were found to co-activate a network of similar-sounding words in both languages when performing a task in one language only.

An interesting finding in this regard comes from a recent study by Arutiunian and Lopukhina (2020) who examined cross-linguistic effects in 4-6-year-old monolingual children and adults. The study claims that the presence of the facilitation or inhibition of the performance on stimuli with varying ND measures depends on the morphological system of the given language and may vary in languages with rich inflectional system (such as Russian or Spanish) from languages with few inflections (English). The study is in line with the suggestion previously made by Vitevitch & Stamer (2006). Potentially, this might have implications on bilinguals whose languages belong to different morphological systems (such as high-inflection versus low-inflection languages like Russian and English). However, these implications require further studies.

To relate these findings to the present task, the most important observation here is that one can expect not only within-language activation of the network of similar sounding words, but also a cross-linguistic activation during a nonword repetition task. Since the NWR task operates with pre-lexical or sub-lexical processing, a facilitative effect can be expected for dense neighborhood stimuli. However, just like it has been done with phonotactic probability above, the decision was made to minimize variation in ND effect by creating a list of uniform items. Therefore, all non-morphological stimuli in the present set have been designed accounting for the uniformity of ND values. As a result, all nonwords in both English and Russian non-morphological nonword sets are extremely low in ND ($n<1$). The morphological set is more difficult to control for ND due to its impracticality: since the stimuli contain some of the earliest acquired phonemes in the language, the presence of phonological neighbors cannot be ruled out for the whole set. The calculations for ND measures are presented in chapter 3.1. *The construction of bilingual English-Russian NWR task (EN-RU NWR).*

2.6. DLD and bilingualism: assessment specifics

As studies have shown, acquiring more than one language in childhood is not a risk factor for DLD, and neither does it increase the severity of the symptoms. A number of studies have found that bilingual children with DLD are comparable in scope of the symptoms to monolingual children with DLD in different languages (e.g., Swedish-Finnish: Paradis et al., 2003; French-English: Westman et al., 2008). Additionally, the disorder occurs across various language pairs (e.g., Russian-Hebrew: Armon-Lotem & Meir, 2016; French-Greek: Stavrakaki et al., 2011). Unsurprisingly, the bilingual context puts additional strain on assessment tools and methods for a number of reasons that will be discussed in the following chapter. The following chapter adds to the discussion of DLD and bilingualism and concludes the literature overview.

2.6.1. The monolingual bias

First of all, DLD assessment in bilingual children is often more complicated because most tests were developed and normed for monolingual children and cannot be applied to bilinguals because there is a risk of receiving results that would not be entirely reliable (Armon-Lotem & Meir, 2016). This risk comes from the fact that in bilinguals, language skills tend to be distributed unequally between the two languages. In one of the languages, grammatical structures or the lexicon might be acquired faster and better than in the other, sometimes to such an extent that the linguistic behavior of a bilingual child might resemble an impairment (Armon-Lotem & Jong, 2015). Therefore, bilinguals with typical language development are at risk of performing below expectations for their age on a test conducted in the weaker language (Armon-Lotem et al., 2015a). Successively, the results of such tests could be validated by applying monolingual norms, which could place the child within the DLD group, when in reality, the test reflected a child's limited lexical and phonological knowledge caused by lack of exposure. (Armon-Lotem & Chiat 2012) Thus, a major complication standing before parents and clinicians is the need to properly disentangle healthy bilingual development from a case of DLD.

2.6.2. Testing both languages

One way to distinguish DLD from normal language acquisition in bilinguals is to conduct tests in both of the languages that a child is acquiring because when the symptoms of DLD are present, they are expected to appear in both languages (Thordardottir, 2015). The particular symptoms, however, tend to depend on specific languages. For example, Leonard (2014a) summarizes that French-English bilingual children with DLD show difficulty with French direct object pronouns but have no difficulty with English direct object pronouns. Similarly, Spanish-English bilingual children with DLD tend to omit verb inflections in English but keep the inflections when speaking Spanish. The implication is that the languages that are being acquired play an important role in screening for the vulnerabilities that may reveal DLD.

Another reason why conducting the tests in both languages is vital is because some aspects of a child's language behavior might be due to a language transfer rather than a language deficiency (Armon-Lotem & Jong, 2015). As mentioned in previous sections, the amount of exposure to a language plays a big role in language acquisition: constructions of one language might be acquired faster and with more fidelity, which might result in one language becoming dominant. Consequently, the dominant language may impose its morphosyntactic or grammatic structures on the weaker language. Thus, in a bilingual context, some aspects traditionally taken as markers of DLD, such as the omission or substitution of case marking, articles, clitics, verbal inflections, etc., discussed in the previous chapter, might actually be the result of the influence of the other language. For example, Russian-Hebrew bilingual children often omit the definite article in Hebrew, which could be explained by the fact that Russian does not have articles (Armon-Lotem & Jong, 2015). This phenomenon has implications for DLD assessment, according to Armon-Lotem and Jong:

> Some grammatical morphemes, that are vulnerable in SLI as well as in typical bilingual development, are less suitable for assessment in bilinguals, or are crucially dependent on bilingual norms (Armon-Lotem & Jong, 2015, p. 9).

The matters are complicated even more by the fact that it is not obvious which morpho-phonological knowledge has been acquired by bilingual children, making traditional assumptions of NWR test design lose ground. For example, the advantage of wordlike nonwords (i.e., the fact that they are easier for both TD and DLD groups) (Graf Estes et al., 2007) might fail with sequential bilinguals, because they might have insufficient exposure to some language aspects and thus fail to retrieve them from the long-term memory. It is also important to mention here that due to variant language exposure, bilingual children might perform significantly below their monolingual peers in the weaker language (Kohnert et al., 2006). Nevertheless, not all studies converge in finding this effect: for example, Chiat (2015) discusses a range of studies showing the opposite, for example, Lee & Gorman (2013) found that groups of bilingual children speaking Korean, Chinese, and Spanish as their L1, were comparable to their monolingual peers.

To return to the design of bilingual assessments, the bilingual context could put a strain on many aspects of task design, including the degree of wordlikeness, the choice of morphological information, and the presence and placement of articulatorily complex material. Moreover, a broader overview reveals that available standardized tools cover a limited number of languages, and when they are available, the tests might not have a counterpart in the second language for bilingual assessment. In addition, the symptoms of DLD must be differentiated from normal bilingual acquisition, which means that the assessment of bilingual DLD calls for a test controlling for a number of factors to avoid over- or under-diagnosis. However complex bilingual assessment is, the present study attempted to use the cumulated knowledge about language impairment and bilingualism to present the NWR task for English-Russian bilingual children.

3. Methodology

3.1. The construction of bilingual English-Russian NWR task (EN-RU NWR)

Drawing upon the theoretical overview provided above, a bilingual English-Russian NWR task (EN-RU NWR) was constructed. The following chapter presents the task, its key parameters, performance expectations, and hypotheses regarding TD bilingual children, as well as predictions for the bilingual DLD performance.

3.1.1. Key design features

The present task was constructed to reveal limitations in the phonological working memory, phonological processing, and lexical processing by means of varying item length, articulatory complexity, and wordlikeness, respectively (with phonotactic probability and neighborhood density being controlled). Two parallel sets of 24 language-specific items were created for each language for a total of 48 stimuli. The items in both language subsets conformed to the phonotactic constraints of the English and Russian languages, including vowel reduction in unstressed syllables.

3.1.2. Item length

The item length varied from 2 to 4 syllables, with 8 items per group. Items longer than four syllables were excluded because they yielded inconclusive results in previous studies (Armon-Lotem & Chiat, 2012; Gathercole et al., 1994). Similarly, one-syllable-long items were not included because they tend to produce insignificant effect-size (Graf Estes et al., 2007). Based on the studies described above, failure to repeat longer items is associated with the phonological loop deficit, whereas poor performance on shorter items can be associated with other deficits.

3.1.3. Articulatory complexity

Articulatory complexity was achieved by adding a consonant cluster. Both English and Russian permit various cluster positions, so initial and medial clusters were added. The clusters were distributed equally among the two positions. In addition, since the absence of onset might add complexity to syllable structure in Russian (Kavitskaya et al., 2011), the VC type of syllable was not used word-initially in either Russian or English to avoid unnecessary complexity. All the items conformed to the typical prosodic structure of the languages. As far as the performance expectancy goes, poor performance on items containing a phonological cluster has been associated with deficient phonological processing.

3.1.4. Wordlikeness

Regarding wordlikeness, the nonwords in the present test are divided into two groups, *morphological nonwords* (MNW) and *non-morphological nonwords* (NNW). The design was inspired by the NWR test by Casalini et al. (2007), who used a similar grouping for monolingual subjects. Thus, the first set contains 12 MNW including an existing root or stem and an existing derivational suffix in a combination that does not occur in either language, for example, [fun] and [ly] making up *funly.* Only early-acquired productive derivational suffixes were chosen for the present purposes. In addition, the suffixes are salient and expected to be recognized by the children. The other 12 items were non-morphological nonwords (NNW) that did not feature any existing morphemes or sub-lexical units in either language.

For MNW in the English subset, the adjectival suffix -*ful* (full of) and the adverbial suffix -*ly* (characteristic of) were used. In Russian, the included affixes were the diminutive suffixes -*ек* [jek] and -*ок* [ok], and an adjectival suffix -*ов* [ov]. All of these affixes are supposed to be familiar to 7–8-year-old pupils. The roots and stems for MNW were taken from books for children, namely Merriam-Webster English Dictionary (2015), Macmillan English Dictionary (2001), and The Orthographic Dictionary of The Russian Language for Pupils (2007).

As follows from the discussion above, NNW have been designed to prevent any lexical support so that they would provide a clearer measure of the phonological working memory. In contrast, The MNW were constructed to be highly language-specific, so TD children would be able to recognize their parts and rely on lexical processing. Since the vocabulary is controlled, bilingual children might be expected to be familiar with these items in both languages and thus receive lexical support equally to monolinguals (as was the case of Engel de Abreu, 2011 after controlling for vocabulary). Thus, the present study hypothesizes that with the present factors controlled for, the bilingual lexical effect should be minimized, i.e., TD children can be expected to perform equally well in both dominant and weak languages.

Nevertheless, it should be noted that the present test cannot control for all the aspects, such as the actual knowledge of the target lexical phonology of bilingual children being tested, so the expectancy of the bilingual effect cannot be ruled out completely.

TD bilingual children are thus expected to perform well on both MNW and NNW, with the MNW set being easier in the dominant language because of the lexical support and the familiarity with the lexical phonology and morphology. In its turn, failure to repeat the items might signal a number of factors, for example:

1. Failure to repeat MNW: possible limitations in LTM retrieval or limited lexical knowledge.
2. Failure to repeat NNW, but successful repetition of MNW: a possible phonological loop deficit or phonological processing deficit (the subject needs to rely on LTM storage during repetition).
3. Failure to repeat both MNW and NNW: limited lexical knowledge and a possible phonological loop deficit or phonological processing deficit.

3.1.5. Phonotactic probability in EN-RU NWR task

As explained above, phonotactic probability introduces a variable that may affect the results of the test, since the accuracy of nonword repetition may differ for high-PP and low-PP sequences of phonemes (see chapters 2.5.5. *Phonotactic Probability* and 2.5.6. *Neighborhood Density* above). Therefore, in the current set-up all nonwords within a language-specific set should have comparable PP values, i.e., the nonwords must not deviate significantly from one another in the PP measures. To create such lists of stimuli, measures for phonotactic probability for the lists of English and Russian items were obtained with the help of computational instruments. For the English set, the web-based Phonotactic Probability Calculator (PPC) by Vitevitch & Luce (2004) was used. To estimate phonotactic probability, the instrument relies on two measures, namely positional segment frequency or "how often a particular segment occurs in a certain position in a word" (2004, p. 482) and biphone frequency or "segment-to-segment co-occurrence probability of sounds within a word" (2004, p. 482).

For the Russian set, no web-based instrument comparable to PPC is currently available, so the Phonological Corpus Tools (PCT) software (Hall et al., 2019) was used. In order to collect the linguistic data, the OpenCorpora corpus of the Russian language (1.989.538 tokens)[2] was used as a source for creating a corpus-based dictionary that was uploaded into the PCT software. Since the original corpus contained no phonological data, it was deemed necessary to modify it to allow for individual phoneme distinction. The most efficient way to add the necessary phonological data was to treat orthographic signs as phonemes and delimit each individual phoneme with a graphic sign, such as the dot to create combinations such as к.о.т. (English equivalent of c.a.t.).

Since Russian orthography is predominantly phonemic with a high phoneme-to-grapheme correspondence rate, this approach is practically possible. However, it has limitations because Russian phonology tolerates positional vowel and consonant alternations. For example, Kerek & Niemi (Kerek & Niemi, 2009, p. 5) note that as it comes to consonants, "the most typical consonant alternations in Russian are progressive assimilations of voiced or unvoiced obstruents", for instance when [d] turns into [t] in the word *boat,* лодка so that [lodka] becomes [lotka] due to progressive assimilation. One thus needs to be aware of the fact that such subtleties cannot be accounted for when threating orthographic corpus data as phonemic data. Nonetheless, knowing about this complication, it is possible to develop stimuli that minimize or eliminate the alternation factor. One way to achieve this is by manually selecting for the nonwords that reduce lenis consonants in the environment where assimilation is expected. Therefore, for the reasons of practicality, the grapheme-to-phoneme correspondence approach was used in accordance with the manual post-hoc selection of the eligible stimuli.

3.1.6. Phonotactic probability calculation for English and Russian

Since variation in PP may influence the difficulty of the test and thus affect the results, the nonwords in both Russian and English sets are required to have comparable degrees of phonotactic probability within the individual set. Morphological nonwords (i.e., the nonwords that are made with an existing root or stem and an existing derivational suffix in a combination that does not occur in either language) fulfil this requirement by definition

[2] Available at http://opencorpora.org

because they are constructed with existing morphemes only, so their PP values are high. To reiterate the explanation from the section above, these items are constructed with early-acquired and salient derivational affixes familiar to children, so no further PP selection is deemed necessary.

In turn, for the non-morphological stimuli (i.e., those that do not include any existing morphemes or sub-lexical units), biphone sequences were analyzed using the software described above. The biphone frequencies (i.e., the sums of the various biphones in each word) were then analyzed to measure the deviation of individual items from the mean (English *M*=0.019; Russian *M*=0.005[3]). The reasonable precision selection criterion for the finalized sets is based on the three-sigma rule (3σ rule): only those items that do not deviate from the mean value by more than three SD were selected (3σ English = 0.027; 3σ Russian = 0,006). The figures for each stimulus are presented in Table 1 and Table 2.

Stimulus	**Biphone frequency**	**Individual deviation** *(Mean – biphone frequency)*
fupeton /ˈfju:ptən/	0.030	-0.011
thoppin /ˈθɔ:pin/	0.005	0.014
plopeck /ˈplɔ:pek/	0.011	0.008
fadip /fædɪp/	0.010	0.009
tumpilit /ˈtʌmpɪlɪt/	0.022	-0.003
tamifol /ˈtæmifɒl/	0.014	0.005
thopilan /ˈθɔ:pɪlən/	0.026	-0.007
fleetidel /ˈfli:tɪdel/	0.021	-0.002
duntimolap /dʌnˈtɪmoləp/	0.034	-0.015
pifabimon /pɪfæbɪmən/	0.026	-0.007
talifisop /tæˈlɪfisəp/	0.022	-0.003
blylifiton /blaɪˈlɪfɪtən/	0.025	-0.006

Table 1. Phonotactic probability measures for NNW English stimuli

[3] Present calculations show that in total Russian stimuli have a lower value of biphone frequency than the English stimuli. The reason for this effect might be in the different tools used for the calculations. While both PPC (Vitevitch & Luce, 2004) and PCT (Hall et al., 2019) use the Vitevitch & Luce algorithm, the results they yield for identical nonwords and real words may differ: during the calculation it was observed that the PCT tends to produce lower phonotactic probabilities for identical items. Since the present work does not have as its objective a deeper investigation into the PP effects and since these calculations allow us to control for significant variation among the items in two sets of stimuli, the calculations provided here are considered sufficient.

Stimulus	Biphone frequency	Individual deviation *(Mean – biphone frequency)*
фи́тла	0.004	0.001
за́пун	0.009	-0.004
дво́си	0.005	0.001
хи́лис	0.005	0.000
тулама́	0.004	0.001
такледу́к	0.008	-0.003
митило́н	0.005	0.000
стуриму́т	0.005	0.000
трукидило́т	0.004	0.001
тирилиди́п	0.003	0.002
липанисо́т	0.006	-0.001
жулитуло́н	0.004	0.001

Table 2. Phonotactic probability measures for NNW Russian stimuli

3.1.7. Neighborhood density calculation for English and Russian

To determine ND, the one-phoneme metric known as the Hamming or Levenshtein distance was used (Vitevitch & Luce, 2016). Following this method, to define neighbors of a target word, the number of real words that differs from each nonword by a single phoneme substitution, addition, or deletion in any position of the target word was counted. For the English set, the calculations were performed using the Similarity Neighborhood tool by Vitevitch & Luce (2004). As for Russian, ND measures were calculated using a Qt 6.2-based tool designed specifically for this project[4]. The tool operationalizes the OpenCorpora corpus of the Russian language to provide a list of words fulfilling the given criteria.

The limitation that comes with the usage of the Russian tool is the same that was described previously for the phonotactic probability measures, namely the absence of phonological data. For the purposes of the current measures, the orthographic sign was treated as the phoneme, which Russian phonology allows for with a number of restrictions described in chapter 3.1.5 *Phonotactic probability.*

The main objective of this step both in Russian and English was to prepare two lists of stimuli with sparse neighborhoods ($n<1$). The motivations behind this decision are described

[4] The tool was designed in collaboration with independent researchers and can be accessed at: https://github.com/IsakovAD/lang_xml_parser

above in chapters 2.5.5 *Phonotactic probability* and 2.5.6. *Neighborhood density* and, generally, lie in the need to minimize variation in the ND effect. The ND measures were applied only to non-morphological stimuli because, as previously stated, the morphological set contains some of the earliest acquired phonemes in the language, and thus the absence of phonological neighbors cannot be guaranteed for the whole set. The final list of nonwords is shown in Table 3 and Table 4 and in the appendix.

WORDLIKE NW				NON-WORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**горбóк /gɐr‘bok/**	**письмóвый /pʲɪsʲ‘movɨj/**	**лесѝстовый /lʲɪ‘sʲistɐvɨj/**	MIDDLE CLUSTER	**фѝтла /‘fʲitlɐ/**	**такледýк /tɐklʲɪ‘duk/**	**тириндилѝп /tʲɪrʲɪndʲɪlʲ‘ip/**
INITIAL CLUSTER	**дрýгик /‘drugʲɪk/**	**глазóвый /glɐ‘zovɨj/**	**внучáтовый /vnʊ‘tɕætəvɨj/**	INITIAL CLUSTER	**двóси /‘dvosʲɪ/**	**стуримýт /stʊrʲɪ‘mut/**	**трукидилóт /trʊkʲɪdʲɪ‘lot/**
NO CLUSTER	**бáлик /‘balʲɪk/**	**городѝк /gərɐ‘dʲik/**	**сеновалóк /sʲɪnəvɐ‘lok/**	NO CLUSTER	**зáпун /‘zapʊn/**	**туламá /tʊlɐ‘ma/**	**липанисóт /lʲɪpɐnʲɪ‘sot/**
NO CLUSTER	**дýбик /‘dubʲɪk/**	**пирóжик /pʲɪ‘roʐɨk/**	**золотóвый /zələ‘tovɨj/**	NO CLUSTER	**хѝлис /‘xʲilʲ ɪs/**	**митилóн /mʲɪtʲɪ‘lon/**	**жулитулóн /ʐʊlʲɪtʊ‘lon/**

Table 3. Russian nonword list

WORDLIKE NW				NONWORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**cavely /ˈkeɪvlɪ/**	**parrotful /ˈpærətful/**	**companily /ˈkʌmpənɪlɪ/**	MIDDLE CLUSTER	**fupeton /ˈfju:ptən/**	**tumpilit /ˈtʌmpɪlɪt/**	**duntimolap /dʌnˈtɪmoləp/**
INITIAL CLUSTER	**dryful /ˈdraɪful/**	**sparrowly /ˈspærəulɪ/**	**grocerily /ˈgrəusərɪlɪ/**	INITIAL CLUSTER	**plopek /ˈplɔ:pek/**	**flitidel /ˈfliːtɪdel/**	**blylifiton /blaɪˈlɪfɪtən/**
NO CLUSTER	**wayly /ˈweɪlɪ/**	**babyful /ˈbeɪbɪful/**	**diarily /ˈdaɪərɪlɪ/**	NO CLUSTER	**fadip /fædɪp/**	**tamifol /ˈtæmifɒl/**	**pifabimon /pifæbimən/**
NO CLUSTER	**beely /ˈbiːlɪ/**	**teddiful /ˈtedɪful/**	**familiful /ˈfæmɪlɪful/**	NO CLUSTER	**thoppin /ˈθɔ:pin/**	**thopilan /ˈθɔ:pɪlən/**	**talifisop /tæˈlifisəp/**

Table 4. English nonword list

3.2. Case study performance expectations

3.2.1. Background tests

As the first step of the present case study, it was necessary to conduct a series of background tests that measure cognitive and linguistic ability of the participant. The cognitive test acts as a screening measure to exclude the participants who might not be suitable for the study, while linguistic tests are included with the purpose of acting as predictors of the performance on the main task (the bilingual EN-RU NWR task). These measures are described in detail below.

3.2.1.a. Raven's CPM

Raven's Coloured Progressive Matrices (CPM) are a set of tests that measure nonverbal intelligence in children and are used here as a background measure of general cognitive ability. It is a simpler version of Raven's Progressive Matrices (RPM) which is, essentially, a visual task of abstract reasoning (Raven & Court, 1938). They comprise three sets of problems with progressively rising difficulty levels. During test administration, the participants are asked to examine a set of pieces with one missing element and then select the best match for the missing piece. On the technical side, Raven's progressive matrices measure the so-called "eductive" ability, or using the author's own words, the "meaning-making ability", and can assess "the person's capacity at the time of the test to apprehend meaningless figures […], see the relations between them […], and develop a systematic method of reasoning" (Raven, J. C., Court, J. H., & Raven J., 1983). The CPM tests were designed specifically for the assessment of intellectual processes of children and less-able adults and are widely used for these purposes (John & Raven, 2003), which confirms the suitability of this measure in the study.

Presently, the CPM are included as a screening measure with the aim to detect a potential cognitive deficit as a variable by comparing the participant's scores to the age-appropriate normative scores. It is expected that the participant selected for the study passes the CPM within the age-appropriate percentile. In detail, the link between cognitive development and DLD has been described earlier in chapter 2.4.1. *Terminological complexity*. In simplified terms, for DLD to be diagnosed as such, the child is expected to show normal cognitive development in the presence of subpar or abnormal language development. Therefore, it is expected that the child performs within the norm on the general cognitive ability tests for the definition of DLD that the current work operates with.

3.2.1.b. YARC

York Assessment of Reading for Comprehension – Passage Reading (YARC) is used to assess the accuracy, rate, and comprehension of oral reading skills in primary school children. The test itself invites the participants to read aloud short texts and answer eight comprehension questions. In technical terms, YARC assesses the following components of reading: decoding skills (reading accuracy), fluency (reading rate), and text comprehension (literal and inferential meaning) (Snowling et al., 2012). Many studies converge in finding that reading skills, and particularly reading rate, is a good indicator of a learner's language proficiency (see for example, Cilibrasi, Adani, et al., 2019; Gráf et al., 2023). In the present work, adequate reading skills attested with the use of YARC act as a predictor of the successful performance on the bilingual EN-RU NWR task. In turn, in clinical assessment environments poor reading skills may be associated with the presence of DLD.

To explain further, many studies indicate that a considerable number of children with DLD have reading difficulties, such as poor reading comprehension and reading decoding skills (see e.g. Baird, 2008; Bishop et al., 2009). The link between the two is assumed to lie in the phonological processing skills that are involved in reading and that are assumed to be significantly affected in children with DLD (as is discussed in the earlier chapter 2.3. *Cognition and bilingualism*). The reading decoding skills are highly dependent on phonological abilities, especially in languages with the alphabetic writing system, such as English and Russian. According to the YARC manual, "phonological processing skills underpin the development of word-level decoding skills and, in particular, 'phonic' reading strategies" (Snowling et al., 2012, p. 1). To simplify it further, to be able to read such languages as English or Russian, one must know the letters, learn to identify the speech sounds represented by these letters, and then learn to match symbols to sounds (Moats, 2019). Therefore, adequate phonological processing skills are essential for what we broadly call reading comprehension skills.

To return to the use of YARC in the present study, the test is expected to provide a sufficient assessment of reading comprehension skills to act as the predictor of adequate phonological processing skills in participants. Specifically, if a child scores within the norm on YARC, then he or she is expected to successfully pass the English part of the bilingual EN-RU NWR task. If a child with a high YARC score fails the English part of the bilingual NWR task, it could indicate issues with the test design.

3.2.1.c. TROG

Test for Reception of Grammar - Version 2 (TROG-2) is used in the present study to assess the grammatical knowledge of the participants. It is a multiple choice test that asks the participants to match the sentences they hear to one one of the four pictures presented to them. The test is designed to assess understanding of grammatical contrasts in English and it enables one to conduct both quantitative studies (to attest how a person's comprehension compares to that of other people of the same age) and qualitative studies (to identify specific areas of difficulty). The test is appropriate for children and is popular in research and clinical use (Bishop, 2003).

As is mentioned in previous chapters, grammar (both syntax and morphology) are a common indicator of DLD. Such issues as erroneous subject-verb agreement, omissions or substitution of inflectional morphemes, and incorrect case marking are prominent DLD markers in a number of languages. Therefore, tests aiming to provide a comprehensive assessment of a child's grammatical skills may indicate the potential presence of a language impairment (see chapters 1-5 in Armon-Lotem et al., 2015 for the examples of such tests used with bilingual children).

In this study, performance on the test of grammar is used as a predictor of performance on the main NWR task with the contention that normal scores on TROG should ideally correspond to successful performance on the bilingual En-Ru NWR task. In turn, the results showing high performance on TROG and lower-than-expected performance on bilingual NWR might indicate test design issues. It is also worth mentioning here that the proposed EN-RU NWR task was not designed to assess the children's knowledge of specific grammatical structures; therefore, the findings gained on TROG are not expected to correlate with specific variables included in the proposed bilingual NWR task. TROG results are expected to show whether the participant's linguistic knowledge is adequate for the recognition of grammatical patterns used in the construction of morphological nonwords.

3.2.1.d. CNRep

The Children's Test of Nonword Repetition (CNRep) is the last performance predictor in the current study. It is a NWR task widely used in DLD studies. The test consists of 40 stimuli that range in length from two to five syllables. Some of the stimuli contain a consonant cluster, some contain a weak syllable with a reduced vowel, and many contain lexical morphemes (Gathercole et al., 1994). During test administration, the children are asked to repeat nonwords, and their repetition attempts are scored immediately as either phonologically correct or incorrect. In cases where a regional accent, or other individual consistent pronunciation effects might influence pronunciation, allowances are made and the answers are scored as correct.

The mechanics of NWR tasks are mentioned above in chapter 2.4.3. *DLD Assessment tools*, while the chapter 2.5. *The nonword repetition task* describes the variables present in NWR tasks, including CNRep. Also, observations on how such tests are assumed to act as a measure of phonological processing are presented there. It is important to note here that CNRep results have been shown to be related to other linguistic abilities, such as vocabulary acquisition, reading, and spoken language comprehension in eight-year-old children (Gathercole et al., 1992). Specifically, CNRep has been found to have an association with grammar proficiency, and SNRep scores are highly correlated with the TROG measures in young children (Gathercole et al., 1994). Therefore, coupled with TROG and YARC, CNRep is expected to show whether the participant's performance on the standardized tests is adequate compared to the performance on the proposed bilingual NWR task. It is expected that if the participant passes CNRep within the expected norm, he or she will be able to pass the proposed NWR and will show a similar pattern of errors in terms of item length and phonological clusters in both tests.

3.2.2. Performance expectations for the bilingual EN-RU NWR

As was previously described, the test was developed to study the interplay of three domains of spoken language processing, namely the phonological processing, short-term-memory processing, and long-term memory (lexical) processing. The assumptions are simplified in that more deficits may occur at once and show a less clear-cut dependence between the particular stimulus and the deficit. During current experiment, however, significant issues with NWR by TD children are not expected to occur. To sum up test construction and hypotheses, the participant is expected to show good performance on all items of his or her stronger language. As shown by previous studies, the participant can perform more poorly on longer nonwords and on nonwords containing a consonant cluster. Also, the child is expected to benefit from the facilitative effect of MNW, especially in the dominant language, and perform better on them rather than on the items that contain no morphological cues.

In the weak language, bilingual participants are expected to show similar item length and consonant cluster effects. However, even though the present test design attempted to minimize the bilingual exposure effect, the children might not contribute from MNW and treat them equally to NNW: the facilitative effect of morphology can be weaker in the weaker language as the children might lack the knowledge of some morphemes. All in all, in the actual clinical environment, the test should be able to reveal the language impairment and successively localize the domain of the difficulty if the child performs poorly on both the dominant and the weak languages. Finally, it is necessary to note that the following experiment aims at assessing the accuracy of the present NWR task on a group of TD bilingual children, not at identifying the impairment.

4. Experiment

4.1. Participant

The participant in this case study was one English-Russian bilingual child (7;5) born and raised in the UK. At the time of the study, the participant was attending local school in English in Year 2. In addition, the child was attending Russian school once a week. A sociolinguistic questionnaire confirmed the language status of the child as well as providing background on his L1 and L2 exposure: the child had been exposed to both languages from birth and had been using both languages on a daily basis in a number of social settings. Additionally, the parental questionnaire revealed the child's inclination towards English (English dominance).

The child had not been previously diagnosed with having a language impairment. However, the parental questionnaire revealed that the child had a certain degree of language difficulties in Russian related to the pronunciation and differentiation of the following consonants: /ʂ/ vs /ɕː/ vs /ʐ/, /r/, and vowels /i/ vs /ɨ/ (in Russian: ш/щ/ж, р and и/ы). The questionnaire also noted that the participant had difficulties with spelling in both languages. The child had never been diagnosed with a cognitive delay or hearing difficulties. Prior to the study, parental consent was received, and the study was approved by the university ethics committee for studies involving schoolchildren.

4.2. Assessment

4.2.1. Telepractice administration

The assessment procedure was conducted online, following the best practice recommendations of remote assessment, or telepractice. Telepractice is a reliable assessment model used by such organizations as ASHA (American Speech-Language-Hearing Association)[5] and Pearson Assessments that enables the administration of speech and

[5] See for example: American Speech-Language-Hearing Association (ASHA). (2021). Considerations for speech, language, and cognitive assessment via telepractice. https://www.asha.org/practice-portal/professional-issues/telepractice/#collapse_1

language testing aimed for the purposes of assessment and/or treatment of various impairments. Thus, the assessment was conducted in the participant's home in a quiet room, with the examinee's parent taking part as the on-site facilitator.

The experimental setup consisted of the laptop with the camera, speakers, and microphone on both the examiner's and examinee's side. The second computer screen was used to record and score the materials, where online scoring was required. The audiovisual information required for the testing was shared during an online meeting via the Google Meet platform. Testing was performed over two sessions, each taking 40-45 minutes. In one session, TROG and the bilingual English-Russian NWR task were administered. In the second session, CNRep, YARC, and Raven's CPM were tested. Sessions were conducted two weeks apart. Some of the tasks (bilingual English-Russian NWR, TROG, and CNRep) were pre-recorded by native speakers of English, while others (YARC and Raven's CPM) were conducted live by the author of the present work, who is not a native speaker of English. The standardized assessments were administered following the instructions in the respective manuals. For the bilingual English-Russian NWR task the details of the assessment are provided below.

4.2.2. Specifics of the bilingual EN-RU NWR administration

Since the assessment of the structural validity of the bilingual English-Russian NWR task was the main interest of the present study, more attention is paid to the description of this task. First of all, the British English version of the test was administered in order reflect the geographical location and linguistic preferences of the examinee. The test was presented via a Power Point presentation created specifically for the purposes of this study. Apart from the auditory stimuli, the presentation contained visual stimuli that had been designed with the purpose of maximizing children's engagement and attention throughout the study. Such visual stimuli are often used in other child-oriented NWR tasks, for example, the LITMUS Crosslinguistic Nonword Repetition test (COST Action IS0804 Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment) that presents the children with a bead game, where the examinees are building up a necklace as they progress in the assessment.

Similarly in terms of the child-friendly design, the bilingual English-Russian NWR task was presented using a "Fishtalk" game developed to accompany the auditory materials. During the game, the child was invited to look at an image of a marine habitat that initially appeared empty. As the child repeated three stimuli, a new object would appear in the sea (thus equaling to 16 various objects for each series of three stimuli repeated). The child was thus asked to populate the sea by repeating "fishwords" or the imaginary words of the sea inhabitants. Selected visual materials are included in the appendix (2). Before the task itself, practice items were administered to make sure that the child understood the task. The examinee's responses were recorded. The recordings were transcribed and scored offline after the assessment.

4.3. Results

4.3.1. Raven's CPM

The results of the Raven's CPM are presented in the Table 5. The percentile rank of the participant's raw score on three sets of tests (32) is 96. This score is within the percentile class 95.5 – 99.5, which corresponds to the IQ class = 125 – 135 (central value = 130); the performance can thus be considered "very high". In terms of the error pattern, the participant made 4 errors in total, where 75% of errors are figure repetition, and 25% are inadequate differentiation errors. In summary, the examinee successfully passed the background IQ test, which confirms that the results of the subsequent tests cannot be attributed to a cognitive delay.

Percentile	IQ Class range	IQ Category	Z score
96 (95.5 – 99.5) 4%	125 — 135 (130)	Very high	1.65

Table 5. CPM Results

4.3.2. YARC

YARC results are presented in Table 6. For this assessment the standard score was selected as the measure of reading skills, with the following result: accuracy = 94, reading rate = 103, and comprehension = 110. This measure expresses a child's performance in relation to the spread of scores obtained by a sample of children of the same age. YARC standard scores have an average of 100 with SD = 15. According to YARC authors, approximately 96% of all standard scores fall between 70 and 130 (Snowling et al., 2012). Then, according to the YARC manual, a pupil with a standard score of 115 is good reader, and one with 125 is an excellent reader (Snowling et al., 2012).

It must be noted that, first, the cutoff points used above are calculated for monolingual children, and second, they can be debated even in some monolingual scenarios. With these cutoff points being accepted, the examinee's performance can be classified as follows: accuracy – below average (standard score = 94; percentile rank = 34), but no reading problem is indicated; reading rate is average (standard score = 103; percentile rank = 58), while comprehension is above average (standard score = 110; percentile rank = 75). The assessment thus places the examinee within the expected range for a pupil with average reading ability.

Name:	Participant 1			Age at assessment:	7:05	
Passages read:	Level 1 A, Level 2 A					
Summary of scores						
	Ability score	Standard score	Percentile rank	Age equivalent		
Accuracy	36	94	34	6:09		
Reading Rate	51	103	58	7:07		
Comprehen sion	55	110	75	8:10		
Analysis of reading errors						
	Mispronunciations	Substutitions	Refusals	Additions	Ommissions	Reversals
Total error type	3	4	1	1	9	1
% of total errors	15.8%	21.1%	5.3%	5.3%	47.4%	5.3%
Analysis of comprenhension questions						
	Cohesive	Elaborative inference	Evaluative	Knowledge-based inference	Literal information	Vocabulary dependent
Number of questions	3	1	1	3	7	1
Percentage correct	66.7%	100.0%	100.0%	33.3%	100.0%	0.0%

Table 6. YARC Results

4.3.3. TROG2

4.3.3.a. TROG2 quantitative interpretation

For the present assessment the TROG2 version of the test was used. Quantitative results of TROG2 are presented in Table 7 and are as follows: the participant successfully passed 12 blocks, which converts to 30th percentile (standard score 92), with age equivalent 6;6. The interpretation of the scores offered by the TROG2 manual states that a percentile of 50 is average for a child's age; thus, the examinee's score is below average. A note must be taken, however, that this is not a cutoff point for a language impairment: many clinicians regard a score that is one SD below the mean as an indicator of an impairment, which corresponds to 16th percentile or less (Bishop, 2003). Based on this principle, the examinee shows below-average performance without indication of an impairment. Another note of caution must be considered: similarly to the YARC test and the other language tests used, the normative sample and the following norms take into account monolingual development.

Block	Grammar	Correct Sequence	Pass/Fail
A	Two elements	13 4314	13 4314 P
B	Negative	3424	3424 P
C	Relative in, and, on	2332	2332 P
D	Three elements	2314	2314 P
E	Relative SVO	4212	4212 P
F	Four elements	2344	2344 P
G	Relative clause in subject	2134	2134 P
H	Not only X but also Y	1444	2444 F
I	Reversible above and below	1233	1133 F
J	Comparative/absolute	1313	1313 P
K	Reversible/passive	3213	3213 P
L	Zero anaphor	4112	4114 F
M	Pronoun gender/number	1232	1332 F
N	Pronoun binding	2212	1212 F
O	Neither X nor Y	2414	4414 F
P	X but not Y	3132	3132 P
Q	Postmodified subject	4143	4143 P
R	Singular/plural inflection	1323	1322 F*
S	Relative clause in object	4134	4134 P
T	Centre-embedded sentence	1124	2313 F
Total blocks passed			12
Standard score			92
Percentile			30
Age equivalent			6:06

Table 7. TROG2 Results

4.3.3.b. TROG2 qualitative interpretation

With qualitative interpretation of the TROG2 results, we can answer the question of which grammatical structures are giving the examinee particular difficulty. Table 7 shows a list of passed and failed blocks along with the corresponding grammatical structures being tested. There are a couple of notes that must be made to the list: first, the block R, despite being considered failed for the quantitative measures, might be considered as a pass in this measure. The failure is suspected to be due to the examinee's lapse of attention and recklessness rather than genuine difficulty with this structure since adequate understanding was subsequently demonstrated after the data for the quantitative analysis had been taken, which is permissible in TROG2 (Bishop, 2003).

Thus, the list of the grammatical structures where the examinee showed poor performance is presented in Table 8. As shown in the table, there are seven types of grammatical errors that were observed in the test: the "not only X but also Y" structure as in *the pencil is not only long but also red,* the reversible above and below as in *the pencil is below the fork*, the zero anaphor as in *the man is looking at the horse and is running,* the pronoun gender and number as in *she is pointing at them* and pronoun binding as in *the man sees that the boy is pointing at him,* the "neither X nor Y" structure as in *the girl is neither pointing nor running*, and the center-embedded sentence as in *the sheep the girl looks at is running*.

Block	Grammar	Correct Sequence	Examinee's sequence
H	Not only X but also Y	1444	2444 F
I	Reversible above and below	1233	1133 F
L	Zero anaphor	4112	4114 F
M	Pronoun gender/number	1232	1332 F
N	Pronoun binding	2212	1212 F
O	Neither X nor Y	2414	4414 F
T	Centre-embedded sentence	1124	2313 F

Table 8. TROG2 Incorrect Items

One of the questions that TROG2 can help to answer is whether the participant's processing memory skills are adequate. Looking at the breakdown of the examinee's errors, one can see that that the examinee is capable of understanding not only single words, but also remembering and combining the meanings of two, three, and four words in a sentence, because the examinee passed the blocks A-G. The items where the examinee made an error (H, I, L, M, N, and O) also require storing and processing of three or four elements. However, one cannot infer from these errors that they necessarily have to do with the processing memory storage as they can also be attributed to a lack of knowledge of these particular grammatical structures. In order to conclude whether thar examinee's processing memory span is adequate, the results obtained on CNRep must be considered (see chapter 4.3.4 CNRep). As demonstrated, CNRep results are in line with this hypothesis, proving that the child's difficulties with TROG can be explained by the knowledge of grammar rather than inadequate processing memory span.

4.3.4. CNRep

The examinee (age 7;5) scored 31 points on the CNRep, which can be converted to either 50th centile (29 points) or 75th centile (33 points) on the CNRep normative data table, which is within the norm. The observed error pattern is the presence of a consonant cluster word-initially, word-medially, or word-finally. The phonological nature of the clusters differs, and there is no particular type of pattern for their constituents. The most difficult error-type is a 5-syllable item with one or more consonant clusters in medial and/or final positions. Next follow 4-syllable items with final CC and 3-syllable items with a cluster, regardless of the placement of the cluster. A breakdown of errors by syllable is presented in Table 9.

Name	Total	Error item 1	Error item 2	Error item 3	Error item 4
Subtotal A (1-20)	16				
Subtotal B (21-40)	15				
Total by syllable	Points				
2 syl	9	ballop			
3 syl	8	brasterer	skitikult		
4 syl	8	contramponist	empliforvent		
5 syl	6	reutterpation	sepretennial	voltularity	versatrationist
Total correct	31				

Table 9. CNRep Errors

4.3.5. Bilingual EN-RU NWR Task

Since the bilingual English-Russian NWR task is the primary test in the current case study, the results will be described with more detail than the results for CNRep. The scoring of this task was conducted per each complete stimulus, i.e., if there is at least one repetition error for an item, the item is treated as incorrect; if there are no errors, the item is assumed to be correct. The stimuli where the participant systematically struggled with specific phonemes which were also confirmed by the parental interview, were not counted as incorrect. For example, the Russian stimuli containing an alveolar trill /r/ as in дру́гик /drugʲɪk/ were scored as pronounced correctly when the participant replaced the phoneme by the alveolar approximant (an English variant of /r/) or by the phoneme /l/, which is a commonly observed difficulty in children learning Russian.

The errors are shown in Table 10 and Table 11, highlighted in red. In summary, the participant achieved the following scores on the two languages of the task: English = 21 (out of 24 total); Russian = 18 (out of 24 total). In the English part of the test, all the errors are located in the non-wordlike (NNW) part of the stimulus list. Also, the majority of the observed errors are in the four-syllable stimuli. The breakdown of the errors is the following:

- fadip /fædɪp/ – two-syllable item; no CC, error type = phoneme substitution;
- duntimolap /dʌnˈtɪmoləp/ – four-syllable item, middle CC, error type = phoneme addition;
- pifabimon /pifæbimən/ – four-syllable item, no CC, error type = phoneme substitution.

In the Russian part of the test, most of the errors are located in the NNW list, similarly to the English part. Two errors are located in the wordlike list (MNW). Most of the errors (4/6) are four-syllable items. The breakdown of the errors is as follows:

- тулама́ /tʊlɐʻma/ – three-syllable item, NNW, no CC, error type = phoneme omission;
- золото́вый /zələʻtovɨj/ – four-syllable item, MNW, no CC, error type = phoneme omission;
- сеновало́к /sʲɪnəvɐʻlok/ – four-syllable item, MNW, no CC, error type = phoneme omission;
- трукидило́т /trʊkʲɪdʲɪʻlot/ – four-syllable item, NNW, initial CC, error type = phoneme addition;
- тириндили́п /tʲɪrʲɪndʲɪlʲʻip/ – four-syllable item, NNW, middle CC, error type = phoneme omission;
- хи́лис /ʻxʲilʲ ɪs/ – two-syllable item, NNW, no CC, error type = phoneme substitution.

	WORDLIKE NW			NONWORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**cavely** **/ˈkeɪvlɪ/**	**parrotful** **/ˈpærətful/**	**companily** **/ˈkʌmpənɪlɪ/**	MIDDLE CLUSTER	**fupeton** **/ˈfju:ptən/**	**tumpilit** **/ˈtʌmpɪlɪt/**	**duntimolap** **/dʌnˈtɪmoləp/**
INITIAL CLUSTER	**dryful** **/ˈdraɪful/**	**sparrowly** **/ˈspærəulɪ/**	**grocerily** **/ˈgrəusərɪlɪ/**	INITIAL CLUSTER	**plopek** **/ˈplɔ:pek/**	**flitidel** **/ˈfliːtɪdel/**	**blylifiton** **/blaɪˈlɪfɪtən/**
NO CLUSTER	**wayly** **/ˈweɪlɪ/**	**babyful** **/ˈbeɪbɪful/**	**diarily** **/ˈdaɪərɪlɪ/**	NO CLUSTER	**fadip** **/fædɪp/**	**tamifol** **/ˈtæmifɒl/**	**pifabimon** **/pifæbimən/**
NO CLUSTER	**beely** **/ˈbiːlɪ/**	**teddiful** **/ˈtedɪful/**	**familiful** **/ˈfæmɪlɪful/**	NO CLUSTER	**thoppin** **/ˈθɔ:pin/**	**thopilan** **/ˈθɔ:pɪlən/**	**talifisop** **/tæˈlifisəp/**

Table 10. Bilingual En-Ru NWR task – English Part Errors

	WORDLIKE NW			NONWORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**горбок** **/gerʻbok/**	**письмо́вый** **/pʲɪsʲʻmovɨj/**	**ле́систовый** **/lʲɪʻsʲistevɨj/**	MIDDLE CLUSTER	**фитла** **/ʻfʲitle/**	**такле́дук** **/teklʲɪʻduk/**	**тиринди́лип** **/tʲɪrʲɪndʲɪlʲʻip/**
INITIAL CLUSTER	**дру́гик** **/ʻdrugʲɪk/**	**глазо́вый** **/gleʻzovɨj/**	**внуча́товый** **/vnʊʻtɕætəvɨj/**	INITIAL CLUSTER	**дво́си** **/ʻdvosʲɪ/**	**стурими́ут** **/stʊrʲɪʻmut/**	**трукиди́лот** **/trʊkʲɪdʲɪʻlot/**
NO CLUSTER	**балик** **/ʻbalʲɪk/**	**городи́к** **/gəreʻdʲik/**	**сенова́лок** **/sʲɪnəveʻlok/**	NO CLUSTER	**за́пун** **/ʻzapʊn/**	**тулама́** **/tʊleʻma/**	**липанисо́т** **/lʲɪpenʲɪʻsot/**
NO CLUSTER	**ду́бик** **/ʻdubʲɪk/**	**пиро́жик** **/pʲɪʻroʐɨk/**	**золото́вый** **/zələʻtovɨj/**	NO CLUSTER	**хи́лис** **/ʻxʲilʲ ɪs/**	**митило́н** **/mʲɪtʲɪʻlon/**	**жулитуло́н** **/ʐʊlʲɪtʊʻlon/**

Table 11. Bilingual En-Ru NWR task – Russian Part Errors

4.3.6. Preliminary results

Before moving on to the NWR task performance result, it is necessary to consider the interplay of the cognitive test, language proficiency test, and the proposed NWR task. First, the child's language proficiency in English was assessed using a test of reading skills, YARC, and a test of grammar, TROG-2. On YARC, the participant scored within the age norm for an average reader. On TROG-2, the child scored below average for a monolingual norm, however, without a severe delay. Coupled together, YARC and TROG demonstrated that the examinee's comprehension of English across the two different modalities (reading and grammar comprehension) is below average, but still adequate for a monolingual norm (i.e., not signaling a language delay). Since the child's performance is adequate for his age, he is expected to be familiar with certain morphological knowledge that children are expected to gain by that age, including the suffixes *-ly* and *-ful* used in the construction of morphological nonwords in the proposed NWR task. Therefore, the participant is expected to show the morphological effect in the NWR task, i.e., items constructed from familiar morphemes should be easier to repeat than the items containing no familiar morphemes.

Then, the participant was tested on the CNRep in order to assess his phonological working memory and provide a standardized benchmark that the proposed NWR task would be compared to. Among all the standardized tasks used, CNRep has the most direct relationship to the studied NWR task because it offers a one-to-one comparison of the participant's behavior on such measures as the cluster effect and item length effect. Therefore, the two tests can be juxtaposed to provide the said comparison. Now, the child's CNRep performance demonstrated the length effect typical for TD monolingual and bilingual children (i.e., the longer the item, the more difficulty for the child to repeat it). Based on this, the participant is expected to show length effect on the proposed NWR task. As it comes to the CC effect, it is not present in the child performance on CNRep (i.e., the child does not find items with a cluster harder to repeat and can repeat them as successfully as items without a cluster). We therefore expect no cluster effect to appear in the results on the proposed NWR task.

5. Discussion

The goal of the current study was to assess the proposed bilingual English-Russian NWR task in order to attest whether it has the potential to detect DLD in bilingual children learning the English-Russian language pair. The study sought to answer the question whether the proposed NWR task can show adequate performance results for a TD bilingual child when evaluated against a set of standardized language ability tests. For the NWR task to be considered adequate, the effects observed on the task are expected to match the effects observed on the other tests in terms of the morphological complexity effect, item length effect, and phonological cluster effect. Also, there should be a visible relationship between the effects observed in L1 (English), and L2 (Russian).

The method selected was a case study with one candidate who was assessed on a set of standardized tests and then on the proposed NWR task. The standardized tests were chosen with the aim to reveal potential dependencies between the child's performance on the proposed NWR task and a specific set of skills: general cognitive development, multi-modal English language proficiency (reading and grammar), and phonological processing and working memory. First, it should be noted that since the candidate's normal intellectual ability was confirmed using Raven's CPM, low scores obtained on some linguistic tasks must be attributed to the linguistic knowledge and not a cognitive delay. Now, the summative results of the other tests can be considered in detail in order to answer the question posed above.

5.1. Cumulative results on three predictors: item length, morphological complexity, and CC

In order to answer the research question, we should summarize the child's performance by the three studied variables: length effect, morphological effect, and the CC effect. To begin with the length effect, longer items were more difficult for the participant on both versions of the bilingual En-Ru NWR task: four-syllable items were the hardest items in Russian and in English. CNRep results match this observation with 5-syllable items being the hardest. Therefore, the tests revealed the presence of the item length effect. As stated before, observed length effects are most likely related to the limitation in phonological working memory, which, to a certain extent, is normal and observed in TD children taking NWR tests. For the participant of this study, normal working memory ability is confirmed by both CNRep and TROG; therefore, normal behavior on the NWR-English is expected. In Russian, a greater item length effect is expected as this language is L2, or the weaker one for the participant, which might explain why the observed effect is larger in this language. Since both English and Russian versions show adequate presence of the length effect, it is possible to conclude that the results of the proposed NWR task match standardized tests on this measure.

Moving on to the morphological complexity effect, morphologically complex stimuli in this NWR task are designed in such a way that they should be perceived as easier to repeat than the nonwords that do not contain any morphological information. This mechanism is explained in the chapters above, but to reiterate, these nonwords contain the morphemes (root + affix) that a child of seven years of age is expected to know in both English and Russian languages. The present study shows this effect in both the English and Russian test versions: in English, all three errors are located in the NNW part of the list, while for Russian, the majority (4/6) of errors are located in the NNW part of the list.

For Russian (L2), the child might be less familiar with the grammatical base of the stimuli (root and/or suffix), so this assumption can explain why the child made two errors in morphologically complex items (possibly, the child did not have the available knowledge to match to the studied stimulus. A test conducted in Russian would be needed to confirm that). For the English (L1) part of test, this result can be matched to an earlier the TROG measure which confirms that the child's knowledge of the English grammar is within the norm and thus acts as a predictor of the adequate performance on a task where age-specific grammar knowledge is expected. It is therefore assumed that the child could rely on the lexical and morphological knowledge stored in the long-term memory when repeating the stimuli, meaning that the familiarity with the parts of the stimuli made the process of nonword repetition easier. To conclude this chapter, the proposed En-Ru NWR task demonstrates expected behavior with regards to the morphological complexity variable.

Finally, the last aspect that remains to be studies is the phonological cluster (CC) effect. With regards to the CC, as has been discussed above, both bilingual and monolingual TD children find it harder to repeat items with a phonological cluster. In the present study, the participant demonstrated a CC effect on CNRep, meaning that the items containing a cluster were harder to repeat than the items without a cluster (8 out 9 errors contain a CC either word-initially, word-medially, or word-finally). However, neither version of the bilingual En-Ru NWR task showed a cluster effect: in English, 1/3 error items contains a cluster, while in Russian, 2/6 error items contain a cluster. This is an interesting observation because, ostensibly, there is a mismatch between the child's performance on CNRep and the proposed task. However, returning to the child's CNRep errors shown in Table 9, it becomes evident that the CNRep items with a CC, where the participant made an error, typically contained more than one cluster: four error items contained 2 clusters, three error items contained 3 clusters, while only one error item contained 1 cluster (*sepretennial*).

It is, therefore, expected that since TD children consistently show difficulty in items with a consonant cluster, the number of clusters might affect the result of a repetition attempt. Consequently, the participant's performance on the proposed bilingual NWR cannot be perfectly matched to CNRep as the latter test uses considerably harder stimuli with more CCs per item. Where there is only one CC in an item, the results obtained on CNRep and the bilingual NWR task are not that far apart: in English there is one error, while in Russian there are two errors in an item that contain only one cluster. It is therefore concluded that the results of both CNRep and the bilingual NWR tests show a relationship, while not direct dependency.

To summarize the results, it is evident that the child's performance on the proposed bilingual NWR task matches the performance on the standardized tests that act as predictors. Therefore, the first question of the study can be answered positively: the proposed task demonstrates adequate performance because its result is dependent on the child' knowledge of grammar, lexical/morphological knowledge, and phonological working memory. In other words, the results obtained from the proposed NWR task with this participant are not distributed at random but behave in a way predicted in the previous chapters of this study. The implications of this result are that the present NWR task has the potential to distinguish DLD in a bilingual child. The study, however, has certain limitations that will be discussed in the chapters below.

5.2. Limitations

As stated above, the present work has demonstrated the given NWR task's potential to distinguish a language impairment in a bilingual child (and conversely whether performance is adequate). However, the study was only conducted as a case study with one participant. The first limitation therefore is the lack of evidence of the test's ability to demonstrate similar results for a larger sample. Thus, in order to make a definitive statement that this test can be used for the purposes of distinguishing DLD from typical language development in bilingual children, a further statistical analysis with a large sample is necessary.

Then, statistical data are essential to answer one of the questions of the present study, which is whether the proposed NWR task has the potential to discriminate between the three different types of variables that could be affected in a child with DLD, namely morphological processing, phonological processing, and working memory. With the case study, the answer to this question can be given only tentatively: from the observation of the child's results, it appears that the test could be able to map language delay on to one of the three variables as there is a clear interplay between predicted performance based on specific competencies and actual results. However, this observation cannot be proved right without statistical data, and thus additional testing is required.

The second limitation concerns the way the two languages were analyzed in this study. As mentioned in chapter 2.6.2. *Testing both languages*, whenever bilingual language development is concerned, it is essential that both languages are assessed. The reasons for that are manifold: first, one should be able to recognize language transfer phenomena and distinguish them from a pathology, and second, one should be able to compare both languages in terms of grammar and lexical knowledge, because in a typical bilingual child, unbalanced development with one language leading and the other lagging is expected, while DLD is known to appear in both languages.

In the present study, the NWR task analyzed both languages the participant speaks, but the standardized tests only assessed English, the dominant language. The data for L2 were only collected via the parental questionnaire. Even though the collected information is reliable, and the NWR performance is in line with it, it cannot certify that the child's lexical and grammatical knowledge corresponds to age norms in L2. This information could be helpful in analyzing the child’s errors on the Russian NWR version as it would supply some insight into the child’s familiarity with the lexical roots and the suffixes used and explain the errors in more detail.

6. Conclusion

The main focus of the present study was to design and test a novel language-specific nonword repetition task (NWR) for bilingual children speaking the English-Russian language pair. The task was created for the assessment of language delay known as developmental language disorder (DLD) in bilingual children in the given language pair. Thus, the present work consists of two main parts: the design of the task and the trial via a qualitative case study with a bilingual child.

As it comes to the construction of the task, the present study has considered a number of cognitive processes that bilingual development involves, such as the interplay of long-term and working memory, phonological development, morphological knowledge, and other phenomena related to the coexistence of two languages in a child's mind. To construct the bilingual language-specific NWR task, variables used in other NWR tasks have been studied and described in the previous chapters. The work proceeded with the three variables: the item length, which is related to the phonological working memory, articulatory complexity, related to phonological processing, and wordlikeness (specifically, morphological complexity) that is related to the knowledge of grammar and long-term memory. These variables informed the final test design that is presented above in the chapter 3.1. and in the appendix (1).

The second part of the study was concerned with the assessment of the novel NWR task in order to ascertain whether the task has the potential to distinguish language delay in a bilingual child. The assessment was conducted via a case study with an English-Russian bilingual with typical language development. A comparison of the child's performance on the proposed NWR task and a series of standardized tests acting as performance predictors was used as the methodology. The case study has demonstrated that that the results obtained by the child on standardized tests that assess specific competencies, such as the knowledge of grammar, reading ability and nonword repetition ability, matched the results obtained on the corresponding parameters of the NWR task. This, in turn, means that the study has confirmed the task's potential to determine the presence of a language delay in a bilingual child speaking English and Russian.

The present thesis also notes on the limitations of this study, being aware of the need to conduct further quantitative research with a larger sample of participants in order to determine the task's reliability and precision. It is also for that reason that the present conclusion does not include a definitive answer whether the variables of item length, morphological complexity and phonological complexity can be used to recognize the source of the child's language difficulty (i.e., whether the impairment is caused by phonological processing, morphological processing, or working memory). Therefore, follow-up studies are suggested to investigate the points above.

In conclusion, the last observation worth making is that the present work has demonstrated how the two languages in a bilingual mind are governed by the principles discussed in the earlier theoretical overview chapters of this thesis. Looking at the case study, one can see how the L1 and L2 exist as two separate and yet coordinated entities: for example, increased length of the stimuli causes errors in L1, but it causes even more errors in L2, so the effect is greater in the non-dominant language. Similarly, non-wordlike (morphologically simple) items are harder for the child than the items that sound like real words in L1, but they are even harder in L2. The implication of this simple observation is the fact that only one of the two languages cannot be selected to act as the informant of the linguistic abilities of a bilingual child as it only shows one side of his or her language ability. As stated earlier by Thordardottir (2015) and Armon-Lotem & Jong (2015) in chapter 2.6.2 *Testing Both Languages* and other recent research in this field, the presence of two languages influences the knowledge of grammar and/or vocabulary aspects of these languages so that they often appear unbalanced or underdeveloped. Because of this factor, there is now a growing interest among researchers in developing specialized tests that are able to test both languages of a bilingual child with equal fidelity and precision. The present study thus contributes to the growing wealth of research that stresses the need for testing bilingual individuals in both of their languages due to the complex interplay of the languages in the bilingual mind.

7. References and sources

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword Repetition: A Comparison of Tests. *Journal of Speech, Language, and Hearing Research*, *49*(5), 970–983. https://doi.org/10.1044/1092-4388(2006/070)

Armon-Lotem, S., & Chiat, S. (2012). *How do Sequential Bilingual Children Perform on Nonword Repetition Tasks?*

Armon-Lotem, S., & Jong, J. de. (2015). Introduction. In S. Armon-Lotem, N. Meir, & J. de Jong, *Assessing multilingual children disentangling bilingualism from language impairment* (pp. 1–24). Multilingual Matters.

Armon-Lotem, S., Jong, J. de, & Meir, N. (Eds.). (2015a). *Assessing multilingual children disentangling bilingualism from language impairment*. Multilingual Matters.

Armon-Lotem, S., Jong, J. de, & Meir, N. (Eds.). (2015b). Sentence Repetition. In *Assessing multilingual children disentangling bilingualism from language impairment*. Multilingual Matters.

Armon-Lotem, S., & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: Evidence from Russian and Hebrew: Accuracy of repetition tasks for identification of SLI in bilingual children. *International Journal of Language & Communication Disorders*, *51*(6), 715–731. https://doi.org/10.1111/1460-6984.12242

Arutiunian, V., & Lopukhina, A. (2020). The effects of phonological neighborhood density in childhood word production and recognition in Russian are opposite to English. *Journal of Child Language*, *47*(6), 1244–1262. https://doi.org/10.1017/S0305000920000112

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158–173. https://doi.org/10.1037/0033-295X.105.1.158

Baird, G. (2008). Assessment and investigation of children with developmental language disorder. In C. F. Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding developmental language disorders in children* (pp. 1–23). Psychology Press.

Baker, C. (2006). *Foundations of bilingual education and bilingualism* (4th ed). Multilingual Matters.

Barke, E. M. (1933). A study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *British Journal of Education Psychology*, *3*, 237–250.

Bialystok, E., & Barac, R. (2013). Cognitive Effects. In F. Grosjean, P. Li, & E. Bialystok, *The psycholinguistics of bilingualism* (pp. 192–206). Wiley-Blackwell/John Wiley & Sons.

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition*, *15*(2), 397–401. https://doi.org/10.1017/S136672891100040X

Bialystok, E., & Senman, L. (2004). Executive Processes in Appearance-Reality Tasks: The Role of Inhibition of Attention and Symbolic Representation. *Child Development*, *75*(2), 562–579. https://doi.org/10.1111/j.1467-8624.2004.00693.x

Bishop, D. V. M. (2003). *Test for reception of grammar: TROG-2: manual* (version 2). Pearson.

Bishop, D. V. M., Brown, B. B., & Robson, J. (1990). The Relationship Between Phoneme Discrimination, Speech Production, and Language Comprehension in Cerebral-Palsied Individuals. *Journal of Speech, Language, and Hearing Research*, *33*(2), 210–219. https://doi.org/10.1044/jshr.3302.210

Bishop, D. V. M., McDonald, D., Bird, S., & Hayiou-Thomas, M. E. (2009). Children Who Read Words Accurately Despite Language Impairment: Who Are They and How Do They Do It? *Child Development*, *80*(2), 593–605. https://doi.org/10.1111/j.1467-8624.2009.01281.x

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword Repetition as a Behavioural Marker for Inherited Language Impairment: Evidence From a Twin Study. *Journal of Child Psychology and Psychiatry*, *37*(4), 391–403. https://doi.org/10.1111/j.1469-7610.1996.tb01420.x

Bloomfield, L. (1984). *Language*. University of Chicago Press.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69. https://doi.org/10.1016/S0010-0277(97)00040-1

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of Early Language Discrimination Abilities in Infants From Bilingual Environments. *Infancy*, *2*(1), 29–49. https://doi.org/10.1207/S15327078IN0201_3

Boysson-Bardies, B. de, Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I., & Arao, F. (1992). Material evidence of selection from the target language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications*. York Press.

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297–332. https://doi.org/10.1016/0010-0277(88)90017-0

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, *9*(3), 247–266. https://doi.org/10.1017/S0142716400007839

Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S., Pecini, C., Roncoli, S., & Burani, C. (2007). Non-Word Repetition in Children with Specific Language Impairment: A Deficit in Phonological Working Memory or in Long-Term Verbal Knowledge? *Cortex*, *43*(6), 769–776. https://doi.org/10.1016/S0010-9452(08)70505-7

Chiat, S. (2015). Non-Word Repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children disentangling bilingualism from language impairment* (pp. 125–151). Multilingual Matters.

Chiat, S., & Roy, P. (2007). The Preschool Repetition Test: An Evaluation of Performance in Typically Developing and Clinically Referred Children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 429–443. https://doi.org/10.1044/1092-4388(2007/030)

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, *1*(2), 215–219. https://doi.org/10.1111/1467-7687.00033

Cilibrasi, L., Adani, F., & Tsimpli, I. (2019). Reading as a Predictor of Complex Syntax. The Case of Relative Clauses. *Frontiers in Psychology*, *10*, 1450. https://doi.org/10.3389/fpsyg.2019.01450

Cilibrasi, L., Stojanovik, V., Riddell, P., & Saddy, D. (2019). Sensitivity to Inflectional Morphemes in the Absence of Meaning: Evidence from a Novel Task. *Journal of Psycholinguistic Research*, *48*(3), 747–767. https://doi.org/10.1007/s10936-019-09629-y

Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *49*(5), 516–525. https://doi.org/10.1111/j.1469-7610.2007.01858.x

COST Action IS0804 Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment. (2019). *LITMUS Crosslinguistic Nonword Repetition Tests*. https://www.bi-sli.org/cl-nonword-repetition

de Groot, A. M. B., de. (2013). Bilingual Memory. In F. Grosjean, P. Li, & E. Bialystok, *The psycholinguistics of bilingualism* (pp. 171–192). Wiley-Blackwell/John Wiley & Sons.

Dollaghan, C., & Campbell, T. F. (1998). Nonword Repetition and Child Language Impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Dorling Kindersley Publishing, Inc, & Merriam-Webster, Inc (Eds.). (2015). *DK Merriam-Webster children's dictionary* (American edition). Dorling Kindersley.

Durkin, K., & Conti-Ramsden, G. (2007). Language, Social Behavior, and the Quality of Friendships in Adolescents With and Without a History of Specific Language Impairment. *Child Development*, *78*(5), 1441–1457. https://doi.org/10.1111/j.1467-8624.2007.01076.x

Edwards, J., Beckman, M. E., & Munson, B. (2004). The Interaction Between Vocabulary Size and Phonotactic Probability Effects on Children's Production Accuracy and Fluency in Nonword Repetition. *Journal of Speech, Language, and Hearing Research*, *47*(2), 421–436. https://doi.org/10.1044/1092-4388(2004/034)

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics*, *16*(3), 513–521. https://doi.org/10.3758/BF03198580

Eimas, P. D., & Miller, J. L. (1980). Discrimination of information for manner of articulation. *Infant Behavior and Development*, *3*, 367–375. https://doi.org/10.1016/S0163-6383(80)80044-0

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech Perception in Infants. *Science*, *171*(3968), 303–306. https://doi.org/10.1126/science.171.3968.303

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Gazanizadand, N., & Kumar, V. (2010). Theory of Mind in Bilingual and Monolingual Preschool Children. *Journal of Psychology*, *1*(1), 39–46. https://doi.org/10.1080/09764224.2010.11885444

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, *20*(1). https://doi.org/10.1111/desc.12427

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*(1), 83–94. https://doi.org/10.3758/BF03210559

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213. https://doi.org/10.1016/0749-596X(89)90044-2

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336–360. https://doi.org/10.1016/0749-596X(90)90004-J

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127. https://doi.org/10.1080/09658219408258940

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

Girbau, D., & Schwartz, R. G. (2007). Non-word repetition in Spanish-speaking children with Specific Language Impairment (SLI). *International Journal of Language & Communication Disorders*, *42*(1), 59–75. https://doi.org/10.1080/13682820600783210

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition*, *6*(1), 1–15. https://doi.org/10.1017/S1366728903001007

Goffman, L. (2004). Kinematic Differentiation of Prosodic Categories in Normal and Disordered Language Development. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1088–1102. https://doi.org/10.1044/1092-4388(2004/081)

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518. https://doi.org/10.1016/0749-596X(89)90009-0

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the Nonword Repetition Performance of Children With and Without Specific Language

Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *50*(1), 177–195. https://doi.org/10.1044/1092-4388(2007/015)

Gráf, T., Huang, L., & Cilibrasi, L. (2023). Oral reading tasks as proficiency indicators: Insights from a learner corpus study. *International Journal of Learner Corpus Research*, *9*(2), 155–179. https://doi.org/10.1075/ijlcr.22013.gra

Grosjean, F. (1997). The bilingual individual. *Interpreting. International Journal of Research and Practice in Interpreting*, *2*(1–2), 163–187. https://doi.org/10.1075/intp.2.1-2.07gro

Grosjean, F. (2012). *Bilingual: Life and reality* (First Harvard University Press paperback edition). Harvard University Press.

Grosjean, F. (2013a). Bilingualism: A Short Introduction. In P. Li & E. Bialystok (Eds.), *The psycholinguistics of bilingualism* (pp. 5–27). Wiley-Blackwell/John Wiley & Sons.

Grosjean, F. (2013b). Spoken Language Processing. In F. Grosjean, P. Li, & E. Bialystok, *The psycholinguistics of bilingualism* (pp. 27–50). Wiley-Blackwell/John Wiley & Sons.

Grosjean, F., Li, P., & Bialystok, E. (2013). *The psycholinguistics of bilingualism*. Wiley-Blackwell/John Wiley & Sons.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. MIT Press.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition*, *29*(3), 503–511.

Hall, K. C., Mackie, J. S., & Lo, R. Y.-H. (2019). *Phonological CorpusTools*: Software for doing phonological analysis on transcribed corpora. *International Journal of Corpus Linguistics*, *24*(4), 522–535. https://doi.org/10.1075/ijcl.18009.hal

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing Cross-Linguistic Lexical Tasks (CLTs) for Bilingual Preschool Children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children disentangling bilingualism from language impairment* (pp. 196–243). Multilingual Matters.

Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36*(2), 149–171. https://doi.org/10.1080/13682820010019874

ICAN/RCSLT. (2018). *Bercow: Ten Years on an Independent Review of Provision for Children and Young People With Speech, Language and Communication Needs in*

England. https://www.bercow10yearson.com/wp-content/uploads/2018/03/337644-ICAN-Bercow-Report-WEB.pdf

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

John, & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Handbook of Nonverbal Assessment* (pp. 223–237). Springer US. https://doi.org/10.1007/978-1-4615-0153-4_11

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-Year Follow-Up of Children With and Without Speech-Language Impairments: Family, Educational, Occupational, and Quality of Life Outcomes. *American Journal of Speech-Language Pathology*, *19*(1), 51–65. https://doi.org/10.1044/1058-0360(2009/08-0083)

Jones, W. R., & Stewart, W. A. C. (1951). Bilingualism and verbal intelligence. *British Journal of Psychology*, *4*, 3–8.

Ju, M., & Luce, P. A. (2004). Falling on Sensitive Ears. Constraints on Bilingual Lexical Activation. *Psychological Science*, *15*(5), 314–318. https://doi.org/10.1111/j.0956-7976.2004.00675.x

Jusczyk, P. W. (1997). *Discovery of spoken language*. A Bradford Book. http://cognet.mit.edu/book/discovery-of-spoken-language

Kavitskaya, D., Babyonyshev, M., Walls, T., & Grigorenko, E. (2011). Investigating the effects of syllable complexity in Russian-speaking children with SLI. *Journal of Child Language*, *38*(5), 979–998. https://doi.org/10.1017/S0305000910000413

Kerek, E., & Niemi, P. (2009). *Russian Orthography and Learning to Read. Linguistics*.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do Language-Based Processing Tasks Separate Children with Language Impairment from Typical Bilinguals? *Learning Disabilities Research and Practice*, *21*(1), 19–29. https://doi.org/10.1111/j.1540-5826.2006.00204.x

Kovacs, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560. https://doi.org/10.1073/pnas.0811323106

Lee, S. A. S., & Gorman, B. K. (2013). Nonword repetition performance and related factors in children representing four linguistic groups. *International Journal of Bilingualism*, *17*(4), 479–495. https://doi.org/10.1177/1367006912438303

Leonard, L. B. (1995). Functional Categories in the Grammars of Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1270–1283. https://doi.org/10.1044/jshr.3806.1270

Leonard, L. B. (2014a). *Children with specific language impairment* (Second edition). The MIT Press.

Leonard, L. B. (2014b). SLI Across Languages. In *Children with specific language impairment* (Second edition, pp. 95–151). The MIT Press.

Levina, R. E. (1968). Характеристика общего недоразвития речи у детей [Describing the general underevelopment of speech and language in children]. In *Основы теории и практики логопедии [Theory and Parctice of Speech-Language Pathology]* (pp. 67–85).

Lopukhina A., Chrabaszcz A, & Khudyakova M. et al. (2019). *Test for assessment of language development in Russian "KORABLIK". In Proceedings of the Satellite of AMLaP Conference "Typical and Atypical Language Development Symposium". [s. l.]*, P. 30.

Macmillan children's dictionary. (2001). Macmillan Heinemann.

Maneva, B., & Genesee, F. (2002). Bilingual babbling: Evidence for language differentiation in dual language acquisition. In B. Skarabela, S. Fish, & A. Do (Eds.), *BUCLD26: The Proceedings of the 26th Boston University Conference on Language Development* (pp. 383–392). Cascadilla Press.

Marian, V., Blumenfeld, H. K., & Boukrina, O. V. (2008). Sensitivity to Phonological Similarity Within and Across Languages. *Journal of Psycholinguistic Research*, *37*(3), 141–170. https://doi.org/10.1007/s10936-007-9064-9

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/S1366728903001068

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth: *Vocabulary deficits*. *International Journal of Language & Communication Disorders*, *48*(3), 307–319. https://doi.org/10.1111/1460-6984.12008

Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-Lambertz, G. (1996). *Coping With Linguistic Diversity: The Infant's Viewpoint*.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143–178. https://doi.org/10.1016/0010-0277(88)90035-2

Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology*, *105*(4), 306–323. https://doi.org/10.1016/j.jecp.2009.12.006

Metsala, J. L., & Chisholm, G. M. (2010). The influence of lexical status and neighborhood density on children's nonword repetition. *Applied Psycholinguistics*, *31*(3), 489–506. https://doi.org/10.1017/S0142716410000081

Moats, L. C., & Smith, C. (1992). Derivational Morphology: Why It Should Be Included in Language Assessment and Instruction. *Language, Speech, and Hearing Services in Schools*, *23*(4), 312–319. https://doi.org/10.1044/0161-1461.2304.312

Montrul, S. (2009). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Benjamins.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness on Nonword Repetitions of Children With and Without Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *48*(5), 1033–1047. https://doi.org/10.1044/1092-4388(2005/072)

National Institute of Deafness and Other Communication Disorders (NIDCD). (2021). *Congressional Justification FY 2022*. National Instutute of Health, Department of Health and Human Services. https://www.nidcd.nih.gov/about/congressional-justification-2022

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756–766. https://doi.org/10.1037/0096-1523.24.3.756

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Norbury, C., Tomblin, J. B., & Bishop, D. V. M. (Eds.). (2008). *Understanding developmental language disorders in children*. Psychology Press.

Novinskaya, N. I. (2007). *Орфографический словарь русского языка для школьников [The Orthographic Dictionary of The Russian Language for Pupils]*. Ripol Classic.

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). French-English Bilingual Children With SLI: How Do They Compare With Their Monolingual Peers? *Journal of Speech, Language, and Hearing Research*, *46*(1), 113–127. https://doi.org/10.1044/1092-4388(2003/009)

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, *121*(2), 130–143. https://doi.org/10.1016/j.bandl.2011.05.003

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Raven, J. C., & Court, J. H. (1938). *Raven’s Progressive Matrices*.

Raven, J. C., Court, J. H., & Raven J. (1983). *Manual for Raven’s Progressive Matrices and Vocabulary Scales (Appendx G).* H.K.Lewis & Co,Ltd.,.

Reuterskiöld, C., & Grigos, M. I. (2015). Nonword Repetition and Speech Motor Control in Children. *BioMed Research International*, *2015*, 1–11. https://doi.org/10.1155/2015/683279

Rice, M. L., & Wexler, K. (1996). Toward Tense as a Clinical Marker of Specific Language Impairment in English-Speaking Children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1239–1257. https://doi.org/10.1044/jshr.3906.1239

Rispens, J., Baker, A., & Duinmeijer, I. (2015). Word Recognition and Nonword Repetition in Children With Language Disorders: The Effects of Neighborhood Density, Lexical Frequency, and Phonotactic Probability. *Journal of Speech, Language, and Hearing Research*, *58*(1), 78–92. https://doi.org/10.1044/2014_JSLHR-L-12-0393

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*(4–6), 635–645. https://doi.org/10.1080/02687030244000220

RuCLAB (Russian Child Language Assessment Battery). (2023). HSE University - RuCLAB (Russian Child Language Assessment Battery). https://www.hse.ru/neuroling/cla_korablik

Sahlen, Christina Reuterskiold-Wagn, B. (1999). Non-word repetition in children with language impairment-pitfalls and possibilities. *International Journal of Language &*

Communication Disorders, *34*(3), 337–352. https://doi.org/10.1080/136828299247441

Singh, L., Loh, D., & Xiao, N. G. (2017). Bilingual Infants Demonstrate Perceptual Flexibility in Phoneme Discrimination but Perceptual Constraint in Face Discrimination. *Frontiers in Psychology*, *8*, 1563. https://doi.org/10.3389/fpsyg.2017.01563

Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1239–1250. https://doi.org/10.1037/0278-7393.18.6.1239

Smolík, F., & Vávrů, P. (2014). Sentence Imitation as a Marker of SLI in Czech: Disproportionate Impairment of Verbs and Clitics. *Journal of Speech, Language, and Hearing Research*, *57*(3), 837–849. https://doi.org/10.1044/2014_JSLHR-L-12-0384

Snowling et al., M. J. (2012). *YARC - York Assessment of Reading for Comprehension*. University of York, Centre for Reading and Language.

Spivey, M. J., & Marian, V. (1999). Cross Talk Between Native and Second Languages: Partial Activation of an Irrelevant Lexicon. *Psychological Science*, *10*(3), 281–284. https://doi.org/10.1111/1467-9280.00151

Stavrakaki, S., Chrysomallis, M.-A., & Petraki, E. (2011). Subject–verb agreement, object clitics and wh-questions in bilingual French–Greek SLI: The case study of a French–Greek-speaking child with SLI. *Clinical Linguistics & Phonetics*, *25*(5), 339–367. https://doi.org/10.3109/02699206.2010.538954

Stille, C. M., Bekolay, T., Blouw, P., & Kröger, B. J. (2020). Modeling the Mental Lexicon as Part of Long-Term and Working Memory and Simulating Lexical Access in a Naming Task Including Semantic and Phonological Cues. *Frontiers in Psychology*, *11*, 1594. https://doi.org/10.3389/fpsyg.2020.01594

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, *259*(5538), 39–41. https://doi.org/10.1038/259039a0

Thordardottir, E. (2015). *Proposed Diagnostic Procedures for Use in Bilingual and Cross-Linguistic Contexts* (S. Armon-Lotem, J. de Jong, & N. Meir, Eds.; pp. 331–359). Multilingual Matters.

Tomas, E., Sevan, D., & Shamaeva, Y. (2019). The Russian Federation. In *Managing children with developmental language disorder: Theory and practice across Europe and beyond* (pp. 398–408). Routledge, Taylor and Francis Group.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural Learning in Adolescents With and Without Specific Language Impairment. *Language Learning and Development*, *3*(4), 269–293. https://doi.org/10.1080/15475440701377477

Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (2000). Neurological and MRI profiles of children with developmental language impairment. *Developmental Medicine & Child Neurology*, *42*(7), 470–475. https://doi.org/10.1017/S0012162200000876

Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology*, *9*(1), 91–96. https://doi.org/10.1037/h0034999

Tribushinina, E., & Dubinkina, E. (2012). Adjective production by Russian-speaking children with specific language impairment. *Clinical Linguistics & Phonetics*, *26*(6), 554–571. https://doi.org/10.3109/02699206.2012.666779

van der Lely, H. K. J., & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, *16*(2–3), 177–217. https://doi.org/10.1080/01690960042000076

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers*, *36*, 481–487.

Vitevitch, M. S., & Luce, P. A. (2016). Phonological Neighborhood Effects in Spoken Word Perception and Production. *Annual Review of Linguistics*, *2*(1), 75–94. https://doi.org/10.1146/annurev-linguistics-030514-124832

Vitevitch, M. S., & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes*, *21*(6), 760–770. https://doi.org/10.1080/01690960500287196

Weiner, P. S. (1969). The Perceptual Level Functioning of Dysphasic Children. *Cortex*, *5*(4), 440–457. https://doi.org/10.1016/S0010-9452(69)80018-3

Weiner, P. S. (1974). A Language-Delayed Child at Adolescence. *Journal of Speech and Hearing Disorders*, *39*(2), 202–212. https://doi.org/10.1044/jshd.3902.202

Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword Repetition Performance in School-Age Children With and Without Language Impairment. *Journal of Speech, Language, and Hearing Research*, *43*(4), 865–878. https://doi.org/10.1044/jslhr.4304.865

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3

Westman, M., Korkman, M., Mickos, A., & Byring, R. (2008). Language profiles of monolingual and bilingual Finnish preschool children at risk for language impairment. *International Journal of Language & Communication Disorders*, *43*(6), 699–711. https://doi.org/10.1080/13682820701839200

Wolfus, B. (1980). Subgroups of developmental language impairment. *Brain and Language*, *10*(1), 152–171. https://doi.org/10.1016/0093-934X(80)90046-2

Yip, V. (2013). Simultaneous Language Acquisition. In P. Li & F. Grosjean, *The psycholinguistics of bilingualism* (pp. 119–137). Wiley-Blackwell/John Wiley & Sons.

Yoshioka, J. G. (1929). A Study of Bilingualism. *The Pedagogical Seminary and Journal of Genetic Psychology*, *36*(3), 473–479. https://doi.org/10.1080/08856559.1929.10532205

8. Resumé

Předkládaná diplomová práce se zabývá návrhem a vyhodnocením nového testu opakování pseudoslov (Eng. nonword repetition task, NWR) schopného detekovat vývojovou jazykovou poruchu řeči (Eng. developmental language disorder, DLD) u anglicko-ruských bilingvních dětí. Mezi odborníky je dobře známo, že rozlišit typický vývoj od vývojové poruchy u bilingvních dětí může být velmi náročné, protože bilingvní děti často mívají některé jazykové potíže v jednom anebo obou svých jazycích, což je u bilingvních dětí normální jev (Armon-Lotem et al., 2015a). Vývojová porucha řeči je častým onemocněním, které postihuje 7 – 10% dětí předškolního věku (Armon-Lotem & Jong, 2015). Proto je hodnocení vývojové poruchy řeči bilingvních dětí často v popředí diskurzu o bilingvní jazykové akvizici. Existuje již řada klinických a výzkumných nástrojů, které se zaměřují na hodnocení jazyka v anglicky mluvící populaci, přičemž oblíbenou metodou hodnocení často jsou testy opakování pseudoslov. Předkládaná práce se pokouší napomoci tomuto výzkumu tím, že představuje jazykově specifický test opakování pseudoslov, který by mohl být potenciálně využit pro hodnocení bilingvních dětí v anglicko-ruském jazykovém páru.

Teoretická část práce představuje problematiku bilingvního osvojování jazyka u dětí a srovnává ji s monolingvním osvojováním jazyka. Přehled začíná raným fonologickým vývojem u monolingvních a bilingvních dětí a pojednává o zrání procesů osvojování jazyka do 12 měsíců věku. Hlavním závěrem těchto kapitol je, že všechny děti, bez ohledu na jazyky, kterými mluví, procházejí v prvních měsících života stejnými klíčovými fázemi, jako je citlivost k rytmickým vlastnostem jazyků, citlivost na fonematické kontrasty, a zúžení vnímání s odkazem na nativní zvukový systém. Bilingvní kontext přidává k tomuto obrazu několik střípků. Důkazy naznačují, že budoucí dvojjazyčné děti se od samého počátku učí rozlišovat mezi svými rodnými jazyky, i když patří do stejné rytmické skupiny. Bosch a Sebastián-Gallés (2001) například ukázali, že čtyřměsíční kojenci, kteří se učí španělsky a katalánsky, registrují různé jazyky, přestože oba tyto jazyky jsou slabičně časované. Bilingvní osvojování tak může pohánět ranou schopnost rozlišovat více jazykových kontrastů, například rozlišovat dva podobné jazyky – což je dovednost, která u jednojazyčných kojenců nebyla prokázána.

Dále se v přehledu zabývá bilingvismem jako pojmem a zaměřuje se na kognitivní aspekty bilingvního vývoje, včetně jeho vlivu na paměť, fonologické procesy, lexikon a obecné aspekty kognitivního vývoje. Zde je bilingvismus popsán s ohledem na další dimenze, jako je věk nástupu a pořadí osvojování. Zde se bilingvní osvojování dělí na simultánní bilingvismus (mluvčí je vystaven dvěma jazykům buď při narození, nebo v raném dětství) a sekvenční bilingvismus (člověk se učí druhý jazyk poté, co byly položeny základy pro osvojení prvního jazyka) (Montrul, 2009). Shrneme-li tuto část, bilingvismus je komplexní pojem, který lze nejlépe definovat v rámci souvisejících funkčních konceptů odkazujících na doménovou specifičnost, kompetenci, odbornost a rovnováhu dovedností mluvčího. Moderní pojetí bilingvismu umožňuje pohlížet na něj široce a zahrnout scénáře, kdy znalost obou jazyků nemusí být nutně stejná.

Další kapitoly teoretického přehledu představují vývojovou poruchu řeči a diskutují její definici a aspekty. Obecně lze říci, že vývojová porucha jazyka je deficit v jazykovém vývoji a schopnostech dětí, který není spojen s žádnou jinou vývojovou poruchou (C. Norbury et al., 2008). Existuje však jiný termín pro DLD uznávaný v oboru: SLI (Eng. specific language impairment). Tato část práce se pokouší o jednoznačnost pojmů a vysvětluje volbu termínu DLD. Všímá si také rozdílného zacházení s DLD v evropsko-americkém a ruském kontextu. Velký rozdíl v ruské klinické tradici spočívá v tom, že DLD existuje jako pojem (*obecná nedostatečnost řeči* rusky: общее недоразвитие речи, OHP), ale standardizované nástroje pro hodnocení DLD nebyly až donedávna k dispozici. Hodnocení DLD se většinou opírá o kvalitativní měřítka, jako je odborné pozorování (Lopukhina A. et al., 2019).

Teoretický přehled postupuje s nástroji hodnocení DLD a popisuje nejpopulárnější metody hodnocení, jako je SRep, neboli úloha opakování vět (Eng. sentence repetition task), kdy subjekt slyší několik vět a doslovně je opakuje. Další oblíbenou metodou hodnocení jsou testy opakování pseudoslov. V nich jsou účastníci požádány, aby poslouchaly sérii nesmyslných slov modelovaných v souladu s fonologií jejich rodného jazyka a poté je opakovaly. V posledních letech se NWR doporučuje jako spolehlivý nástroj pro hodnocení monolingvních osob hovořících řadou jazyků, například angličtinou (např. Bishop et al., 1996), italštinou (Casalini et al., 2007). Tyto testy jsou nyní široce studovány i u bilingvních populací (přehled viz Chiat, 2015). Je také důležité poznamenat, že v současné době existuje a používá se celá řada NWR testů (např. NRT, CNRep) a každý test může být více či méně přesný při hodnocení fonologické pracovní paměti, morfologického zpracování a dalších domén díky návrhu podnětů úlohy.

Metodologie je zakončena popisem klíčových aspektů konstrukce testu NWR, kterými jsou délka položky, artikulační složitost a slovní podobnost podnětů. Každá z těchto proměnných může ovlivnit výkon člověka v úkolu opakování pseudoslov. Začněme délkou položky: předchozí výzkum ukázal, že delší pseudoslova se dětem opakují hůře než kratší. Vysvětlení spočívá ve fonologickém zpracování: aby bylo možné opakovat neznámé slovo, je třeba uložit do pracovní paměti neznámou fonologickou sekvenci (Baddeley et al., 1998). Další je artikulační složitost: položky se shlukem souhlásek se obvykle hůře opakují než položky bez shluku. I když neexistuje definitivní odpověď na to, proč se to děje, vysvětlení může být částečně nabídnuto špatnými "motorickými schopnostmi řeči" (Archibald & Gathercole, 2006). Poslední proměnná, slovní podoba (Eng. wordlikeness), je nejsložitější. Stručně řečeno to znamená, jak je pseudoslovo podobné skutečnému slovu existujícímu v daném jazyce. Ve skutečnosti však slovní podobnost zahrnuje takové jevy, jako je morfologická složitost, prozódie, fonotaktická probabilita (Eng. phonologic probability) a hustota sousedství (Eng. neighborhood density). Když se tyto faktory řídí vzorci daného jazyka, je pro dítě snazší opakovat pseudoslova a naopak. V další metodologické části práce jsou vysvětleny důvody výběru konkrétních proměnných pro faktory návrhu daného NWR testu.

Hlavním cílem metodologie je popsat návrh a prezentovat bilingvní anglicko-ruský test opakování pseudoslov. Závěrečný test představuje celkem 48 podnětů, 24 pro každý jazyk. Navrhovaný test opakování pseudoslov manipuluje se třemi proměnnými: délkou položky, morfologickou složitostí a fonologickou složitostí, o nichž se předpokládá, že odpovídají specifickým mechanismům zpracování jazyka: fonologické pracovní paměti, fonologickému zpracování a znalosti gramatických pravidel a dlouhodobé paměti. Konečné seznamy podnětů jsou proto složeny z položek o 2-4 slabikách, s jedním nebo žádným souhláskovým shlukem, přičemž polovina podnětů je morfologická (MNW) tj. sestávající z existujících morfémů, a druhá polovina – nemorfologická a neobsahující žádné existující morfémy (NNW). Tato kapitola pokračuje stanovením očekávání pro výkon typicky se vyvíjejících dětí a dětí s DLD v navrhovaném testu. Očekává se tedy, že typicky se vyvíjející bilingvní děti budou mít dobré výsledky jak na MNW, tak na NNW, přičemž sada MNW je snazší v dominantním jazyce díky lexikální podpoře a obeznámenosti s lexikální fonologií a morfologií. Neopakování položek může signalizovat řadu faktorů, například možná omezení při vybavování dlouhodobé paměti nebo omezené lexikální znalosti, možný deficit fonologické pracovní paměti nebo deficit fonologického zpracování.

V závěrečných kapitolách metodologie jsou stanovena očekávání výkonu případové studie. Jako první krok této případové studie bylo nutné provést sérii základních testů, které měří kognitivní a jazykové schopnosti účastníka. Kognitivní test (RCPM) funguje jako screeningové opatření k vyloučení účastníků, kteří by nemuseli být vhodní pro studii, zatímco lingvistické testy (hodnocení gramatiky TROG2, hodnocení čtení YARC a neslovní opakování CNRrep) jsou zahrnuty za účelem působení jako prediktory výkonu při plnění hlavního úkolu.

Experimentální část navazuje na metodiku a popisuje případovou studii, její cíle a předběžné výsledky. Ona představuje účastníka, který je simultánně bilingvním dítětem (7;5) s angličtinou jako dominantním jazykem. Během experimentu bylo dítě hodnoceno na sérii standardizovaných testů nad rámec navrženého testu NWR. Výsledky RCPM ukázaly, že dítě je přijatelným kandidátem z hlediska jeho kognitivního vývoje. Výsledky testů jazykových znalostí byly následující: účastník prošel YARC s průměrným skóre, TROG2 s podprůměrným skóre, ale bez indikace postižení, a CNRep s průměrným výkonem. S ohledem na navrhovaný úkol NWR dosáhl účastník v obou jazycích následujících bodů: angličtina = 21 (z 24 celkem); ruština = 18 (z 24 celkem). V anglické části testu jsou všechny chyby umístěny v nemorfologické (NNW) části seznamu podnětů. Také většina pozorovaných chyb v obou jazycích je ve čtyřslabičných podnětech.

Další, analytická část práce interpretuje výsledky dítěte ve třech standardizovaných testech (YARC, TROG2 a CNRep) a zkoumá, zda výsledky získané v těchto testech odpovídají výsledkům získaným v navrhovaném testu opakování pseudoslov z hlediska vzorců obtížnosti. V této části je vidět, že výsledky navrženého testu NWR odpovídají očekáváním kladeným na začátku práce. Shrneme-li to, navrhovaná úloha NWR prokázala adekvátní výkon, protože se ukázalo, že její výsledek závisí na znalostech gramatiky, lexikálních/morfologických znalostí a fonologické pracovní paměti dítěte.

Závěrečné kapitoly práce představují omezení případové studie, přičemž se zmiňují především o dvou aspektech. Prvním aspektem je, že výsledek ruské části testu NWR je třeba dodatečně potvrdit testovací baterií v ruštině. V této studii test opakování pseudoslov analyzoval oba jazyky, kterými účastník mluví, ale standardizované testy hodnotily pouze angličtinu, dominantní jazyk. Údaje pro druhý jazyk byly shromážděny pouze prostřednictvím rodičovského dotazníku. Druhým omezením je nedostatek důkazů o schopnosti testu prokázat podobné výsledky pro větší vzorek. Aby bylo možné definitivně říci, že tento test lze použít pro účely odlišení DLD od typického jazykového vývoje u bilingvních dětí, je nutná další statistická analýza s velkým vzorkem. Práce je zakončena zjištěním, že prezentovaný bilingvní anglicko-ruský test opakování pseudoslov má potenciál odlišit DLD od normálního bilingvního vývoje, jak bylo demonstrováno na případové studii s jedním účastníkem. V závěrečných odstavcích je naznačen další směr výzkumu a diskuze o relevanci práce v širším kontextu lingvistického hodnocení bilingvních dětí.

9. Appendix

The Appendix contains lists of stimuli for the bilingual EN-RU NWR task and sample images used for the presentation of the test in a child-friendly format. The design of the child-friendly format and the images were created by the author of the thesis.

1. Lists of stimuli for bilingual EN-RU NWR task

Russian

WORDLIKE NW				NON-WORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**горбóк /gɐrʻbok/**	**письмóвый /pʲɪsʲʻmovɨj/**	**леси́стовый /lʲɪʻsʲistɐvɨj/**	MIDDLE CLUSTER	**фи́тла /ʻfʲitlɐ/**	**такледýк /tɐklʲɪʻduk/**	**тириндили́п /tʲɪrʲɪndʲɪlʲʻip/**
INITIAL CLUSTER	**дрýгик /ʻdrugʲɪk/**	**глазóвый /glɐʻzovɨj/**	**внучáтовый /vnʊʻtɕætəvɨj/**	INITIAL CLUSTER	**двóси /ʻdvosʲɪ/**	**стуримýт /stʊrʲɪʻmut/**	**трукидилóт /trʊkʲɪdʲɪʻlot/**
NO CLUSTER	**бáлик /ʻbalʲɪk/**	**городи́к /gərɐʻdʲik/**	**сеновалóк /sʲɪnəvɐʻlok/**	NO CLUSTER	**зáпун /ʻzapʊn/**	**туламá /tʊlɐʻma/**	**липанисóт /lʲɪpɐnʲɪʻsot/**
NO CLUSTER	**дýбик /ʻdubʲɪk/**	**пирóжик /pʲɪʻroʐɨk/**	**золотóвый /zələʻtovɨj/**	NO CLUSTER	**хи́лис /ʻxʲilʲ ɪs/**	**митилóн /mʲɪtʲɪʻlon/**	**жулитулóн /ʐʊlʲɪtʊʻlon/**

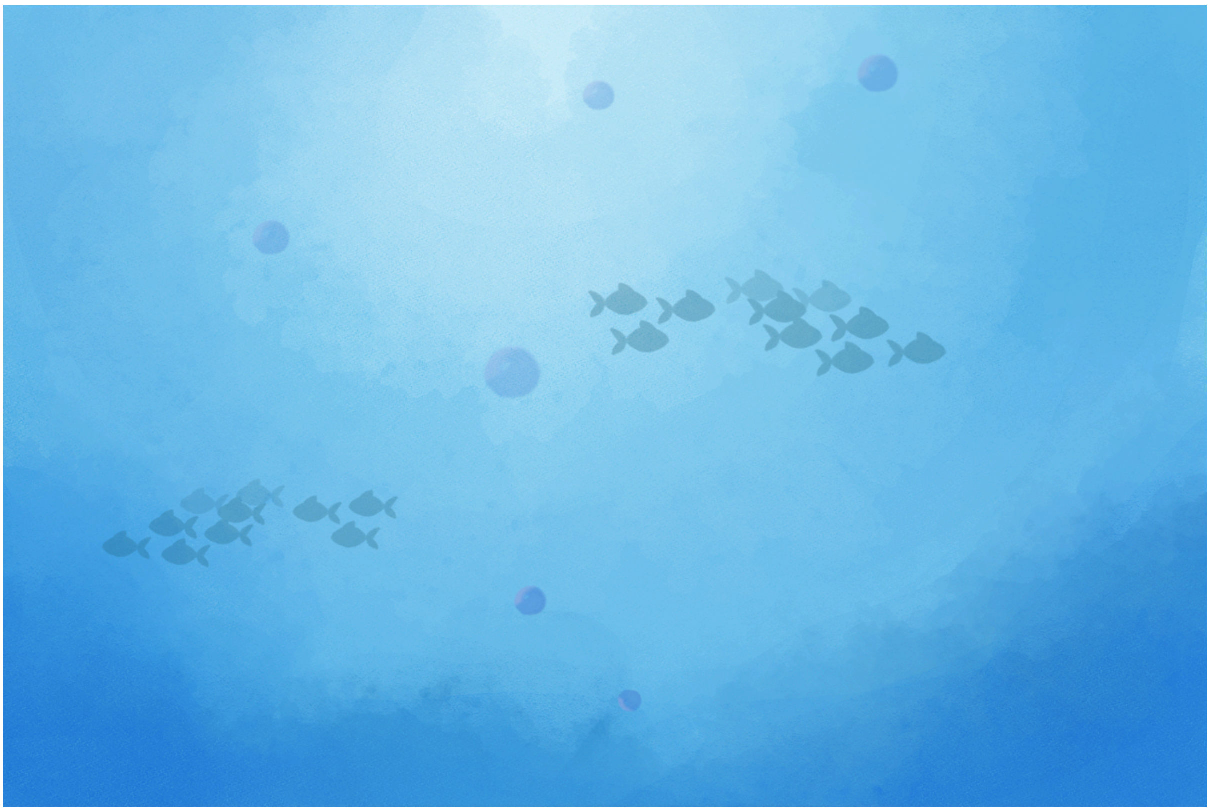
English

WORDLIKE NW				NONWORDLIKE NW			
	2 syl	3 syl	4 syl		2 syl	3 syl	4 syl
MIDDLE CLUSTER	**cavely /ˈkeɪvlɪ/**	**parrotful /ˈpærətful/**	**companily /ˈkʌmpənɪlɪ/**	MIDDLE CLUSTER	**fupeton /ˈfju:ptən/**	**tumpilit /ˈtʌmpɪlɪt/**	**duntimolap /dʌnˈtɪmoləp/**
INITIAL CLUSTER	**dryful /ˈdraɪful/**	**sparrowly /ˈspærəulɪ/**	**grocerily /ˈgrəusərɪlɪ/**	INITIAL CLUSTER	**plopek /ˈplɔ:pek/**	**flitidel /ˈfliːtɪdel/**	**blylifiton /blaɪˈlɪfɪtən/**
NO CLUSTER	**wayly /ˈweɪlɪ/**	**babyful /ˈbeɪbɪful/**	**diarily /ˈdaɪərɪlɪ/**	NO CLUSTER	**fadip /fædɪp/**	**tamifol /ˈtæmifɒl/**	**pifabimon /pifæbimən/**
NO CLUSTER	**beely /ˈbiːlɪ/**	**teddiful /ˈtedɪful/**	**familiful /ˈfæmɪlɪful/**	NO CLUSTER	**thoppin /ˈθɔ:pin/**	**thopilan /ˈθɔ:pɪlən/**	**talifisop /tæˈlifisəp/**

2. Child-friendly format presentation (excerpts)

a. Gamified presentation of the test

b. Initial slide



c. Final slide