School of Psychology and Speech Pathology

Assessing Dual Language Learners with Primary Language Impairments: Developing an assessment approach

Naomi Li'el

This thesis is presented for the Degree of
Master of Philosophy (Psychology)
of
Curtin University

February 2017

DECLARATION

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Human Ethics The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number # HR 225/2014; and the Western Australian Department of Education, Approval Number # D15/0015395 (see appendix A and B).

Signature: ……………………………………………

Date: … February 2017……..

ABSTRACT

Similar to other western countries, bilingualism in Australia is on the rise. However, high diversity creates a unique cultural and linguistic mosaic with no dominant language other than English. While awareness of bilingual matters is high among Speech Language Pathologists (SLP), providing equitable services to all clients regardless of the languages they speak is a challenge. Specific difficulty is evident for initial assessment, as the gold standard requires assessment in both/all the languages the child speaks. However, the large range of mother tongues does not easily allow to implement this recommendation in practice, yielding the need for innovative and creative assessment approaches to discriminate bilingual children (dual language learners [DLLs]) with Primary Language Impairment (PLI) from typically developing (TD) DLLs.

The aim of this project was to investigate a unique approach that incorporates a parent questionnaire (ALDeQ) and two language processing tasks (nonword repetition [NWR] and recalling sentences [RS]) administered in English. In addition, we sought to explore the impact of internal and external factors on the performance in language processing tasks.

Performance of 42 typically developing DLLs and 19 DLLs with PLI was compared. Groups were matched for age (*M*=5;10) socioeconomic status (*M*=1,023 SEIFA), length of exposure to English (*M*=33.4 months) and language typology (21% non-tense-marking mother tongues).

Results found that DLLs with PLI had significantly ($p < .005$) lower scores than TD DLLs on all three assessment tools at the group level and that the ALDeQ provided the highest diagnostic accuracy (100% sensitivity, 95.2% specificity, LR+ =21, LR- =0.00, AUC =0.991). Two other combinations also provided good diagnostic accuracy (above 80% sensitivity and 80% specificity):.combination of ALDeQ and NWR; scores below the cut-off line on any combination of assessment tools. In addition, results showed that TD DLLs outperformed DLLs with PLI on language processing tasks, but fell below monolingual standard scores. Correlation analysis showed that performance was impacted by internal/external factors.

While best practice calls for assessment of DLLs in both/all the languages in which they function, a combination of indirect and carefully selected direct assessment tasks may support SLPs to correctly distinguish TD DLLs from DLLs with PLI. This approach may have the potential to minimize the disadvantage of minimal or reduced exposure to English in the assessment process of DLLs while providing a practical and evidence based solution for speech language pathologists.

ACKNOWLEDGEMENTS

On completion of this thesis, I would like to reflect on the people who have supported and encouraged me along the way. Foremost, I would like to express gratitude to my supervisor Dr Cori Williams whose expertise, understanding and patience were invaluable to this project and my graduate experience.

I would also like to thank additional faculty members from the school of Psychology and Speech Pathology at Curtin University: Dr Robert Kane, my statistical advisor, and Dr Suze Leitao, a senior lecturer. Thank you for your insightful comments, hard questions and for sharing your immense knowledge with me.

Thank you to all the families that kindly took interest in the study and volunteered to participate. I was overwhelmed by the willingness to take part and assist in recruitment. A very special thanks to the dedicated staff at participating schools and particularly to the cultural brokers and interpreters. Without their tireless work and invaluable relationships with families, the outreach and diversity of this study would not have been possible. I would also like to acknowledge Rosemary Simpson for her ongoing support of local research in the field of language impairment through her role as Language Development Centre principal.

On a personal note, I am extremely grateful of the support of my family, close and far. I wish to acknowledge my partner and best friend Yoni, without whose love, encouragement and endless patience, I would not have finished this thesis. Last but not least, very special thanks to my parents in Israel who, despite the distance, were always willing to help and provide vital insights every step of the way.

CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Meaning
ALEQ	Alberta Language Environment Questionnaire
ALDeQ	Alberta Language and Development Questionnaire
AUC	Area Under the Curve
BIPAQ	Bilingual Parents Questionnaire
CCC-2	Children’s Communication Checklist, Version 2
CELF-P-2	Clinical Evaluation of Language Fundamentals Preschool- Australian Standardised Second Edition
CELF-4	Clinical Evaluation of Language Fundamnetals- Australian Standardised Fourth Edition
CPM	Coloured Progressive Matrices
CTOPP	Comprehensive Test of Phonological Processing
DLL	Dual Language Learner
ENNI	Edmonton Narrative Norms Instrument
ESL	English as a Second Language
LDC	Language Development Centre
LR	Likelihood Ratio
NWR	Nonword Repetition
PABIQ	Parents of Bilingual children Questionnaire
PLI	Primary Language Impairment
PPVT	The Peabody Picture Vocabulary Test
ROC	Receiver Operating Characteristic
RS	Recalling Sentences

SD	Standard Deviation
SEIFA	Socio-Economic Index For Areas
SES	Socio-Economic Status
SLP	Speech-Language Pathologist
SPA	Speech Pathology Australia
TEGI	Test of Early Grammatical Impairment
TNL	Test of Narrative Language
TOLD-P	Test of Language Development - Primary
UK	United Kingdom
US	United States
WA	Western Australia
ZPD	Zone of Proximal Development

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1- INTRODUCTION

This research was undertaken to inform those involved in child health and education within a multilingual context. The overarching principle driving this project is equality, and the belief that health and education service providers can actively influence fair and equal service provision to all clients regardless of linguistic, cultural or religious differences. Of particular importance are speech language pathologists (SLPs) who are at the front line of early language assessment and have reported many challenges limiting their ability to assess children from diverse backgrounds in a reliable, valid and evidence based manner.

1.1. Objectives

The main objective of this research project was to explore the use of a unique assessment approach to address current challenges in assessing bilingual children (dual language learners [DLLs]) within the Australian context. The Australian cultural-linguistic mosaic and an overview of monolingual and multilingual language development in this population are discussed in chapter 2. Chapter 3 unravels the complex challenges of assessment worldwide and specifically in Australia.

Secondary objectives aim to explore the effects of language environment and specific factors of multilingualism on task performance. The factors that influence language development in both monolingual and DLL children are presented in chapter 2.

1.2. Methodology

When building the assessment approach tested in this study the unique characteristics of both DLLs and SLPs in Australia were taken into consideration. This was done in an attempt to produce a solution that could resonate beyond the research community and be readily used in clinical practice. Therefore, assessment tools were chosen to allow the following three points to occur simultaneously: (1) detect hallmarks of primary language impairment (PLI), (2) reduce cultural/environmental bias and (3) gain information on both/all languages to which the child is exposed.

1.3. Methods

In order to address the points detailed above, the assessment approach explored in this study is based on a larger Canadian study (J. Paradis, Schneider, & Sorenson Duncan, 2013) that used a combination of direct and indirect tools in the child's second language to identify impairment. Following recommendations for future research suggested by Paradis, Schneider and Sorenson Duncan, and current research in the field, the present study used a parent questionnaire to gain information on the child's mother tongue and language processing tasks in English to detect hallmarks of primary language impairment (PLI)[1] directly. Language processing tasks were chosen as they are less impacted by compounding factors and could minimize cultural/environment bias. Detailed accounts of previous research are presented in chapter 3.

A case control study design was chosen to compare DLLs with and without PLI. Recruitment focused on schools and communities with a high population of DLLs within the Perth metropolitan area in order to capture a sample that would present a cohort of Australian DLLs. Demographic characteristics of the sample are presented in chapter 4.

[1] The term primary language impairment (PLI) will be used throughout this thesis. In past research and clinical caseloads, different terms are used to describe this sub-group of children: *specific language impairment (SLI), expressive/receptive language delay, late talkers and language-based learning disabilities.*

The researcher consulted with cultural brokers and worked closely with interpreters during fieldwork. This was important during recruitment to ensure families could provide informed consent and comprehensive parent interviews. Chapter 4 describes the recruitment process and provision of the assessment approach.

1.4. Results and Limitations

Interpretation of the results and discussion (chapters 5 and 6) and particularly the conclusions (chapter 7) should take into account the limited scope of the current study. Although limitations will be discussed in detail in chapter 7, to set the stage for further reading it should be noted that the study was small (N=61) and independent assessment of language abilities was not possible. Hence group classification (typically developing or with language impairment) was based on education setting.

1.5. Contribution

While research on multilingual language assessment has significantly expanded in recent times, few studies have been conducted in Australia. In addition, no study to date has explored the specific assessment approach used in the current study. Results can contribute to the growing cross-linguistic global evidence in the field of multilingual assessment and provide preliminary data specific to Australia. On a general level, the findings can assist health and educational service providers to better understand multilingualism and how it presents within the Australian context. Moreover, the results of this study provide a specific and practical approach which could be readily adapted by SLPs thereby improving their professional toolbox.

CHAPTER 2 -LANGUAGE IN A MULTILINGUAL CONTEXT

The main focus of this thesis is language impairment in a multilingual context, specifically identification of language impairment among Dual Language Learners (DLLs). A clear understanding of what language is and how it typically develops is fundamental to any discussion on impairment, particularly when attempting to correctly identify individuals with and without a disorder. In addition, the multilingual context of this thesis requires a clear definition of multilingualism, understanding of multilingual language development, and how impairment presents within this population.

2.1. Multilingualism

Who is considered multilingual and which measures are used to determine multilingualism? A conventional use of the term ‘multilingual’ refers to an individual who speaks **more than** two languages. This conventional definition differs from the term ‘bilingual’, an individual who speaks **no more** than two languages. However, in some publications bilingualism and multilingualism are used interchangeably, referring to individuals who speak **two or more** languages (Fielding, 2015).

A key measure of multilingualism is proficiency. In its most basic form, an individual is considered multilingual if she[2] has balanced proficiency in all the languages she speaks. Yet, as language is a complex tool used for communication in a range of environments for different purposes, equal proficiency across languages is uncommon. A broader definition includes **varying** proficiency across languages.

[2] Pronouns will be used interchangeably throughout this thesis (he/she, her/his)

Another measure frequently used to define multilingualism is the pattern of acquisition. Some individuals acquire the languages they use simultaneously. This refers to children who are exposed to different languages from birth or shortly after. An alternative pattern of acquisition is successive or sequential learning of languages. This refers to children who learn additional languages after acquiring foundations in one language. Although there is no definitive cut off age to differentiate simultaneous from sequential patterns of language acquisition, many researchers (e.g. Roesch and Chondrogianni, 2016; Sorenson Duncan and Paradis, 2016; Summers, Bohman, Gillam, Peña and Bedore, 2010) have accepted the age of 3 as a point of reference. By this age established knowledge in one language can be visible in terms of vocabulary and grammar (J. Paradis, Genesee, & Crago, 2011).

In line with recent publications in the field of Speech-Language Pathology and stemming from the theoretical framework on which this thesis is based (see 2.4), Kohnert's (2013) , a broad definition of multilingualism is used in this thesis. Hence a multilingual child is defined as a child who functions in two or more languages in her typical daily environments. This definition includes both sequential and simultaneous patterns of acquisition and allows for varying proficiency across languages. To reflect this broad definition and to emphasize the diverse group of languages in the current study, the term *dual language learners* (DLLs), which includes children learning two or more languages, will be used throughout.

2.2. Global Trends of Multilingualism

Global migration has expanded to the point that multilingualism is more common than ever before: "Vast population mobility has converted most countries of the world into plural societies, so that the experience of community-level multiculturalism and multilingualism is now itself universalized" (Lo Bianco, 2014 p. 314).

Following the wide spread effect of migration, one might expect similar changes to have occurred in all predominantly English speaking countries. A comparison reveals that Australia, Canada, the United Kingdom and the United States have similarities in some domains. They all have a rich history of immigration, similar socio-demographic characteristics and relatively open concepts of citizenship (Washbrook, Waldfogel, Brabury, Corak, & Ghanghro, 2012). In addition, Speech-Language Pathology ethical frameworks and policies in these countries specifically refer to providing equal access to care and non-discriminant service provision (e.g. Speech Pathology Australia's Competency-Based Occupational Standards, 2011).

However, some important differences also exist. In Australia, the United Kingdom and most states in the United States, English is the only official language, while in Canada French is also an official language. In addition, the proportion of speakers of minority languages differs greatly from country to country. The proportion of Spanish speakers in the United States is relatively large, accounting for 11% of the general population. In contrast the proportion of each language group other than the official languages in Australia, the United Kingdom and Canada is on average 1% or less (Australian Bureau of Statistics, 2011; Statistics Canada, 2011; United Kingdom Office for National Statistics, 2013; United States Census Bureau, 2015). Lastly, policies concerning criteria for selecting immigrants, settlement schemas, and schooling, differ greatly.

Overall, the most comparable countries may be Canada and Australia. In addition to sharing English as an official language and similar variety and proportion of non-official languages, they share similar migration policies. Both countries focus on admitting higher skilled migrants under specific labour categories. This similarity may be of importance when focusing on language abilities, as previous literature shows a link between selection policies and child immigrant outcomes on comparative tests during school years (Washbrook et al., 2012). Studies conducted in Canada formed the foundations of the current study. Despite these similarities, Australia has unique characteristics that will be discussed in the following section.

2.3. Multilingualism in Australia

Immigration waves to Australia over the past two centuries have shaped the country’s cultural and linguistic landscape, adding languages from across the globe to the existing indigenous languages. Data from the most recent census show that 23.2% of households (Australian Bureau of Statistics, 2011) use a language other than English at home. The general proportion of these households and the specific proportion of each minority language are closely related to migration trends and thus have changed since the 2006 census. Current data shows a 1.7% increase of households using a language other than English since 2006 (from 21.5%). The latest census (2011) recorded Mandarin (1.6%), Italian (1.4%), Arabic (1.3%), Cantonese (1.2%) and Greek (1.3%) as the top 5 minority languages. Although these languages were the top 5 in 2006, the proportion of each was different with Mandarin moving from fifth place to first and Greek dropping from second to fifth by 2011.

When focusing on children aged 4-5 years old, more recent data is available. Verdon, McLeod and McDonald (2014) reported that 15.3% were exposed to a language other than English at home and that the top 5 minority languages differed from the data reported in the 2011 census: Arabic (1.5%), Italian (1.2%), Greek (0.9%), Spanish (0.9%) and Vietnamese (0.9%).

In summary, Australia is a multilingual nation in which a wide range of languages is spoken. Australia's linguistic landscape is dynamic and continuously changing as a reflection of migration from different parts of the world. This linguistic context will be considered throughout the current study within theoretical frameworks of language development.

2.4. Theoretical Frameworks of Language Development

The theoretical framework of this thesis draws on an understanding of what language is and how it develops using an integrated approach of theories from different disciplines. Many areas of research, such as psychology, cognitive sciences and, most relevant to this thesis, speech-language pathology, include language under their realm of interest. As theories attempt to provide principles or overarching statements regarding complex phenomena based on empirical data, they determine which aspects are put in the spotlight and influence interpretation of new data. Hence different theories may highlight different areas and methodology, and terminology may not be consistent. The following section will present complementary theories. First, general theories based on primarily monolingual data will be presented, followed by adaptations based on multilingual data. Together these theories form the base that guided the construction and interpretation of this study.

2.4.1. Social Interactivism.

Social interactivism, also known as social constructivism, highlights the broad context in which language development occurs. According to this perspective, language develops through meaningful, constructive interactions between a child and her caregivers within dynamic environmental contexts. Based on the work of Vygotsky (1986), language is seen as an integrated part of a child's general development. Social interactivism views integration of different developmental areas as a central to language development (e.g. the impact of the interaction between cognition and motor development on language development). Gradually children draw on their growing understanding of the world and linguistic experiences to draw inferences about meaning and structure of the language being learned.

According to social interactivism the driving force of change is the child who is internally and intentionally motivated to engage with others. Bloom and Tinker (2001) argue that there is tension between the will to engage and be an active part of social interactions and the effort needed to do so using cognitive, emotional and physical resources. This tension internally motivates the child to learn language. Vygotsky (1986) claims that learning occurs within the boundaries of the child's "zone of proximal development" (ZPD), ranging from what the child can do independently to what the child can do with mediation from a more skilled individual. Language learning occurs when the child and adult collaborate and the adult mediates novel information within the ZPD.

The social interactivism perspective places importance on the individual's pattern of learning, rather than static outcomes or set criteria. The ZPD and the type and rate of mediation needed to make gains are used to reflect the child's learning style and ability to learn. Dynamic assessment (see section 3.3.4) uses these measures to detect language impairment. In addition, as language development is seen within context, the ability to make gains is interpreted in relation to the different environments in which the child functions.

In the current study an interactivism perspective guided the choice of assessment tools. The aim was to use assessment tools that provided information about the child's language environments, while minimizing this impact of environment on performance.

2.4.2. Psycholinguistic Processing Models

Processing models focus on the underlying systems that ‘work’ in service of language. The mental systems used to process and produce oral language according to most psycholinguistic models are the conceptual system, the language system, and the articulatory system. The conceptual system consists of non-linguistic information (understanding of concepts based on experience/sensory input and influenced by semantic knowledge and pragmatics); the language system consists of semantic, syntactic, morphological and phonological information; and the articulatory system consists of motor information. These systems interact to process and comprehend language and transform internal thoughts to spoken language (Levelt, 1989; Stackhouse & Wells, 1997).

According to these models, there are strong links between the systems and cognitive mechanisms. Cognitive mechanisms power the capacity of each system and influence the speed and efficiency of information processing. Specifically, perception, memory, attention and emotion impact how quickly, in what manner, and to what extent linguistic information is identified, stored and retrieved. The full process consists not only of simultaneous exchanges of information within language processing systems (e.g. semantics interacting with syntax), but also of exchanges between general cognitive mechanisms and language processing systems (e.g. memory interacting with phonology) (Leonard, 2014; M. Paradis, 2004). Strong cognitive abilities paired with rich and well organized mental systems allows relatively quick and efficient learning of novel linguistic information.

The internal organization of items within each system is determined by thresholds. The more frequently an item is used the lower its threshold and the ‘easier’ it is to access (M. Paradis, 2004). Individuals decode speech by paying attention to word boundaries, identifying speech sounds and comparing them to sounds stored in the phonological system.

When thresholds are met, phonological sequences are matched to words in the semantic system and the sounds are interpreted as meaningful. When thresholds are not met, the speech sounds are held in the phonological short term memory (phonological loop) and compared to information in the long term memory (Baddeley, 2012; Rönnberg et al., 2013). This is how known words are understood and new words are learned. While in the short term memory, links begin to form connecting the new information to information stored in the long term memory. These links embed the new word in mental networks which represent all the linguistic and conceptual information associated with the new word (Collins & Loftus, 1975).

Tracking the processes of identifying, storing and retrieving novel auditory information can shed light on an individual's underlying language processing system or, put simply, how an individual learns new words or sentences. In the context of the current study, processing models and specifically the idea of creating a novel learning experience independent of language experience, guided the choice of some assessment tools (the language processing tasks).

2.4.3. Usage-Based Perspective

A usage-based perspective views language as a dynamic phenomenon guided by principles for the purpose of communication. Therefore the focus is on detecting patterns that underpin language development (Bybee, 2010). Similarly to the interactivism perspective, language development is seen as part of general development and language as one of many domains. Different languages provide listeners with varying cues to encode meaning. Cues include structural elements, such as syntax, as well as prosodic elements, such as intonation.

When a proficient listener hears a sentence a number of cues may contribute to decoding meaning such as: word order; subject – verb agreement; case markings; inflectional morphemes; intonation or stress. For example in the sentence "She said she did not take his money", the meaning changes according to the word that is stressed. "She said she did not take his **money**" means that she did take something, but not money. However, "She said **she** did not take his money" means that money was taken, but not by her. Hence stress is an important cue in understanding the meaning of this sentence.

From the usage-based perspective the ability to decode sentences can be viewed as a window to the individual's ability to detect linguistic cues. In addition to intonation, children learn how to map grammatical form to function based on exposure from the environment. Children acquire a grammatical form when they reach a 'critical mass' of input that allows them to abstract a pattern or rule from accumulated relevant structures (V.C.M. Gathercole, 2007). The frequency of exposure and consistency of form-to-function mapping are believed to predict developmental outcomes to a greater degree than how frequent a specific form is in a particular language (token frequency) (Bybee, 2008). Therefore tracking the process of imitating a sentence conveys information on what linguistic cues the individual can decode and reconstruct. For example, if a 6-year-old child who can produce the phenome /s/ omits plural 's' ("Did the boy eat the apple?" instead of "Did the boys eat the apples?"), we can hypothesise that he did not encode the grammatical cue for plurals, or did not have the grammatical knowledge to reconstruct this feature. More evidence would be needed to confirm or reject if indeed this hypothesis is true.

In summary, an integrated approach including different perspectives on language development formed the framework of the current study, highlighting how language development is embedded in one's environment and how typical function of underlying mechanisms is integral to language growth. Yet, it is vital to remember that these theories are predominantly based on monolingual data. The next section will focus on multilingual adaptations to language development theories.

2.4.4. Theories of Multilingual Language Development

The consideration of multilingual language development has raised many questions and debates among researchers in the field. Does the multilingual brain process and produce language in a unique manner? Do multilingual individuals have one united linguistic system that processes all languages or separate systems for each language? Can theories of language be adapted when more than one language is at play?

The first widely used theoretical account of simultaneous multilingual language development was published in 1978 by Volterra and Taeschner: the Unitary Language System Hypothesis. According to this hypothesis simultaneous DLLs pass through three stages. Initially DLLs have one linguistic system that includes words from both or all languages. Later, between the ages of two and three, DLLs separate words from different languages into separate mental lexicons, but use one grammatical system. In the last stage, both semantic and grammatical knowledge is separated into two different systems. This hypothesis was supported by a detailed, yet small (n=3), longitudinal study (Volterra & Taeschner, 1978). They found that during the first stage, based on data collected from 18 to 23 month old children, simultaneous DLLs know different words in each language and have few translation equivalents (a word for the same concept in both languages). This evidence showed that concepts rarely had two mental representations in different languages and therefore supported the hypothesis that all words were part of one united system. In addition they demonstrated that between the ages of 24 to 33 months, albeit acquiring many translation equivalents, grammatical rules from one language were used in the other even though it resulted in error. They concluded that while two mental lexicons had developed, only one syntactic system was acquired. Finally, data collected between 33 and 42 months of age showed correct use of grammatical structures in both languages. These final data lead the authors to conclude that at the final stage the child was “truly bilingual”, and had reached monolingual competence in both languages.

Later research has interpreted these findings differently and has produced evidence that supports an alternative hypothesis, the Subsystem Hypothesis, according to which simultaneous DLLs exhibit two separate but interconnected linguistic systems from the initial stage of the language development process (de Bot, 1992; Hambly, Wren, McLeod, & Roulstone, 2013; M. Paradis, 2004). Early evidence can be found in multilingual infants who can distinguish between different languages acoustically (Bosch & Sebastian-Galles, 2001) and whose babbling-style reflects linguistic elements of the dominant language, as opposed to a mixture of both languages (Poulin-Dubois & Goodz, 2001). Additional studies attribute the low number of translation equivalents to the different environments in which each language is learned (Kan & Kohnert, 2005) and grammatical errors to cross linguistic influences (see detailed explanation in section 2.7.1) (Bedore & Peña, 2008; Gawlitzek-Maiwald & Tracy, 1996; Matthews & Yip, 2003). Moreover, recent studies have shown evidence that some DLLs are sensitive to two different grammatical systems before the age of five and use correct word order and specific morphology and syntax for each language, further strengthening the hypothesis that each language has its own sub-system (Meisel, 2001; J. Paradis, Nicoladis, & Genesee, 2000; Thordardottir, 2015b). Throughout the learning process, the languages the child knows continue to influence each other, especially when knowledge that exists in one is lacking in the other (Yip & Matthews, 2007).

When focusing on the underlying mechanisms and neuropsychological processes, current research supports the theory that both monolingual and multilingual (simultaneous and sequential) minds process language in a similar fashion. There is no evidence to suggest that the mental processes are different in a qualitative way (de Bot, 1992; M. Paradis, 2004; Runnqvist, Fitzpatrick, Strijkers, & Costa, 2012), yet quantitative differences are present. For example, a multilingual infant has two or more phonological sub-systems allowing for discrimination between sounds of different languages, while a monolingual infant has one phonological sub-system. Likewise multilingual children have two or more grammatical sub-systems allowing different grammatical encoding for different languages while monolingual children have one. The difference is in the number of sub-systems and therefore could be described as a quantitative difference.

Psycholinguistic models infer that infants or children are unaware they are choosing one language or inhibiting the other. The Direct Access Hypothesis (M. Paradis, 2004) suggests that similarly to monolingual individuals, multilingual individuals access the information stored in the language system according to thresholds. The speed and accuracy of access to and retrieval of items depends on the threshold of each item. The more frequently an item is used the lower its threshold and the 'easier' it is to access. Multilingual individuals have direct access to the most appropriate linguistic information for a specific context (constrained by pragmatics), in the same way monolingual individuals do. However multilingual individuals are obligated to include a pragmatic consideration that monolingual individuals do not consider. They must implicitly choose which language to use, or decide whether code-mixing (using elements from both languages interchangeably within one utterance) is appropriate (M. Paradis, 2004).

There are different implications for thresholds between sequential and simultaneous DLLs. Sequential DLLs acquiring a second language approach the task with an established body of knowledge in their first language. Hence the first language sub-systems will initially contain more items than the second language sub-systems and the thresholds of the items in the first language will be lower. The quantity of items in both simultaneous and sequential DLLs sub-systems will reflect the environment in which each language is used. For example a Vietnamese-English DLL who attends school in English will have a larger school-related vocabulary (e.g. recess, slip-slap-slop, cubby etc.) in English than Vietnamese. The exposure to, and use of, each item affects its threshold creating unequal representations of each language across the domains (J. Paradis, Genesee, et al., 2011).

The usage-based and social interactivism perspectives also identify thresholds as a key element in interpreting differences between monolingual and multilingual language acquisition. These perspectives argue that DLLs will present with initial delays in acquisition of grammatical structures since it will take them longer to reach the critical mass of linguistic information required to generalise context based knowledge into internal patterns and rules. However once the critical mass is reached monolingual children and DLLs will perform similarly. Therefore DLLs will go through similar stages of development, in the same order, but at a different pace.

Previous research has shown that TD DLLs acquire grammatical structures at a faster pace than might be expected based on amount of input (Blom et al. 2016; Blom and Paradis, 2013; Paradis, 2010; Thordardottir, 2015). V. C. M. Gathercole (2007) proposes a social interactionist (constructivist) model that can be used to explain the pace of multilingual language development. According to her model, although sub-systems are language specific, two or more language sub-systems may interact when languages share structural similarities.

Thus ‘critical mass’ is achieved faster through shared information across languages. DLLs do not have to learn everything about language twice even if rate of acquisition is sensitive to reduced and variable input.

Lastly, the social interactivism perspective can be adapted to address the social context of multilingual language development. The driving force to communicate socially is shared across monolingual children and DLLs, however within the multilingual context, this may require a child to attend to and acquire more than one language. Applying this perspective will highlight the differences between monolingual and multilingual language development as both the pattern of learning and modifications used to facilitate language learning will differ due to context.

Similar theoretical frameworks account for both monolingual and multilingual language development. Psycholinguistic processing models argue that the mechanisms and neuropsychological processes that underpin monolingual language development are no different to those of multilingual individuals. Likewise the principals of usage-based and social interactivism theories, that link the linguistic cues in one’s environment to the ability to encode language, can be adapted within the multilingual context. Lastly, according to the interactivism perspective the impetus to develop language is shared among monolingual and multilingual individuals. They both develop language in order to communicate socially. Nevertheless, key differences are apparent. Multilingual individuals’ language processing system consists of sub-systems that contain language specific information. In addition multilingual individuals develop language within a linguistic environment that is different to that of a monolingual individual. Therefore cues to encode language and social communication demands and opportunities are different.

2.5. Language Development within a Multilingual Context

Language development refers to the process of learning and mastering skills that lead to comprehension and production of language. It involves a gradual increase in proficiency in the different language domains (semantics, syntax, phonology, morphology and pragmatics). This section will address the unique features of multilingual language development. First, monolingual language development will be briefly reviewed to form a baseline for comparison, and then multilingual language will be addressed.

2.5.1. Monolingual Language Development

Although language development among monolingual children is far from homogeneous due to a child's environment (e.g. socioeconomic status, maternal education) and individual differences (e.g. memory, attention), Meisel (2011) identifies three characteristics of typical language development: ultimate success, rate of acquisition and uniformity.

All typically developing children regardless of environment or individual differences develop language. That is, they experience success in acquiring language to varying degrees of proficiency. This is not to suggest that developing language is an easy or effortless task. Another characteristic of typical language development is rate of acquisition. All typically developing children move from one milestone to the next at a similar rate. For example during one year, between the ages of 2 and 3, most typically developing children learn to produce utterances with more than one word (Paul, 2012). Thus within a year children learn to connect two or more words together to form simple sentences. A deviation from this rate, such as staying at the one-word-utterance level for more than one year, is considered a deviation from typical development. Lastly the stages of language development (or milestones) follow a rather uniform path, even across different languages. All typically developing children first develop gestures and then develop first words (Paul, 2012). These three characteristics, as well as theories of language development, will guide the comparison between monolingual and multilingual language development in the following section.

2.5.2. Multilingual Language Development

Evidence from recent research shows that typically developing simultaneous DLLs reach the same early milestones at the same rate as monolingual peers in many domains (Bedore & Peña, 2008; J. Paradis, Genesee, et al., 2011; Petitto et al., 2001). Hence many similarities to monolingual language development can be found. With specific regard to expressive word milestones, when the number of words from both languages is combined, it is consistent with age matched monolingual children (Pearson, 1998; Pearson, Fernández, & Oller, 1993). Similarly, simultaneous DLL's morphological and syntactic systems evolve according to the stages known in typical monolingual development; that is, showing a gradual increase in length and complexity reaching a well-established foundation by the age of five (Bedore & Peña, 2008). Furthermore, DLLs show similar stages for the acquisition of narratives (Kohnert, 2013) and phonological awareness skills when compared to monolingual children (Werker & Weikum, 2006).

Despite the underlying similarities, Grosjean (1989) emphasizes that bilingual children are not the sum of two monolingual children, but rather a fused whole with a unique linguistic profile. Central to this difference is the presence of two or more sub-systems across the language domains and the thresholds of the items within each sub-system.

One unique developmental difference between DLLs and monolingual children is the distribution of vocabulary. Although consistent with monolingual norms when both or all languages are combined, DLLs have two or more semantic sub-systems as opposed to one

(Oller, Pearson, & Cobo-Lewis, 2007). The semantic system is usually larger in the dominant language (Kan & Kohnert, 2005) and when each language is tested separately, the raw number of words is smaller than expected from monolingual peers (J. Paradis, Genesee, et al., 2011; E. D. Peña, Bedore, & Zlatic-Giunta, 2002; Verhoeven, 2000).

Under the premise that the main task in word learning is learning the concept, and the form is secondary, some researchers have questioned the use of raw number of words and have suggested counting conceptual vocabulary (Marchman, Fernald, & Hurtado,2010; Pearson et al., 1993). Conceptual vocabulary counts concepts, rather than words, with the result that a concept which is expressed in both languages is counted only once (for example, ‘ball’ and the Hebrew equivalent /kadur/ would be counted once when assessing a Hebrew-English DLL). In studies that compared raw word and conceptual scores (in both languages) amongst Spanish-English DLLs, the DLLs conceptual scores were significantly lower than monolingual children, but their raw word total matched the norm of their monolingual peers (Core, Hoff, Rumiche, & Senor, 2013; Pearson et al., 1993). These studies support the theory that learning phonological forms is a major part in word learning and thus takes time (Storkel, 2001). Therefore, DLLs need more time than monolingual children to expand their conceptual vocabulary since they learn more than one phonological representation for certain concepts. This should be viewed as a unique developmental difference and not a sign of language impairment (Core et al., 2013). Similarly, DLLs need more time to acquire some grammatical aspects of their dominant languages since they need to learn specific rules for each language (J. Paradis, Genesee, et al., 2011; Thordardottir, 2015b).

A further difference in language development in DLLs is cross linguistic transfer: using knowledge from one language to develop the other. This phenomenon occurs when linguistic knowledge from one language boosts the learning of the other language (Gawlitzek-Maiwald

& Tracy, 1996). For example, a young child from a German speaking household who lives in Italy may have broader grammar knowledge in German. When attempting to construct a sentence in Italian, the child may draw on available knowledge in German producing a sentence with words in Italian based on German grammar. In this instance German temporarily 'crosses over' to Italian and influences the way the sentence is constructed. The result is a form of mixed knowledge using unusual grammar and morphology. This should be viewed as productive use of language knowledge rather than grammatical errors (Bedore & Peña, 2008; Yip & Matthews, 2007).

Similarly, DLLs often fill semantic gaps in one language with semantic knowledge from the other. This unique phenomenon of multilingual language is referred to as code-mixing and is common in multilingual communities during informal interactions (Zentella, 1999). Whereas it may reflect lexical gaps in one language, it is not necessarily a sign of language impairment (Kohnert, 2013).

DLLs attend to speech in the same way as monolingual children, however new sounds can be compared to phonological and semantic items in all the language sub-systems, while monolinguals only have access to one (M. Paradis, 2004). Similarly, when producing sounds, simultaneous DLLs can draw from two phonological systems and produce specific kinaesthetic plans for the articulators which may account for simultaneous DLL's ability to speak in a native accent in more than one language (Serratrice, 2012).

Additionally, a unique feature of language development, especially in countries like Australia where monolingual English is dominant in most environments, is language attrition. Typically, language proficiency or dominance shifts from the mother tongue to the majority language according to changes in exposure. This shift is common when DLLs start school in the majority language, increasing the quantity and quality of input in that language while decreasing input in their mother tongue. As a result language development in the mother tongue may plateau or differ from monolingual speakers of the majority language. Hence multilingual language development can differ from monolingual language development in both the mother tongue and the second language. Similarly to gaps in the second language, relatively low proficiency in the mother tongue may be a sign of language attrition, and not impairment (Kohnert, 2010; J. Paradis, Genesee, et al., 2011).

It is important to note that differences can be found between simultaneous and sequential DLLs. While simultaneous DLL's grammatical development generally follows monolingual trajectories (albeit at a slower pace) sequential DLLs may not (Abrahamsson & Hyltenstam, 2008; Roesch & Chondrogianni, 2016). Research in this area is scarce and thus it is difficult to explain this discrepancy. The underlying reason may be related to environmental and internal variables such as age of first exposure to the second language, length of exposure and typological similarities or differences between the spoken languages. In addition, sequential DLLs have reduced perception of speech sounds in their second language which could impact their ability to imitate new words correctly or speak with a native accent (Bosch, Costa, & Sebastian-Galles, 2000; Sebastian-Galles & Soto-Faraco, 1999).

In summary, both typically developing DLLs and monolingual children ultimately develop language in a relatively uniform manner. However, in terms of vocabulary and grammar acquisition, DLLs may show different rates of acquisition for each language that are not comparable to the typical monolingual rate of acquisition. In addition, DLLs can draw linguistic information from multiple language specific sub-systems and are exposed to varying cultural-linguistic environments. The combination of these factors creates a unique linguistic profile. However, underneath these surface differences, multilingual and monolingual minds process language in a similar fashion. Awareness of both similarities and differences is central to the design and interpretation of the present study.

Thus far the comparison between groups has been discussed, however it has long been established that there is substantial variation in the development of language across different children. Therefore general factors that influence the individual linguistic profiles of children will be discussed in the next section.

2.6. Factors Influencing Monolingual Language Development

Research among monolingual children (Hart & Risley, 1995; Hoff, 2003; Schwab & Lew-Williams, 2016) has shown an overall correlation between the quality and quantity of linguistic input provided to children during early language development and performance on language tasks across linguistic domains. Rich and frequent experience in a range of language activities (such as pretend play, book sharing and word games) contributes to higher overall language outcomes. Quality and quantity of language input has been associated with socioeconomic status and socioeconomic related factors such as parental education, parental occupation, family income or a combination of all three (Hoff, 2003; Schwab & Lew-Williams, 2016). It is important to note that different studies use different variables to measure quality and quantity of language exposure, including varying means to establish socioeconomic status. Therefore not all studies are comparable. However some interesting findings can be made from an overlook of past studies.

Initially, the introduction of language processing tasks as a means of language assessment was thought to mitigate the correlation between task performance and the quality and quantity of linguistic input the child received prior to the assessment. Researchers argued that language processing tasks, such as nonword repetition and digit span recall, were not impacted by compounding factors, essentially socioeconomic related variables (Dollaghan &

Campbell, 1998; Engel, Santos, & S. E. Gathercole, 2008; Law, McBean, & Rush, 2011), as they did not require previous language knowledge. These tasks were seen as a 'clean' representation of underlying generative processing mechanisms bypassing grammatical and vocabulary knowledge.

Yet recent data has contradicted this argument to various degrees. In a study conducted in Chile among 126 typically developing Spanish monolingual children aged 5;6, Ballandares, Marshall and Griffiths (2016) found that while one language processing task, recalling sentences (see section 3.1.3), was impacted by socioeconomic status, NWR was not. In contrast, a relatively large study conducted in the United Kingdom found that Children from low socioeconomic status scored significantly lower than peers from mid-high socioeconomic status on tasks believed to require minimal exposure and language knowledge. Participants included 387 monolingual English speaking children aged 3;6-5;0, 219 from low socioeconomic status and 168 from mid-high socioeconomic status. Researchers concluded that all tasks, which included language processing tasks (word/nonword repetition and recalling sentences), were influenced by socioeconomic status (Roy & Chiat, 2013).

In addition, extensive research has found that children are more accurate at repeating nonwords when the task consists of nonwords that sound like real words, contain real morphemes of the test language and consist of sounds and combinations of sounds with high phonotactic probability (Chiat, 2015; Ieclercq, Maillart, & Majerus, 2013; Jones, Tamburelli, Watson, Gobet, & Pine, 2010; Messer, Leseman, Boom, & Mayo, 2010; Sharp & V. C. M. Gathercole, 2013). Hence experience with and knowledge of the phonology and vocabulary of the test language contribute to repetition accuracy. This conclusion is even more applicable to performance on RS tasks, since repeating a sentence requires experience and knowledge of all language domains (Polišenská, Chiat, & Roy, 2015).

Based on this evidence from monolingual children and the theoretical approaches discussed above, researchers have hypothesised that language input would influence multilingual language development similarly to the impact on monolingual children (De Houwer, 2009). Generally, previous research has confirmed this assumption, however, relative to monolingual research, evidence is scarce and results are mixed.

2.7 Factors Influencing Multilingual Language Development

2.7.1 Socioeconomic Status

Scheele et al. (2010) investigated the impact of socioeconomic status on language proficiency among Turkish-Dutch and Moroccan-Dutch DLLs. They found a positive correlation between parent education and job status and receptive vocabulary among Moroccan-Dutch DLLs, but no correlation for Turkish-Dutch DLLs. A similar discrepancy was found in Armon-Lotem et al.'s study (2011). They found that Russian-German DLLs performance on standardized tests in their second language (German) was strongly correlated to socioeconomic status measured by parent education and occupation, but no correlation was found between the same factors and Hebrew-Russian DLLs who were tested in their second language (Hebrew). More conclusive results were found by Golberg, Paradis and Crago (2008) who reported a correlation between maternal education and lexical development among DLLs from various linguistic backgrounds. A recently published large scale study by Rojas et al. (2016) supports this correlation showing that maternal education level predicted Spanish-English DLLs expressive language skills in English.

To date, few studies have explored the impact of socioeconomic status on language processing tasks among DLLs. Findings so far are generally in line with monolingual findings. Chiat and colleagues (2013) reported that socioeconomic status impacted the performance of typically developing Russian-German DLLs on a RS task conducted in the second language, German. DLLs from lower socioeconomic status (measured by years of maternal education) scored significantly lower than DLLs from higher socioeconomic status and the later were more likely to score within the monolingual norms than the former. Socioeconomic status was also reported to influence performance on a NWR task. However not all NWR tasks were impacted in the same manner. Chiat & Polišenská (2016) found that of the three tasks they used, only the language specific NWR task was marginally correlated to socioeconomic status among DLLs from diverse linguistic backgrounds (mainly Spanish and Turkish).

Within the multilingual context, additional factors may impact the quantity and quality of language experience in the test language beyond socioeconomic related factors. As multilingualism inherently means input is divided between languages, recent research (Armon-Lotem & Ohana, 2016; Armon-Lotem et al., 2011; Chiat et al., 2013; J. Paradis, Emmerzael, & Sorenson Duncan, 2010; Thordardottir, 2011) has addressed the following factors that might influence the quantity of input in a specific language: (1) age of first exposure, (2) length of exposure, (3) amount of current exposure, and (4) language use in the home. Note that these factors are closely linked and influence one another. For example, age of first exposure is usually linked to length of exposure, since children who are exposed at a younger age necessarily have longer length of exposure to a specific language than same age peers with a later age of first exposure.

However the fact that heterogeneous results among DLLs could not be explained solely by the factors above (Blom & Baayen, 2013; Chondrogianni & Marinis, 2011; Hammer et al., 2012; S. A. S. Lee & Gorman, 2013; J. Paradis, Nicoladis, Crago, & Genesee, 2011), or by theoretical accounts of usage based acquisition (Bybee, 2010), led researchers to conclude that the impact of input went beyond quantitative variables. Qualitative factors, unique to multilingual language development, have recently been explored, specifically parent's self-rated proficiency in the languages they speak and language richness (i.e. which activities and how many native speakers are included in an experience in a specific language). Lastly, language knowledge from one language can impact performance in the other. Similarities and differences in language typology between the mother tongue and target language can either boost or interfere with performance in the second language and vice versa (J. Paradis & Gruter, 2015; Yip & Matthews, 2007). For example, differences between English and Dutch grammar have been shown to interfere with performance in Dutch. In Dutch, gender marking of definite determiners is obligatory, while in English definite determiners do not have grammatical gender. Unsworth (2015) showed that typically developing English-Dutch DLLs made more errors on this specific grammatical feature compared to monolingual Dutch children. Hence, cross linguistic influence negatively impacted performance in the second language. On the other hand, differences in syllable structure in English and Spanish may boost performance in English. Words in Spanish tend to consist of more syllables than words in English. Enhanced performance of typically developing Spanish-English DLLs on nonword repetition tasks in English (i.e. the ability to correctly repeat longer nonwords) has been linked to language experience in Spanish. Hence, cross linguistic influence positively impacted performance in the second language (Summers et al. 2010).

2.7.2 Language Exposure

Thordardottir (2011) argues that amount of previous and current language exposure is the strongest predictor of language development among DLLs. She demonstrated this in a study that included monolingual French/English children and simultaneous French-English DLLs. Input in each language was converted to percentages on a five point scale: 0%, 25%, 50%, 75%, 100%. Results showed that DLLs were more likely to be correctly identified with a language disorder if tested in a language in which they received 50% or more input. When input was below 25% their scores were 1-1.5 SD below the mean even if they did not have a language disorder. Similarly, Paradis (2010) compared language outcomes of three groups of simultaneous English-French DLLs based on the relative use of each language at home: (a) predominantly French, (b) equal use of both languages, and (c) predominantly English. The children who received balanced exposure to both languages at home performed within the monolingual norms on tests in both languages. However the children in the other groups were more likely to score within the normative range only on the tests in the language they primarily used at home. Similarly Gillam Peña, & Perez (2013) found that sequential TD Spanish-English DLLs in their first year of primary school were more likely to be misdiagnosed with language impairment if they received less than 30% input in the test language (English) or were exposed to English for less than a year. Recently Armon-Lotem and Ohana (2016) conducted a study that further strengthens the association between quantity of input and language outcomes. In a study of 40 Hebrew-English 2 and 3-year-old DLLs, it was found that children produced relatively more words in the language they were more frequently exposed to in different environments.

With regards to length of exposure, Armon-Lotem and colleagues (2015) found different results for different language combinations using a screening language assessment in the participants' second language (Hebrew). Length of exposure accounted for a significant proportion of variance of language scores for Hebrew-Russian TD DLLs, but not for Hebrew-English TD DLLs. In addition, Hopp (2011) reported that length of exposure was strongly correlated to performance on grammar tasks in German among 60 TD DLLs aged 3;5 to 7;0 years old from diverse linguistic backgrounds. Additional results from Hebrew-Russian DLLs show that children with more than 4 years exposure to Hebrew (their second language) perform within monolingual norms (less than 1 SD below the cut-off point) with respect to both total and subtest scores (Armon-Lotem et al., 2011). DLLs with less than two years exposure fell more than 1.5 SD below the norm. However, those with 2-4 years exposure fell below the monolingual norms on a vocabulary task but stayed within the norm on imitation tasks including a recalling sentences (RS) task.

Pertaining to language processing tasks, similarly to monolingual studies, multilingual studies have shown that levels of knowledge and experience in the test language impact results. Hence an overlap between typically developing (TD) DLLs and monolinguals with primary language impairment (PLI) may occur. Kohnert et al. (2006) compared three groups of children aged 7;10 to 13;11: TD Spanish-English speakers, TD English speakers and English speakers with PLI. Children were tested using NWR tasks in English. Results showed that although both TD groups' means were higher than the PLI mean, Spanish-English children's average score was lower than the monolingual average and some individual scores overlapped with scores from the PLI group. Researchers concluded that previous exposure to the task's phonology improves performance. In addition, S. A. S. Lee and Gorman (2013) found that a child's mother tongue sound inventory impacted their ability to imitate nonwords in their second language. Hence limited experience with specific sounds could negatively impact performance. These results are in line with other multilingual studies that support the effect of exposure on NWR tasks (Gibson A. et al., 2015; Sharp & V. C. M. Gathercole, 2013; Summers, Bohman, Gillam, Peña, & Bedore, 2010).

A further language processing task is recalling sentences (RS) (see 3.1.1). These tasks are more dependent on language knowledge than NWR and therefore it would be expected that they are more influenced by exposure variables (Thordardottir, 2015a). However, to date, few studies have explored this connection in the context of DLLs. Chiat and colleagues (2013) reported the effects of age of first exposure and length of exposure on RS. Results showed a significant correlation between age of first exposure and a trend towards length of exposure, with longer exposure correlating to higher scores amongst Hebrew-Russian and Hebrew-English children. Yet no significant correlations between these variables and scores on RS tasks were found amongst Russian-German or Turkish-English children. Armon-Lotem and colleagues (2011) reported additional information on the Russian-German and Russian-Hebrew cohorts reported by Chiat and colleagues (2013). They found that for both groups, children with more than 2 years exposure to the test language scored within the monolingual norms and concluded that RS tasks can be considered a valuable measure for identification of PLI among DLLs even when using monolingual tasks and norms.

Thordardottir and Brandeker (2013) continued to explore the effect of language exposure on language processing tasks by conducting a study using NWR and RS tasks on 84 five-year-old typically developing children. This group was divided into five subgroups according to the amount of current relative exposure to English and French: 16 monolingual English-speakers, 16 bilingual English-French speakers dominant in English, 13 bilingual English-French speakers with equal exposure to both languages, 20 bilingual English-French speakers dominant in French and 19 monolingual French speakers. Each bilingual group was tested in both languages and each monolingual group in their mother tongue using NWR and RS tasks. Results showed both tasks were less affected by amount of exposure than were the vocabulary measures (expressive and receptive standard vocabulary tests in English and French) used in an earlier study with the same participants (Thordardottir, 2011). NWR was less sensitive to exposure than RS and thus judged to be less culture and content dependent. Thordardottir further developed this area by researching the relationship between bilingual exposure and morphosyntactic development (Thordardottir, 2015b). She found that grammatical development was strongly correlated to exposure to the target language, which could explain why RS is more vulnerable to exposure than NWR.

2.7.3 Linguistic Environment

Linguistic environment relates to the natural language context an individual is exposed to in daily life. Based on the correlation between quality of input and outcome measures among monolingual children, previous research has shown a correlation between variables associated with DLL's linguistic environment and outcome measures. Paradis (2011) found that language richness, defined as complexity and variation of input gained through language and literacy activities and exposure to native speakers, accounted for the variance in language outcomes among DLLs from a range of language backgrounds. The study included DLLs aged 4;10 to 7;0 in Canada. This correlation was further demonstrated by Blom and Vasic (2011), who found that Turkish-Dutch DLLs aged 6-9-years-old presented with grammatical errors that mirrored those made by adults in their community.

Chondrogianni and Marinis (2011) included maternal self-rated proficiency in English (the test language) as a possible input factor. Results of 43 Turkish-English sequential DLLs aged 6;1 to 9;8 reflected previous research in that variance in vocabulary and some grammar tasks could be attributed to length of exposure and age of first exposure. Multiple regression analysis that took into account multicollinearity effects showed that mother's self-rated proficiency in English also contributed to variance in results in addition to length of exposure and age of first exposure to English. Moreover, mother's self-rated proficiency was the strongest predictor of inability to reach monolingual norms. Hence many of the DLLs who did not reach monolingual norms had mothers with low English proficiency. Similar results were found in a large scale study conducted by Hammer and colleagues (2012) on 191 Latino families and their 4-5 year old children in the US. They showed that maternal English proficiency was a significant predictor of performance on vocabulary and story recall tasks in English. To date, the impact of factors related to linguistic environment on performance on language processing tasks has not been reported.

2.7.4 Language Typology

Languages can be classified according to their typology which is a linguistic classification based on structural and functional features. Similarities or differences in language typology between the child's mother tongue and second language have also been shown to influence performance on tasks in the second language. Paradis (2011) revealed that accuracy on verb morphology tasks was impacted by whether tense and agreement were marked in the child's mother tongue. Similarly Blom and Paradis (2013) found that mother tongue typology predicted accuracy on a specific morphological element in English- past tense marking. Blom and Baayen (2013) found that DLLs whose mother tongue had rich verbal morphology outperformed those with relatively scarce verbal morphology on verbal inflection tasks in Dutch. They concluded that children whose mother tongue has verbal inflections use this morphological feature as a cue to encode sentences and therefore are more attentive to verbal inflections in their second language than children whose mother tongue does not have verbal inflections. These findings, in addition to further evidence (Blom, Chondrogianni, Marinis, & Vasic, 2016; Blom, Paradis, & Duncan, 2012; Meir, Walters, & Armon-Lotem, 2015), demonstrate that the way morphosyntactic features are configured in each language impacts DLL's grammatical accuracy in their second language and thus impacts performance on tasks that require grammatical knowledge. One task that requires such knowledge is the RS task. Meir and colleagues (2015) analysed the error patterns of typically developing 6-year-old Russian-Hebrew DLLs and found that errors were generally traceable to cross-linguistic influence of morphosyntactic features.

Yet another typological difference between languages is phonology. Languages vary in regard to the speech sounds (phonemes) that are used and the frequency of each sound. In addition languages vary in regard to permitted sound sequences and syllable structures. Sorenson Duncan and Paradis (2016) used an English NWR task to show that children whose mother tongue had limited variety and frequency of phenomes in the final position within a syllable (coda) compared to English, were less accurate at imitating phenomes in this position.

Within the heterogeneous multilingual context it is important to consider the unique factors that contribute to the individual differences of DLLs. It is evident that quality and quantity of input impact the language development of typically developing DLLs. However evidence to explain discrepancies in the factors having the most impact and the impact on different language tasks is still lacking.

To this point, group differences between monolingual and multilingual languages have been discussed, as well as the factors that influence the individual differences between typically developing children from both groups. Yet the main focus of this thesis is language impairment among DLLs. To this end, the next section will define a sub-set of children who do not develop language according to the typical monolingual or multilingual language trajectory.

CHAPTER 3- LANGUAGE IMPAIRMENT IN A MULTILINGUAL CONTEXT

With Meisel's (2011) language development characteristics in mind, a sub-group of children may not show ultimate success in acquiring language, rate of acquisition may be delayed, and the pattern of acquisition may be atypical. In past research and clinical caseloads, different terms are used to describe this sub-group of children: *primary language impairment (PLI), specific language impairment (SLI), expressive/receptive language delay, late talkers and language-based learning disabilities.* The term primary language impairment (PLI), will be used throughout this thesis despite a prevalent use of SLI in many publications, to reflect the theoretical framework discussed above (see section 2.4) and the most current research in the field that views language impairment as an integrated part of general development rather than a specific or independent domain (see Leonard, 2014 for review). PLI will first be addressed in general followed by specific reference to the multilingual context.

3.1. Primary Language Impairment (PLI)

PLI is defined as a chronic deviation from normal language development in the absence of psychological, neurological and physical impairment; language being the primary area affected although some secondary deficits may be found in other areas (Leonard, 2014). Hence language is the main area of concern, yet the impairment may not be specific to language. The impact of PLI in later life may cause psychological and academic difficulties and affect participation in society (Bishop & Leonard, 2000; Leonard, 2014). The general prevalence of longstanding PLI ranges between 6-10% (Law, Boyle, Harris, Harkness, &

Nye, 2000) with Australia-specific studies reflecting between 2-19% of the population (McLeod & Harrison, 2009).

Comprehensive research over the last two decades has pinpointed difficulties in morphosyntactic computing and phonological short term memory as two prominent symptoms of PLI (Leonard, 2014). A small vocabulary compared to similar aged peers, deficits in comprehension (Ellis Weismer & Hesketh, 1996) and word finding deficits (Kail & Leonard, 1986) are also indicative, yet are regarded as less prominent (Rice & Hoffman, 2015).

3.1.1. Phonological Short Term Memory

It is hypothesized that limited information-processing capacity accounts for many of the deficits characterized in PLI (Leonard, 2014). Capacity to process information correctly in a timely fashion requires sufficient cognitive space, energy and time (Medwetsky, 2011). As discussed in section 2.4.2, when new auditory information is received, it is held in the short term memory in the phonological loop where it is kept unchanged long enough to be stored more permanently in long term memory. If capacity in the phonological memory is not sufficient to retain the information long enough, the information will be stored partially or not at all.

Theoretically, limited capacity in phonological short term memory may contribute to a cascading sequence of language deficits. When phonological memory lacks the capacity to ensure that the phonological representation of a new word is stored in the long term memory, more encounters with the new word will be necessary to complete the learning process. This will require greater exposure to the new word, resulting in limited vocabulary compared to peers that reach the same size vocabulary in a shorter period of time (Jackson, Leitao, & Classen, 2016; Leonard, 2014).

A deficit in phonological memory has also been linked to morphology; a specific area of difficulty among children with PLI. According to Leonard and colleagues (1997), children with PLI do not have sufficient capacity to perceive word-final consonants (e.g. ‘s’ as in dog*s*) and weak non-final syllables (e.g. ‘un’ as in *un*well) when such forms have a morphological role. This theory is known as the surface account. The challenge of identifying and registering the grammatical function of these forms while processing other information required to understand input has a taxing effect resulting in frequent omission of inflections and function words.

3.1.2. Grammatical Computation

The surface account is most readily applicable in languages such as English where grammatical forms are presented using short or rapid phonetic forms, thus placing stress on the phonological memory. However, difficulties in morphosyntax have been well documented in many cross-linguistic studies (for reviews see Leonard, 2014; Schulz & Friedmann, 2011) including a wide range of grammatical systems that differ from English. These studies gave rise to the development of another theory: the Morphological Richness Account (Leonard, 2014). According to this account, children with PLI struggle to hypothesize, retain and retrieve correct grammatical forms when the number of dimensions required in the specific grammatical system exceeds capacity. Essentially, an additional hallmark of PLI to limited phonological short term memory, is difficulties computing (encoding and manipulating) grammatical features of language regardless of its phonological features.

3.1.3. Identification of PLI

Despite the fact that PLI is a neurodevelopmental disorder with strong genetic correlations (Kalnak, Peyrard-Janvid, Forssberg, & Sahlen, 2014; Rice, 2013), it is diagnosed using a combination of inclusionary criteria (e.g. delayed onset of first words) and exclusionary criteria (e.g. the existence of a syndrome). Identification is based on lower than expected language performance compared to similar age peers (Kohnert, 2013; Rice, 2004). To evaluate a child’s language abilities, assessments include gathering a detailed case history and information on comprehension and production modalities across different language domains. Often both informal and formal tests are used (Owens, 2014; Paul, 2012). The following sections will briefly cover some common approaches to assessment of PLI.

3.1.3.1. Standardised Norm Referenced Assessment

Standardized norm referenced assessments are the most common tools used when clinicians need to come to a diagnostic decision. These assessments facilitate the comparison between the child being assessed and her peers in a timely and reliable manner. Most standardized tools provide clinicians with static information, that is, the clinician sees a ‘snap shot’ of the child’s areas of relative strength and difficulty as presented at a specific point in time. The child’s proficiency or skill in these domains is gauged on a scale of normative data. It is not surprising then, that Peña, Spaulding and Plante have referred to standardized assessments as the “cornerstone of diagnosis in the field” (2006, p. 247) and Stow and Dodd described them as “the tools of the profession” (2003, p. 363). In addition, the use of standardized assessments allows one to convert scores to percentiles which provide common terminology to support collaboration with other health professionals including general practitioners, occupational therapists and others. As a consequence, many service providers require standardized assessments to determine eligibility and funding (for example: Department of Education, 2015).

Yet the use of these tests has raised diagnostic dilemmas among monolingual children. Foremost is the determination of cut-off scores for diagnostic purposes. As a score below the cut-off point reflects impairment, the selection of a specific cut-off point could drastically change the proportion of children identified with language impairment. The following will address terms and analyses that are used in the field of speech-language pathology to determine diagnostic accuracy.

Sensitivity (true positive) and specificity (true negative) values may be used to report diagnostic accuracy. The use of these values is increasing, but still not universal. Sensitivity indicates the proportion of participants with a disorder (e.g. PLI) identified by a diagnostic test. Specificity indicates the proportion of participants without a disorder (e.g. TD) who were not identified using a diagnostic test. Overall accuracy combines both sensitivity and specificity to reflect the proportion of individuals who were correctly identified. In the field of speech pathology, independent judgments by experienced clinicians, current attendance in language treatment, and use of different standardized tests with adequate diagnostic accuracy, are considered the “gold standard” to determine diagnostic accuracy of a particular assessment procedure or approach (Leonard, 2014). Hence true positives and negatives are determined when test results and the gold standard measurement reflect the same diagnosis.

Ideally, 100% sensitivity, specificity and overall accuracy would be achieved, with the result that all individuals would be correctly identified. However, as language assessments are behavioural, this rarely occurs in practice. As sensitivity and specificity reflect true positive and true negative values, there is a trade-off between the two. Higher sensitivity may result in lower specificity which might result in over-diagnosis of a disorder. Conversely, higher specificity with lower sensitivity may result in under-diagnosis. Therefore, optimal accuracy occurs when both sensitivity and specificity are relatively high. Plante and Vance (1994) proposed that standardized language tests have a minimum of 80% sensitivity and specificity, with levels above 90% indicative of good discriminant accuracy.

Recently an additional measure, likelihood ratios, has been argued to provide a more reliable indication of diagnostic accuracy than sensitivity and specificity values (Leonard, 2014). Likelihood ratios determine the probability of a true positive or true negative result. Positive likelihood ratios (LR+) reflect the probability of 'ruling in' a disorder. This computes the likeliness of a true positive (e.g. how many times more likely a DLL with PLI will score below the cut-off point than a TD DLL). Negative likelihood ratios (LR-) reflect the probability of 'ruling out' a disorder. This computes the likeliness of a false negative (e.g. how many times more likely a DLL with PLI will score above the cut-off score than a TD DLL) (Portney & Watkins, 2009).

Likelihood ratios above 10.0 and below 0.1 (LR+≥ 10.0; LR-≤0.1) indicate high diagnostic accuracy. Ratios above 3.0 and below 0.3 (LR+≥ 3.0; LR-≤0.3) indicate suggestive accuracy and ratios between 0.3 and 3.0 (LR+< 3.0; LR->0.3) do not allow ruling a disorder in or out with confidence (Dollaghan, 2007). In subsequent sections, if published data from previous studies consists of sufficient information to calculate likelihood ratios or likelihood ratios were calculated by the author, they will be noted.

The selection of cut-off points with the best diagnostic accuracy cannot be generalized from test to test and from sample to sample (Leonard, 2014; Thordardottir, 2015a). Some tests will show high diagnostic accuracy at 1SD below the cut-off point (Conti-Ramsden, Botting, & Faragher, 2001; Thordardottir et al., 2011), while others show greater accuracy using a lower threshold (Tomblin, Records, & Zhang, 1996; Wiig, Semel, & Secord, 2013), and these thresholds could vary with different samples.

3.1.3.2. Language Processing Tasks

Language processing tasks are also used for PLI assessment and are often included in standardized omnibus assessments (e.g. CELF-4: Semel, Wiig, & Secord, 2006; CTOPP: Wagner, Torgesen and Rashotte, 1999). Language processing tasks aim to reveal the child's underlying language learning mechanisms, or how the material is mentally manipulated, rather than depict a snapshot of accumulated language knowledge. As a result, they have become a valuable tool in the assessment of language skills in DLLs. In the current study two language processing tasks were used: nonword repetition (NWR) and recalling sentences (RS). As their names suggest, in NWR tasks the child repeats a nonword (a string of sounds that is not a real word) and in RS tasks the child imitates a sentence.

In NWR tasks children repeat nonsense words consisting of phonemes and phonotactic structures present in the test language. These forms gradually increase in syllable length. Verbatim repetition of new auditory information taps into short term memory and reflects the capacity of long term memory and the inherent thresholds of the items (Baddeley, 2000). Impaired short term memory is a hallmark of PLI (Leonard, 2014). When conducting NWR tasks the researchers or clinicians can observe a child's attempt to learn new syllables in real time. NWR tasks tap into the child's ability to acquire new linguistic information from the environment. Based on psycholinguistic language processing theories, we create an opportunity for the child to hear new auditory input, retain it in the phonological short term memory and repeat it using phonological and semantic knowledge in the long term memory that produces a plan for the articulatory system.

A large range of studies in different languages around the world have shown that NWR tasks can distinguish monolingual children with PLI from their typically developing (TD) peers. Some examples include: English (Coady & Evans, 2008; S. E. Gathercole, 2006), Swedish (Kalnak et al., 2014), Italian (Casalini et al., 2007; Dispaldro, Leonard, & Deevy, 2013), Spanish (Girbau & Schwartz, 2007; Guiberson, Rodríguez, Nippold, & Troia, 2013), Dutch (de Bree, Wijnen, & Gerrits, 2010) and Arabic (Shaalan, 2010). However, in two studies in Cantonese (Liu et al., 2010; Stokes, Wong, Fletcher, & Leonard, 2006) NWR failed to distinguish children with PLI. This may be attributed to the relatively simple phonological structure of Cantonese at the phoneme level compared to other languages mentioned above (Liu et al., 2010).

Another language processing task argued to be a stronger clinical marker of PLI amongst monolingual children by Conti- Ramsden (2001), is the recalling sentences (RS) task. This task requires children to imitate sentences that increase in morphosyntactic complexity and length. In order to repeat the sentence correctly a child must encode the sentence and then reconstruct it using internal linguistic mechanisms (Potter & Lombardi, 1998). Hence, imitation is confined to the language knowledge of the child (Polišenská, 2011; Seeff-Gabriel, Chiat, & Dodd, 2010), and her grammatical computation capacities. Limited grammatical computation capacities is a hallmark of PLI (Leonard, 2014).

Polišenská, Chiat and Roy (2015) recently conducted a detailed study exploring how different types of linguistic knowledge affect performance on RS tasks. Participants included 100 monolingual children: 50 English-speakers and 50 Czech-speakers, aged 4 and 5 years. Each group was tested separately using RS tasks in their native language alone. When pseudo or nonwords were used, they were based on either English or Czech phonology or syllable structure respectively. The use of two languages allowed the researchers to explore the effects of different linguistic conditions on performance when comparing the groups. Stimuli included seven linguistic conditions: (1) grammatical sentence, (2) sentence with unnatural prosody, (3) semantically implausible sentence, (d) ungrammatical sentence, (e) pseudo-sentence with content words replaced by nonwords, (f) pseudo-sentence with function words replaced by nonwords, and (g) pseudo-sentence with all words replaced by nonwords. In each condition participants repeated blocks of successively longer sentences. Results revealed that recall capacity varied according to the child's linguistic knowledge. All types of linguistic knowledge (morphosyntax, semantics, and prosody) affected recall accuracy, however higher recall capacity was generated from sentences with correct grammatical structure in both English and Czech leading the authors to conclude that morphosyntax had more influence on repetition than familiarity of lexical items and their semantics and prosody. It can be suggested, therefore, that this task provides information on the individual's grammatical computation and is less affected by exposure to vocabulary.

RS tasks have also been proven accurate in identifying monolingual children with PLI over a range of different languages including English (Conti-Ramsden et al., 2001; Seeff-Gabriel et al., 2010), French (Thordardottir et al., 2011), Arabic (Shaalan, 2010), Vietnamese (Hoang, Schelstraete, Tran, & Bragard, 2014) and Cantonese (Liu et al., 2010; Stokes et al., 2006). Interestingly, RS tasks were successful in Cantonese, the only language to date that showed no discrimination power for NWR tasks. It is important to keep in mind that RS tasks are more dependent on language knowledge than NWR tasks and therefore depends more on language experience that may be lacking in DLLs.

3.1.3.3. Parent/Caregiver Questionnaires

In addition to direct assessment, past research supports the use of parent/caregiver questionnaires for identifying language impairment in monolingual children. This approach is cost effective and relatively time efficient and can serve as a supplement to direct assessment.

A study in Iceland compared language outcomes collected via direct assessment to indirect data retrieved through parent interviews for 81 3-6-year old monolingual Icelandic speaking children (Gudmundsson & Gretarsson, 2013). They found a strong correlation between both forms of assessment concluding that parent-based assessment tools of child language development are a viable possibility for both research and clinical purposes. Similar results were found in Cantonese (K. Lee, Chiu, van Hasselt, & Tong, 2009) among 3-6 year old children with hearing impairment. Fair to good agreement was found between parent report and direct assessment. In addition earlier publications exploring the efficiency and accuracy of identification of language impairment among pre-school children using parent questionnaires in English speaking countries (Squires, 1996; Thal, O'Hanlon, Clemmons, & Fralin, 1999) have shown correlations between parent report and direct assessment. Notably, in a longitudinal study that included 93 children with current or a history of PLI, Conti-Ramsden and colleagues (2006) found that parents were less accurate identifiers of PLI when the impairment was mild.

3.1.3.4. Combination of assessment tools

Given the strengths and weaknesses of the tools used to identify PLI among monolingual children, Bishop and McDonald (2009) explored the diagnostic accuracy of a combination direct and indirect assessment tools among 9-10-year old monolingual children in the UK. Parents of participants completed the Children’s Communication Checklist (CCC-2: Bishop, 2003) and children underwent a battery of standardized tests that included receptive and expressive language tasks and language processing tasks. Results showed that the best combination of sensitivity and specificity (84% and 59% respectively) was obtained when at least one measure from both the language battery and the CCC-2 subtests were more than 2 SD below the mean. In addition, the strongest predictors of PLI were tasks and sub-tests that reflected short-term memory function. Specifically RS, NWR, imitation of oromotor sequences, understanding directions and memory for names. Despite not reaching minimal diagnostic criteria of over 80% sensitivity and specificity (Plante & Vance,1994), researchers concluded that integration of direct and indirect assessment tools provide an objective diagnosis if direct tasks include those that act as good PLI markers (i.e. reflect short term memory function). Hence, a combination of tools can be more accurate than the use of standardized assessments alone.

Thus far it has been established that children who present with a chronic language deficit alongside otherwise typical development in most developmental domains are classified as having a primary language impairment (PLI). Diagnosis is based on exclusion and inclusion criteria and data elicited by a range of assessment approaches including standardized tests, language processing tasks, parent questionnaires and a combination of assessment tools. It is established that impaired phonological short term memory and grammatical computation are two main hallmarks of PLI (Bishop & McDonald, 2009; Conti-Ramsden, 2003; Leonard, 2014; Schulz & Friedmann, 2011). With monolingual language impairment as a baseline for comparison, the following section will discuss PLI among DLLs.

3.1.4. Primary Language Impairment and Multilingualism

Based on the theoretical framework and cross linguistic evidence of multilingual language development discussed in sections 2.4 and 2.5.2, it has been recognized that the underlying mechanisms and psycholinguistic processes are shared across monolingual and multilingual minds. Therefore deficits due to PLI are expected to affect internal language foundations of both DLLs and monolingual children in a similar manner. With respect to DLLs, underlying difficulties are anticipated to impact all languages the child speaks (Gutierrez-Clellen & Simon-Cereijido, 2009; J. Paradis, Genesee, et al., 2011) and prevalence is expected to reflect the same proportion as monolingual children (Bedore & Peña, 2008;

Engel de Abreu, Cruz-Santos, & Puglisi, 2014). In other words, language impairment is not language specific and DLLs with PLI are expected to demonstrate comparable underlying difficulties to their monolingual peers with PLI (Bedore & Peña, 2008).

Despite the theoretical clarity on the similarities between multilingual and monolingual PLI, on the surface PLI may look different across languages. Different languages map forms into meaning in different ways, therefore the same child may present language deficits differently across languages. For example, a complex grammatical aspect of Hebrew and therefore an area of difficulty that can be used as a hallmark of PLI is subject-verb agreement (Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999). In English this is not a hallmark of PLI since subject-verb agreement does not require use of complex morphology and thus does not test the capacity of grammatical computation abilities. A Hebrew-English DLL with PLI may present with more subject-verb agreement errors in Hebrew than English. From a diagnostic point of view this may incorrectly appear like a deficit in grammar is present in only one language.

As discussed in section 2.7, DLLs have additional factors that influence the quality and quantity of language input, such as length of exposure to languages spoken at home and in the environment and parent proficiency in these languages. As language experience impacts language skill, exposure differences may disadvantage DLLs with PLI in comparison to monolingual children with PLI when assessed in either language (De Lamo White & Jin, 2011; Kohnert, 2013; J. Paradis, Genesee, et al., 2011).

Thus far language and language impairment have been discussed in reference to both the mono- and multilingual contexts. The next section will turn to the challenging task of identification of PLI among DLLs.

3.2. The Challenges of Assessing DLLs

Early assessment of and intervention for language difficulties is an effective way to address language impairment and minimize long term psychosocial and academic difficulties (Durkin, Conti-Ramsden, & Simkin, 2011; Leonard, 2014; Ullrich, Ullrich, & Marten, 2014; van Agt, Verhoeven, van den Brink, & de Koning, 2011). While awareness of multilingual issues is high among Speech Language Pathologists (SLPs) in Western countries (Mennen & Stansfield, 2006; Williams & McLeod, 2012), providing equitable services to all clients regardless of their linguistic and cultural background is still a challenge (Verdon et al., 2014). Specific difficulty is evident for initial assessment, when SLPs are required to disentangle multilingualism from language impairment and provide a valid diagnosis as a base for effective treatment (Armon-Lotem, de Jong, & Meir, 2015; De Lamo White & Jin, 2011).

The challenges associated with diagnostic assessment of DLLs can lead to both over and under-identification of language impairment (Bedore & Peña, 2008; Kohnert, 2013; Roseberry-McKibbin, Brice, & O'Hanlon, 2005). A disproportionate number of DLLs are in special education settings and receive speech-language therapy (Klinger & Artiles, 2003; J. Paradis, Genesee, et al., 2011) while others are delayed or denied intervention (Morgan et al., 2015; Ziethe, Eysholdt, & Doellinger, 2013). At times, a crucial window for intervention may be missed due to failure to identify a language difficulty (Kohnert, 2013; Samson & Lesaux, 2009). Recent research suggests that under-diagnosis is more prevalent than the former (Morgan et al., 2015; Tuller et al., 2013).

Many SLPs in the US, UK and Australia find it difficult to answer the main diagnostic question at hand: is the child’s language different due to multilingualism or is there an underlying language impairment that requires intervention? A mismatch between the characteristics of SLPs and their clients, based on linguistic background, socioeconomic status and culture, creates barriers when collecting diagnostic information (De Lamo White & Jin, 2011; Verdon et al., 2014; Williams & McLeod, 2012). Most SLPs in these countries are monolingual English speaking, white, middle-class women (De Lamo White & Jin, 2011; SPA, 2015; Roseberry-McKibbin et al., 2005). Moreover, SLPs have reported a lack of specific training in the area of multilingualism (Caesar & Kohler, 2007; Kimble, 2013) and a shortage of appropriate and accessible assessment tools and normative data (De Lamo White & Jin, 2011; Sánchez, 2006; Williams & McLeod, 2012). Previous studies have reported inconsistent beliefs and practices as to how much exposure is needed before a diagnostic decision can be made and when to recommend intervention (Broomfield & Dodd, 2004; Jordaan, 2008; Kritikos, 2003). Roseberry-McKibbin and colleagues (2005) found that even when changes in university coursework were made to enhance training in the area of multilingualism, lack of assessment tools remained an unsolved issue.

This being the case, we must ask why there is a shortage of assessment tools. One reason is the heterogeneous characteristics of DLLs (Kohnert, 2010; Oliveri, Ercikan, & Simon, 2015; Uziel-Karl et al., 2014; Verdon, 2015). As discussed in sections 2.5.2 DLLs differ greatly from one another in terms of language proficiency and social-cultural and linguistic backgrounds. In the Australian context it is hard to find a group of DLLs who speak the same language, are exposed to a similar quantity and quality of input in all languages, and are from the same social-cultural background. Therefore creating a norm referenced test battery, similar to those available for monolingual children, is highly problematic. Furthermore, even if an assessment for a specific language combination was to be designed in Australia, very few SLPs would be able to use it due to the lack of multilingual SLPs (Verdon et al., 2014).

Another key challenge in identification of language impairment in multilingual children is the linguistic similarities between TD DLLs and monolingual children with PLI. In the early stages of second language learning, it is common to hear a DLL use telegraphic speech, imitation and use of formulaic constructions (Roseberry-McKibbin, 1995). While this linguistic profile indicates low proficiency in monolingual speakers, in the case of DLLs it may represent a stage of typical learning and not impairment. In addition, TD DLLs may present with morphosytactic errors typical of language impairment. Studies conducted in different second languages (English, Hebrew, Dutch and Swedish) showed an overlap between TD DLLs and monolingual children with PLI (Armon-Lotem, 2010, 2012; Håkansson, 2001; J. Paradis, Genesee, et al., 2011; Tuller et al., 2013). However, more in-depth analysis has found qualitative differences. TD DLLs are more likely to use the wrong morpheme, while monolingual children with PLI are more inclined to omit the morpheme altogether (Armon-Lotem, 2012; J. Paradis, Genesee, et al., 2011; Vender, Garraffa, Sorace, & Guasti, 2016). It has also been shown that DLLs show more errors in vocabulary-related tasks (such as similarities and differences, functions and category generation) in their second language compared to their mother tongue (Bedore, Peña, Garcia, & Cortez, 2005).

Previous research has clearly stated that if a language impairment exists, it will present in both or all the languages in which the child functions (Armon-Lotem, de Jong, et al., 2015; Gutierrez-Clellen & Simon-Cereijido, 2009; J. Paradis, Genesee, et al., 2011). Accordingly, the “gold standard” for assessment, advocated by professional Speech Pathology organizations around the world, is to assess all the languages the child uses in daily life (i.e. ASHA 2004; SPA 2009). A deficit across languages is sufficient evidence to suggest the presence of an underlying, cross linguistic language deficit and therefore is the best way to identify hallmarks of PLI without the interference of confounding factors (Armon-Lotem, de

Jong, et al., 2015; Kohnert, 2013). Nevertheless, this recommendation poses many practical difficulties. It requires SLPs to speak more than one language and have access to assessment tools in a range of languages.

In practical terms, assessment in more than one language is uncommon. Jordaan (2008) conducted an international study surveying 99 SLPs from 13 different countries. These SLPs provided information on 157 DLLs in their caseloads. Although the proportion of bilingual SLPs in this study was relatively high, at 26%, only 5 children (3.2%) were assessed in both/all the languages they function in. This data clearly shows a very large gap between the expectations from SLPs based on current theoretical knowledge and the ability of SLPs to put the recommendations into practice.

In light of these challenges, the next section will discuss the approaches SLPs use in practice to assess DLLs. The evidence base of these current assessment approaches will be addressed.

3.3. Current Assessment Approaches

3.3.1. Standardised Assessments

In Australia PLI is often identified using standardised assessments (e.g. Department of Education, 2015). These assessments were built and normed predominantly on monolingual linguistic development trajectories in English (e.g. CELF-P-2, PLS-5).While these assessments do provide instructions and recommendations for use among DLLs (and the samples included a small percentage of DLLs), it is important to note that they are unsuitable for DLLs as they reflect current knowledge in English that could still be in the process of being learned or influenced by the mother tongue (Bedore & Peña, 2008). In addition they are inadequate due to content and cultural bias (De Lamo White & Jin, 2011). In other words, DLLs' performance could be below the norm for reasons other than an underlying language deficit (Caesar & Kohler, 2007; Hoff et al., 2011; Marinis & Chondrogianni, 2010; Tuller et al., 2013; Wiig, Secord, & Semel, 2006).

Despite the relatively large body of evidence in this area, and the awareness of multilingual issues among SLPs in Australia, many still use English-only standardized assessments for DLLs. Specifically Williams and McLeod (2012) reported that these assessments were frequently used by 81.9% (N=86) of SLPs in their sample.

Previous research has demonstrated that TD DLLs could be underestimated if tested in their less dominant language alone (Armon-Lotem et al., 2011; Ben-Zeev, 1977; J. Paradis, 2010; J. Paradis, Schneider, & Sorenson Duncan, 2013; Thordardottir, 2011). Gillam and colleagues (2013) demonstrated that use of standardized tests following monolingual guidelines developed by Tomblin, Records, and Zhang (1996) underestimated the language skills of many TD DLLs. According to these guidelines PLI is diagnosed when a child scores -1.25 SDs below the mean on two or more composites scores out of five composite scores (expressive, receptive, vocabulary, grammar and narration). Gillam and colleagues (2013) derived these scores from performance on a standardised language development test and a narrative test (for more detail see section 6.2.2). This assessment produced 45% specificity and 95% sensitivity. Hence more than half of the TD DLLs in their sample scored below the monolingual cut-off point.

3.3.2. Revised Cut-Off Scores

Gillam and colleagues (2013) attempted to improve diagnostic accuracy by using the same data, but configuring it differently by using alternative cut-off points. Logically, if the cut-off point of tasks were lowered, the proportion of TD DLLs below the cut-off line would decline. The full range of sensitivity and specificity of each possible criterion score was analysed to produced optimal revised cut-off scores using ROC curve analysis (see section 4.5). These scores yielded 86% sensitivity and 68% specificity. Hence although specificity improved, it still did not reach the minimal requirement according to Plante and Vance (1994) and resulted in reduced sensitivity.

Recent research using language processing tasks in the child's second language with revised cut-off scores derived from ROC curve analysis has shown a noteworthy improvement of diagnostic accuracy. Armon-Lotem and Meir (2016) showed that when monolingual cut-off scores were applied to NWR and RS tasks, sensitivity levels were above 95%, but specificity levels were poor (48% and 63% respectively). Use of revised cut off scores improved specificity on both tasks. The specificity of the NWR task increased to 79% (marginally below the minimal requirement) and on the RS task to 89%, while sensitivity on both tasks stayed above 80%. Therefore, adjustment of the cut-off scores brought the RS task to acceptable levels of diagnostic accuracy and the diagnostic accuracy on the NWR to merely 1% below criteria.

3.3.3. Informal Assessments

An alternative assessment often used in practice by SLPs in Australia is informal assessment. Williams and McLeod (2012) reported that 98.2% (N=108) of SLPs used informal procedures in addition to, or as an alternative to, standardized assessment. However this solution relies heavily on the individual experience of an SLP in choosing assessment tasks and analysing results. Informal assessments are hard to regulate and replicate. They cannot function as a widespread solution as validity and reliability is difficult to demonstrate.

3.3.4. Dynamic Assessment

Recent publications have suggested using dynamic assessment; an approach that evaluates the child’s ability to learn rather than current knowledge in a specific language (Hasson, Camilleri, Jones, Smith, & Dodd, 2013; Petersen & Gillam, 2015; Verdon, 2015). This form of assessment is strongly linked to the interactivism approach to language development. The SLP collects information on the child's learning strategies and the number and type of prompts the child needs in order to make changes. The most common framework for this type of assessment is test-teach-retest. Using this process a SLP can teach strategies, and according to evaluation of the child’s ability to change during the teaching phase, discriminate between a difference due to multilingualism and a disorder due to impairment (Peña, 1993). However this approach is time consuming and also relies on the individual experiences of the SLP (Kapantzoglou, Restrepo, & Thompson, 2012).

3.3.5. Translations

Another possible solution could be to translate standardized tests to different languages. However, cross linguistic research has shown that language development milestones differ from one language to the other and important features in one language may not exist in the other (Zimmerman, Steiner, & Pond, 2002). Moreover, cultural differences can cause bias when vocabulary and context reflect that of the origin language (the language in which the test was constructed) and not the target language (the language into which the test is translated) (Laing & Kamhi, 2003).

Thus far the approaches described have different drawbacks, either disadvantaging DLLs or imposing practical limitations. It can be concluded that clinicians need assessment tools that have the power to reveal an underlying language impairment with minimal impact of cultural and linguistic influences and that these tools need to be easily available, realistic in terms of time constraints and able to be used by monolingual SLPs. Current literature includes a small number of approaches that have the potential to meet these requirements. Nevertheless they too may have drawbacks. The following section will discuss these approaches in detail.

3.3.6. Appropriate Reference Groups

Under the premise that DLLs are more likely to present with similar language development patterns to one another than to monolingual children, creating a comparison between DLLs with and without PLI could help minimise assessment bias. Previous studies have shown encouraging support for this approach with clear between group differences between DLLs with and without PLI on a range of different tasks covering many language domains (Armon-Lotem & Meir, 2016; J. Paradis et al., 2010; J. Paradis, Schneider, & Sorenson Duncan, 2013; Restrepo & Gutierrez-Clellen, 2001; Windsor, Kohnert, Lobitz, & Pham, 2010). However, the heterogeneous characteristics of DLLs may make this difficult to achieve. Even if an appropriate reference group is found, this approach should be used with caution taking into consideration that differences may still exist and the impact they may have.

3.3.7. Parent/caregiver Questionnaire

Recent studies have explored the potential of parent questionnaires as an indirect approach to the assessment of DLLs. An important Canadian study by Paradis, Emmerzael and Sorenson Duncan (2010) explored the use of a parent/care-giver questionnaire (Alberta Language and Development Questionnaire- ALDeQ) to identify DLLs with PLI. This tool consists of four sections: early milestones, current first language abilities, behaviour patterns and activity preferences, and family history. The combination of all sections aims to draw a detailed picture of the child's mother tongue profile without any direct assessment. One hundred and thirty nine TD DLLs were compared to 29 DLLs with PLI, with a mean age of 5;7 and a total of nine different mother tongues. PLI status was based on current participation in individual speech-language intervention or special kindergarten programs for children with language delays. Results showed significant differences between the DLLs with PLI and the TD DLLs when total and sub-test scores were compared. Lower scores were associated with PLI and higher scores with typical development. This test showed high specificity (96%), but poor sensitivity (66%), indicating that it is not to be used independently to make diagnostic decisions (LR+ =16, LR- =0.35).

Following Paradis et al., pilot studies incorporating use of the ALDeQ have been carried out in 13 different countries (Tuller, 2015), including Australia (May & Williams, 2012). Some of these studies had participants with diverse mother tongue backgrounds, while others explored specific language combinations (i.e. German-Russian). Most studies used a questionnaire based on the ALDeQ together with additional questions related to the child's linguistic environment (e.g. BIPAQ, Abutbul-Oz, Armon-Lotem, & Walters, 2012).

May and Williams (2012), explored the use of the ALDeQ among 17 DLLs, 3 with apparent language difficulty, in Perth, Australia. Classification of apparent language difficulty was determined based on external clinical identification of language difficulty or current referral to language therapy. Participants had a mean age of 6;7 (range 5;3-8;7) and came from Vietnamese, Romanian and Macedonian backgrounds. They found that TD DLLs scored higher than the DLLs with apparent language difficulty and that the total scores of TD DLLs were consistent with the Canadian norming population. The authors concluded that further research was required to continue to explore the validity of this approach.

Detailed evidence from the remaining 12 pilot studies was not available. However, general information provided by Tuller (2015) shows that reliable information was provided by parents and caregivers from a wide range of cultural and linguistic backgrounds describing their child's language abilities in their mother tongue. These studies suggest that indirect information collected using parent questionnaires has the potential to accurately distinguish between TD DLLs and DLLs with PLI, especially when other means to collect information about the mother tongue are limited.

It is important to state that in most of these studies the parents of children with PLI were aware that their children were seeing a SLP or attending a special education program. Being aware of a diagnosis for PLI might have influenced how they answered questions on the questionnaire (J. Paradis et al., 2010; Tuller, 2015). Tuller (2015) recently reported two studies that aimed to overcome this drawback by interviewing parents before a clinical diagnosis was given. The majority of these children were referred to services due to parent or teacher concern and a small group was randomly selected to complete a parent questionnaire without presenting with apparent difficulties. The first study included 53 DLLs, 41 who had been referred to a private SLP clinic and 12 who had not, aged 2;6-6;6 in Israel. The second study included 30 3-8-year-old DLLs referred by their schools to a university language clinic in Lebanon. Preliminary results suggest that parent report is positively correlated to clinical diagnosis. Parents were able to accurately report early milestones and current language skills regardless of prior knowledge of a diagnosis.

3.3.8. Language Processing Tasks

Language processing tasks tap into the ability to decode acoustic stimuli and encode it using internal linguistic knowledge. Two tasks are of interest: nonword repetition (NWR) and recalling sentences (RS).

3.3.8.1. Nonword Repetition

As described in section 3.1.3.2, a large range of studies in different languages around the world have shown that NWR tasks can distinguish monolingual children with PLI from their typically developing (TD) peers and therefore low performance is seen as a clinical marker for PLI (S. E. Gathercole, 2006).There is debate regarding the use of monolingual NWR tasks in attempting to identify language impairment in DLLs. Some studies show that DLLs perform similarly to monolingual peers while others show that monolingual NWR tasks underestimate TD DLLs. A recent study found no differences in the performance of TD DLLs compared to monolingual Italian-speaking peers on NWR tasks in Italian. Vender and colleagues (2016) conducted a study comparing the performance of 40 TD Italian monolingual children with a mean age of 4;9 to that of 120 age matched TD DLLs from three different linguistic backgrounds: Albanian, Arabic and Romanian. Although the monolingual z-score for monolingual children was slightly higher than the DLLs scores, the difference was not statistically significant leading the authors to conclude that performance was independent of mother tongue influence.

Additional evidence in support of this conclusion can be found in a study that included both TD children and children with PLI. Armon-Lotem and Meir (2016) explored the diagnostic potential of three different repetition tasks: NWR, forward digit span and RS. A total of 230 children aged 5;5 to 6;8 were recruited. Participants were divided into six groups: monolingual Russian-speaking children with and without PLI, monolingual Hebrew speaking children with and without PLI and Hebrew-Russian DLLs with and without PLI. DLLs were identified with PLI if parents, SLPs or teachers reported concerns in regard to language development and if they scored -1.25 SDs below bilingual specific norms on tests in both Russian and Hebrew. Monolingual children with PLI were identified if they attended speech-language therapy or if concerns were raised by parents, teachers or SLPs and scored -1.25 SDs below monolingual norms on a standardized language test in their language (either Hebrew or Russian). Results showed that TD children, regardless of how many languages they spoke, performed similarly to one another and outperformed children with PLI on NWR tasks in both languages.

Similarly, S. A. S. Lee and Gorman (2013), found no significant differences between monolingual English speaking children and TD DLLs on a NWR task in English. They compared 15 monolingual English speakers to DLLs with three different linguistic backgrounds: Korean, Chinese and Spanish. All participants were 7 years old and enrolled in year 1 in English instruction schools.

Moreover, a longitudinal study conducted in Spain among 44 simultaneous Spanish-Catalan children (22 with PLI) supports the use of monolingual NWR tasks among DLLs (Buil-Legaz, Aguilar-Mediavilla, & Adrover-Roig, 2016). In this study DLLs with PLI were identified if concerns were raised by their teachers and they scored below the norm on a standardized language test in their dominant language (either Spanish or Catalan). Results showed that across a six year period: at age 6, 8 and 12, TD DLLs continued to produce significantly higher scores on a NWR task in Spanish than DLLs with PLI. Hence the NWR task was able to distinguish between the groups independent of age, reflecting the chronic deficit related to PLI.

However, other studies have found that TD DLLs scored lower than their monolingual peers on NWR tasks in their second language. Windsor et al (2010), compared 4 groups of children aged 6;0 to 11;6: 65 TD Spanish-English speakers, 69 TD monolingual English speakers, 19 Spanish-English speakers with PLI and 34 monolingual English speakers with PLI. PLI status was confirmed by scoring no higher than 5/10 on subtests of standardized language test (CELF-4: Semel, Wiig and Secord, 2003). Monolingual English speakers with PLI were identified using the CELF-4 English edition and DLLs with PLI were identified using both English and Spanish editions of the CELF-4. Findings showed that in NWR tasks in English, TD monolingual children outperformed TD DLLs in words of greater than 4 syllables. Unsurprisingly, the TD DLLs scored higher than the DLLs with PLI, but no significant difference was found between the TD DLLs and the monolingual children with PLI (LR+ =2.20, LR- =0.09). Results were mirrored using a Spanish NWR task. TD DLLs outperformed the other three groups, both TD groups scored higher than the PLI groups and DLLs with PLI outperformed monolingual children with PLI (LR+ =3.14, LR- =0.52).

Additional studies also found that TD DLLs outperformed monolingual children with PLI, but did not perform as well as their TD monolingual peers (Boerma et al., 2015; Kohnert et al., 2006). Boerma and colleagues (2015) compared four groups of 30 5-6-year-old children from the Netherlands: Dutch speaking monolinguals with and without PLI and DLLs from a range of linguistic backgrounds with and without PLI. PLI status was confirmed by scoring -1.5 SDs below the mean on at least two out of four subscales of standardized language assessment in Dutch. TD monolingual mean scores were significantly higher than TD DLLs mean scores on Dutch specific NWR tasks (LR+= 9.00 LR-=0.14).

However Boerma and colleagues (2015) also explored a unique NWR task developed as part of an international cross-linguistic research network: Cooperation in Science and Technology (COST) Action IS0804. This new task aims to minimize the effects of language exposure by creating 16 nonwords with a simple CVCV sequence using consonants and vowels that occur in many languages. When administered, the assessor produces phonemes in her own mother tongue. Thus, performance may still be influenced by exposure, making the task quasi-universal rather than universal. Boerma and colleges (2015) reported high diagnostic accuracy for DLLs using this task (83% sensitivity and 93% specificity, LR+ =11.9 LR- =0.14) with no significant difference between TD DLLs and monolingual peers. French and English versions of the same quasi-universal NWR task were found to correctly distinguish between TD DLLs and peers with PLI (Chiat & Polišenská, 2016; Tuller et al., 2013).

3.3.8.2. Recalling Sentences

Another language processing task, argued to be a stronger clinical marker of PLI than NWR amongst monolingual children by Conti-Ramsden and colleagues (2001), is the recalling sentences (RS) task. RS tasks require children to imitate sentences that increase in morpho-syntactic complexity and length. Recently a growing body of evidence has shown RS tasks to successfully discriminate PLI amongst DLLs.

As part of a wider cultural group study undertaken during norming, the RS subtest of the Clinical Evaluation of Language Fundamentals Preschool-Second Edition test (CELP-P-2) (Wiig et al., 2006) was tested on TD Australian children from different linguistic backgrounds. Children from three linguistic backgrounds were selected representing three major migrant groups in Australia: Italian (n=11), Chinese (n=21) and Arabic (n=15). The participation criteria in this sub-study demanded that at least one parent was born overseas and that the child had no evidence of language impairment based on teacher, parent or school SLP reports. No substantial evidence that a language other than English was spoken at home was provided, nor information regarding language dominance or length of exposure to English. Participants from the three cultural groups were compared to monolingual English-speaking children matched for age, gender and primary caregiver's education level. For all three groups, mean RS scores were lower than their monolingual peers, but still within the normal range (less than 1 SDs from the monolingual norm). The Italian cultural group showed the smallest mean difference (1.00 below the monolingual mean) while the Chinese and Arab group showed a slightly larger difference, 1.57 and 1.33 below the monolingual mean respectively.

A study by Komeili and Marshall (2013) compared the performance of children between the age of 5;7 and 12;5 (18 TD English speaking monolinguals, 18 TD Farsi-English speaking bilinguals) on a RS task in English (32 sentences). The monolingual group scored higher than the DLL group, but when receptive vocabulary was taken into account by adding receptive raw scores into an ANOVA analysis model, the difference was not significant (group classification [monolingual or bilingual] did not contribute to variance in RS scores, but receptive language scores did).

Ziethe et al. (2013) conducted a study in German using RS and digit span tasks (tasks that test one's ability to hold a sequence of numbers reflecting short-term memory function). The study consisted of four groups of children aged 4;9 to 8;9 years from a wide range of linguistic backgrounds: 25 TD monolingual children, 19 monolingual children with PLI, 14 TD DLLs and 15 DLLs with PLI. Identification of PLI was based on -1 SD below the mean on at least two of three language tests in German (i.e. language comprehension, morphological rule building and vocabulary expression). Results showed that TD children, both monolingual and DLLs, scored significantly higher than the children with PLI (both monolingual and DLLs) on RS and digit span tasks in German. No differences were exhibited between the PLI groups on the two processing tasks. The authors concluded that since both tasks distinguished children with PLI from TD peers and that the results of both PLI groups were similar; these tasks could be potential markers of PLI among DLLs. The longitudinal Spanish study mentioned in the previous section provides additional support for this conclusion. Results showed that TD DLLs outperformed DLLs with PLI on a RS task persistently across time (at age 6, 8 and 12) (Buil-Legaz et al., 2016).

As part of the study by Armon-Lotem and Meir (2016) mentioned above, a unique RS task was created. The construction of the sentences was based on a detailed study (Meir et al., 2015) that analysed the crosslinguistic influence of Russian linguistic knowledge on performance on RS tasks in Hebrew. Results enabled creation of a RS task in Hebrew that excluded morphosyntactic features that could be affected by Russian morphsyntax. Hence, this task is more sensitive to errors typical of PLI and less affected by crosslinguistic influence, disentangling hallmarks of PLI from unique features of typical Russian-Hebrew dual language development. This study found that RS provided high sensitivity and specificity (100% and 89% respectively; LR+ =8.92 LR- =0.00).

In sum, current evidence shows the potential of using appropriate reference groups, parent questionnaires about the child's mother tongue and language processing tasks. The joint impact of these promising components has attracted additional attention. The following section will explore how a combination of diagnostic tools and approaches influences diagnostic accuracy.

3.3.9. Combination of Diagnostic Tools

In practice SLPs often use a combination of tools during assessments to gain sufficient information to guide clinical decision making (Owens, 2014; Paul, 2012). With specific reference to DLLs Kohnert advised that " a valid assessment can be achieved by using a combination of methods and triangulating data outcomes from multiple sources to evaluate developmental history, current levels of achievement in both languages as well as to assess the integrity of the child's more general ability to learn or use language" (2010, p. 465).

Previous research has supported the use of different combinations of tools that include, among others, parent questionnaires and language processing tasks. Conti-Ramsden and colleagues (Conti-Ramsden, 2003; Conti-Ramsden et al., 2001) explored the use of a range of tasks including NWR and RS among monolingual English speaking children. Results from both studies showed a slight increase in diagnostic accuracy when combining tasks. A combination of NWR and RS showed the most impressive values with 96% sensitivity, 78% specificity and overall 89% accuracy. However, this was only 1% higher than the RS overall accuracy (88%). It is important to note that mild language difficulties were identified at 50% level of accuracy or higher only when both tasks were combined, hence the value of combining both tasks is important on a clinical level. Similarly, Thordardottir and colleagues

(2011) found that a combination of a NWR task with a vocabulary comprehension task yielded better diagnostic accuracy than each task separately among monolingual French speaking 5-year-old children.

Paradis, Schneider and Sorenson Duncan (2013) were the first to explore the diagnostic potential of combing indirect and direct assessment tools to identify PLI among DLLs. They used a parent questionnaire to gain information on the child's mother tongue and standardized tasks in English. The idea of supplementing the results of the parent questionnaire (ALDeQ) followed the conclusions drawn by Paradis and colleagues (2010) that the ALDeQ could not be used as an independent diagnostic tool due to its low sensitivity (66%).

Two groups were compared: 152 TD DLLs and 26 DLLs with PLI. Participants had a mean age of 5;10 and were from a wide range of linguistic backgrounds who spoke English as their second language. PLI was identified based on current participation in speech-language therapy or special education programs for children with language difficulties. The study procedure included testing each child using four standardized tasks in English: NWR (CTOPP, 1999), Test of Early Grammatical Impairment (TEGI, 2001), Story Grammar (The Edmonton Narrative Norms Instrument [ENNI], 2005), and the Peabody Picture Vocabulary Test – III (PPVT, 1997) in addition to the ALDeQ. Results showed that DLLs with PLI had significantly lower scores on all tasks, except for vocabulary, compared to TD DLLs. The strongest assessment tool proved to be the ALDeQ, followed by NWR and the TEGI. A combination of all tools excluding the PPVT (i.e. ALDeQ, NWR, TEGI and ENNI) increased the overall diagnostic accuracy in comparison to the ALDeQ alone to 92% and yielded 92% specificity and 91% sensitivity (LR+ =11, LR- =0.10).

Additional studies have also explored the diagnostic accuracy of combinations of language processing tasks, shown to enhance monolingual identification of PLI, among DLLs. Thordardottir and Brandeker (2013) conducted a study using RS and NWR tasks in French. They compared 28 monolingual French speaking children with and without PLI to 28 DLLs with and without PLI with a mean age of 5. The TD DLLs spoke English as their mother tongue, while the DLLs with PLI spoke a wide range of mother tongues. The NWR task produced 85% sensitivity and 79% specificity (LR+ = 4.05; LR- =0.19). The RS task produced higher sensitivity, but lower specificity than the NWR task with 92% and 57% respectively (LR+ =2.14; LR- =0.14). However when both tasks were combined, sensitivity improved to 100% and specificity remained at the result obtained from the NWR task alone (79%) resulting in 89% overall accuracy (LR+ =4.75, LR- =0.00).

A further study conducted in French among 29 French-Arabic DLLs age 6-8 years old by Tuller and colleagues (2013) found that a combination of three tasks (NWR, RS and wh-question comprehension) produced promising results. The tasks chosen for this study emerged from COST Action IS0804. They used the French version of the quasi-universal NWR task and a unique RS task for French-Arabic DLLs. The RS task was created in a similar fashion to the RS task reported in Armon-Lotem and Meir’s study (2016). First Tuller and colleagues identified qualitative differences between the expressive language of TD DLLs and monolingual French children with PLI. Then they created a RS task that aimed to bypass cross linguistic influences that could underestimate TD DLLs. Lastly they tailored a wh-question comprehension task by allowing both verbal and non-verbal responses (pointing at a corresponding picture). A combination of all three tasks yielded 89.5% specificity and 90% sensitivity with overall 90% accuracy (LR+ =8.57, LR- =0.11).

The diagnostic accuracy of these three tasks was supported by a parent questionnaire regarding the child's mother tongue (i.e. Arabic). Tuller and colleagues used the Parents of Bilingual children Questionnaire (PABIQ: Tuller, 2015), which is based on the ALDeQ (J. Paradis et al., 2010) and ALEQ (J. Paradis, 2011). They found a significant difference between the scores of TD DLLs and DLLs with PLI on specific items they defined as risk factors: age of first word, age of first sentence and number of family members with language difficulties.

The assessment approach used by Tuller and colleagues was further supported by a parallel study conducted in Arabic among 42 French-Arabic DLLs in Lebanon, 10 with PLI (Henry, Tuller, Prevost and Zebib as cited in Tuller et al., 2013). The Arabic study used the same process described above to create the NWR and RS tasks in line with COST Action IS0804 conclusions. The combination of NWR and RS provided 100% sensitivity and 94% specificity with overall 95% accuracy (LR+ =16, LR- =0.00); and the parent questionnaire clearly distinguished between the two groups. In conclusion the findings from both Tuller et al. and Henry et al. show that carefully selected direct assessment in the child's second language supported by indirect information via parent questionnaire can provide important diagnostic information.

However, Armon-Lotem and Meir (2016) found that a combination of tasks did not improve the diagnostic accuracy of a single task. The combination of NWR and RS did not improve the accuracy compared with the RS task alone (91%; LR- =0.00, LR+ =8.92). Singularly the RS task produced 100% sensitivity and 89% specificity. In combination with the NWR task specificity improved slightly to 93%; however sensitivity declined to 80%, and no change was made to the overall accuracy (91%).

In sum, as can been seen in table 1, a combination of assessment tools has the potential to improve diagnostic accuracy. Nevertheless, when one assessment tool produces high sensitivity and specificity levels, combination with an additional tool may not add diagnostic value.

Table 1

Diagnostic Accuracy of Combinations of Assessment Tools

Publication	Combinations of assessment tools	Sensitivity %	Specificity %	Overall %	LR+	LR-
Paradis et al. 2013	ALDeQ + NWR + TEGI + ENNI	91	92	92	11	0.10
Thordardottir & Brandeker 2013	NWR + RS	100	79	89	4.75	0.00
Tuller et al. 2013	NWR + RS + Wh-question comprehension	90	89.5	90	8.57	0.11
Armon-Lotem & Meir 2016	NWR + RS	80	93	91	11	0.22

Note. Sensitivity=% correct PLI; Specificity=% correct TD; LR+ = positive likelihood ratio; LR- =negative likelihood ratio; ALDeQ= Alberta language and development Questionnaire; NWR= Nonword repetition task; TEGI =Test of Early Grammatical Impairment; ENNI = Story Grammar; RS= Recalling Sentences

3.4. The Current Study

Recent research has clearly outlined the challenges associated with DLL assessment and intervention within the Australian context and provided general guidelines to enhance service provision (Verdon, 2015; Verdon et al., 2014; Williams & McLeod, 2012). However to date no study has explored the diagnostic potential of a specific assessment approach in Australia.

The main aim of the current study was to investigate a unique assessment approach that would assist SLPs in Australia with the task of reliably identifying DLLs with PLI. One noticeable challenge is the diversity of languages other than English. In addition, most SLPs are monolingual English speakers. Therefore, the current study chose assessment tools in English alone with the practical prospect of exploring an approach that could be easily and readily adapted in clinical work. This study stems from the findings of a Canadian study that used indirect and direct assessment tools in English alone to discriminate between TD DLLs and DLLs with PLI (J. Paradis, Schneider, & Sorenson Duncan, 2013). The use of an assessment approach developed in Canada for an Australia study is supported by the many demographic and linguistic similarities both countries share (Washbrook et al., 2012).

The conscious decision to use tools in English posed two main drawbacks. First the gold standard of assessment of DLLs requires assessment of all the languages the child functions in to determine if impairment is evident in all languages (i.e. ASHA 2004; SPA 2009). Second, previous research has shown that TD DLLs are at risk of over identification when tested in their second language alone (Caesar & Kohler, 2007; Hoff et al., 2011; Marinis & Chondrogianni, 2010; Tuller et al., 2013; Wiig et al., 2006). In an attempt to minimize the drawbacks while still exploring an all English approach, the following decisions were made. It was initially decided that an indirect means would be used to collect developmental information about the child's mother tongue. Based on strong evidence (J. Paradis et al., 2010; Tuller, 2015; Tuller et al., 2013), including a pilot study in Perth Australia (May & Williams, 2012), the ALDeQ was chosen. Next, additional language tasks were chosen to directly assess the child's performance as recommended by Paradis and colleagues (2010). Specifically, tools with good ability to detect hallmarks of PLI were sought that were less impacted by language experience to maximise the chances of correct identification. Based on encouraging results using NWR and RS among DLLs around the globe (Armon-Lotem & Meir, 2016; Buil-Legaz et al., 2016; Komeili & Marshall, 2013; S. A. S. Lee & Gorman, 2013; Tuller et al., 2013; Vender et al., 2016; Wiig et al., 2006; Ziethe et al., 2013), the current study also chose these language processing tasks.

Maintaining a cautious perspective on the use of direct assessment tools in English for assessment of DLLs, the current study also investigated what internal and external factors contributed to performance on language processing tasks in an attempt to reject or verify their use within this population.

3.5. Research Questions

Specifically the present study asked the following questions:

1. In DLLs aged 5;0 to 6;11 (years; months) can the following assessment tools discriminate children with primary language impairments (PLI) from typically developing (TD) children, and which combination of tools is optimal?
    - Parent/caregiver questionnaire obtaining information about the child's first language.
    - Nonword repetition (NWR) task in English.
    - Recalling sentences (RS) task in English.
2. What proportion of DLLs reach monolingual norms on NWR and RS tasks in English? Can use of revised cut-off scores improve diagnostic accuracy on these tasks?
3. What is the relationship between child internal (age at time of testing, age of first exposure to English, length of exposure) and external factors (socioeconomic status,

parent's self- rated proficiency, parent's education, language use in the home and language richness) and performance on language processing tasks?

4. What is the relative contribution of each factor, and which combinations of factors predict performance on language processing tasks?

CHAPTER 4- METHODOLOGY

This case control study examined the diagnostic efficacy of three assessment materials by comparing the performance of typically developing (TD) Dual Language Learners (DLLs) and DLLs with Primary Language Impairment (PLI). In addition, the impact of internal/external factors on performance on language processing tasks was explored.

Data were collected over a period of seven months. Each assessment consisted of two sessions: One session with the child and one with their caregiver/parent (see Figure 1). All recruitment and data collection was completed by the researcher.

4.1. Ethics Statement

This study received ethical approval from Curtin University Human Research Ethics Committee (HR 225/2014) and the Western Australian Department of Education (D15/0015395) (see appendix A and B). Ethical guidelines followed specific requirements for children as well as individuals from a non-English speaking background. All recruitment and assessment data were coded and kept on a password protected computer. Hard copies were stored in a locked cabinet on the Curtin University campus.

The researcher maintained a current Working With Children Check throughout the course of the study as required by the Department of Child Protection.

4.2. Participants

Participants included in the study fitted the following criteria: functioned daily in two or more languages, aged between 5;0 and 6;11 (years; months), had no pre-existing cognitive sensory, psychological or neurological impairment (according to parent report) and hearing within the typical range on day of testing.

In total 61 participants took part, 42 TD DLLs and 19 DLLs with PLI. Seventy seven percent (*n* = 47) were exposed primarily to English after the age of 2 years and 34% (*n* = 21) were born outside of Australia. All participants had at least one parent who was born in a foreign country and for 80% (*n* = 49) of participants both parents were born in a non-English speaking country.

Table 2

Frequency of Mother Tongue

Mother Tongue	**Frequency**	
	TD DLLs	DLLs with PLI
French	15	1
Vietnamese*	3	3
Japanese	1	2
Arabic (Middle Eastern)	2	1
Italian	2	1
German	1	1
Spanish	1	1
Hebrew	2	
Malayalam	2	
Turkish	1	
Mandarin*	1	
Cantonese*	1	
Telugu	1	
Dutch	1	
Dzonka	1	
Bahasa Indonesia	1	
Farsi	5	
Chin*	4	
Hindi		2
Fillipino*		1
Kirundi		1
Arabic (African)		1
Tamil		1
Serbian		1
Shilluk		1
Masedonian		1
Kurdish		1

Note. *Non-tense-marking Mother Tongue. Frequency of languages exceeds total number of participants as three participants were trilingual

The sample included a total of 27 mother tongues (see Table 2). Three participants spoke more than two languages, two spoke Italian, French and English and one spoke Italian, Spanish and English. Overall the largest language group was French (*n* = 16) followed by Vietnamese (*n* = 6) and Farsi (*n* = 5). The distribution of languages between TD DLLs and

DLLs with PLI was not matched. Among TD DLLs the largest language group was French (*n* = 15), while Vietnamese was the largest language group among DLLs with PLI (*n* = 3).

The large range of languages is reflective of the large variety of over 400 languages spoken in Australia (Australian Bureau of Statistics, 2011), but not the proportion of each language reported in the population. The most commonly reported mother tongues for 4-5-year-old children in Australia, based on information collected from a population based study of 10,000 children were Arabic, Italian, Greek, Spanish and Vietnamese (Verdon et al., 2014). Excluding Greek, these languages are represented in the current study. While Arabic, Vietnamese and Italian appear in the top five languages in both the population based and present study, French and Japanese are more common in the present study than the population based study.

The TD DLLs had a mean age of 70.14 months (5;10 years; months) and the DLLs with PLI had a mean age of 68.58 months (5;9 years; months). The DLLs with PLI had on average longer exposure to English with over 3 years of exposure (39.32 months) compared to the TD DLLs who had 2;6 years of exposure (30.64 months). Independent-group *t* tests revealed no differences between groups for length of exposure to English, age, SEIFA (Socio-Economic Indexes For Areas: see section 5.1.2. for more details on this measure) and parent education (see Table 3). All participants in this study classify as "intellectually average/above average/superior" (Raven, Raven, & Court, 1998) scoring above the 25th percentile on Raven's Coloured Progressive Matrices.

In addition groups were matched based on language typology characteristics. Following criteria from previous studies (Blom & Paradis, 2013; J. Paradis, Genesee, et al., 2011) languages were split into two groups: tense-marking and non-tense-marking languages.

Twenty one percent of DLLs with PLI (n = 4) and 21.4% of TD DLLs (n = 9) had non-tense-marking mother tongues.

Table 3

Characteristics for TD DLLs and DLLs with PLI

Characteristics	TD	PLI	*p-value*
	n= 42 *M* (SD)	*n*= 19 *M* (SD)	
Age (months)	70.14 (7.67)	68.58 (7.28)	.457
Length of exposure (months)	30.64 (18.89)	39.32 (16.64)	.091
Nonverbal IQ	20.48 (4.31)	18.79 (3.36)	.179
SEIFA	1028.46 (71.24)	1011.23 (61.41)	.366
Mothers' education (years)	13.69 (5.68)	14.13 (2.24)	.731
Fathers' education (years)	14.00 (5.02)	13.16 (3.70)	.515

Note. TD= typically developing; PLI= primary language impairment; SEIFA= socioeconomic indexes for areas

Nonverbal IQ= according to Raven's CPM

4.3. Materials

4.3.1. Background Information.

Background information was collected from caregivers/parents using the Alberta Language Environment Questionnaire (ALEQ:J. Paradis, 2011) developed in Canada. Minor alterations were made to fit the questionnaire to the Australian population (e.g., referring to

Australia instead of Canada), without changing the meaning of the questions or the scoring scheme. The questionnaire is not language specific and was therefore administered to all the participants in the study (see full questionnaire in appendix C).

The questionnaire provided information on internal and external factors. Internal factors included: age of child at testing, child's length of exposure to English and age of first exposure to English. External factors included: socio-economic status, parent education, parent self-rated English proficiency, language use in the home and language richness.

4.3.2. Nonverbal Ability.

Nonverbal abilities were measured using Raven's Coloured Progressive Matrices (CPM). This test is used with children aged 5;0 to 11;0. It consists of 36 multiple choice questions increasing in difficulty. In each test item, the child is asked to identify, by pointing, the missing element that completes a pattern. One point is scored for each correct answer and total scores are compared to percentiles according to age. In various cultures the CPM has yielded adequate test-retest reliability and criterion and construct validity. The CPM was derived from the Standard Progressive Matrices in 1948. Since then, the test has been used widely around the world and was normed in Australia in 1975. The sample consisted of 700 children from the state of Queensland. Although the test has been normed more recently in the US, these norms have been proven unsatisfactory for the Australian population (Kamieniecki & Lynd-Stevenson, 2002) and therefore the Australian norms were used.

4.3.3. Hearing.

Hearing was tested using a GSI 38 automatic tympanometer and portable audiometer. Hearing tests and analysis of results followed the guidelines of the Community Health Manual (Government of Western Australia, 2013). All participants were required to respond to a pure tone at 25dB at 4 frequencies: 500, 1000, 2000 and 4000 Hz and produce a typical tympanogram (type A- normal, static compliance). Responses at amplitude greater than 25dB in one or more frequency and/or an atypical tympanogram resulted in exclusion of the participant.

4.3.4. Alberta Language and Development Questionnaire.

The Alberta Language and Development Questionnaire (ALDeQ: J. Paradis et al., 2010) was used to collect information regarding early and current language development in the child's first language. This questionnaire is a norm referenced questionnaire developed in Canada on a sample of 129 TD DLLs and 29 DLLs with PLI aged 5 to 7 years. The questionnaire consists of 18 questions divided into four sections: early milestones, current first language abilities, behaviour patterns and activity preferences, and family history (see full questionnaire in appendix D). The ALDeQ is not language specific and the norming sample included children from a range of language backgrounds. Answers are scored on rating scales that can be calculated to yield a total proportional score between 0 and 1. Scores below 0.66 are more consistent with PLI. The test yields excellent specificity correctly classifying 96% of typically developing DLLs, with moderate sensitivity detecting 66% of DLLs with PLI. In a small scale study conducted in Perth on 14 TD DLLs and 3 DLLs with apparent language difficulties, Australian total scores of TD DLLs were consistent with the Canadian norming population (May & Williams, 2012).

4.3.5. Nonword Repetition.

The NWR task chosen for this study is English-based and taken from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999). The CTOPP was normed on a sample of 1,656 individuals in the US aged 4 to 24 years. The composite scores demonstrate appropriate internal consistency ($\alpha > .8$), test-retest reliability ($r = .78$ to .95), inter-rater reliability ($r = .98$) and criterion and construct validity. NWR tasks in general, and the NWR sub-test in the CTOPP specifically, have proven to be reliable tests for distinguishing PLI in monolingual children (Hintze, Ryan, & Stoner, 2003) as well as DLLs (J. Paradis, Schneider, & Sorenson Duncan, 2013).

The task includes eighteen recorded nonwords (e.g., meb) that increase in syllable length and phonological complexity. Each response scores one point if the child produces an accurate repetition that includes all the sounds of the target word in the correct order. Standard scores have a mean of 10 and *SD* of 3.

4.3.6. Recalling Sentences.

The RS task chosen for this study is a sub-test of the Clinical Evaluation of Language Fundamentals Preschool-Second Edition-Australian Standardised Edition (CELF-P-2). CELF-P-2 is a test that was normed in Australia and New Zealand with a group of 342 English-speaking children aged 3;0 to 6;11 years old. The composite scores of the CELF-P-2 RS sub-test demonstrate excellent internal consistency ($\alpha > .90$) and test-retest reliability ($r = .90$) with appropriate criterion and construct validity. This sub-test yields a scaled score with a mean of 10 and a range of 7-13.

This RS task consists of a series of 13 sentences that increase in length (word number) and syntactic complexity. The content of the sentences relate to preschool and common children's activities in Australia using simple vocabulary. An accurate response required the child to repeat the sentence without changing the meaning while keeping to the original syntax and morphology. Scores for each sentence ranged from 0 to 3, relative to the number of errors made.

4.4. Procedure

4.4.1. Recruitment

Following ethics approval, principals of schools with a high percentage of DLLs, principals of Language Development Centres (LDCs) and bilingual families in the community were approached.

Thirteen schools in the Perth metropolitan area were contacted by phone and a detailed description of the study was sent by email. Five schools agreed to participate, however one school dropped out during the recruitment process. Two LDCs were contacted in the same manner and one agreed to participate.

Principals of schools and LDCs who agreed to participate in the study signed a written consent letter (see appendix E and F). The researcher consulted school staff to determine which students were DLLs and the level of their caregiver or parent's proficiency in English. Accordingly caregivers/parents were provided with information letters in English or a relevant language (see appendix G and H). Information letters were translated by professional translators to Arabic, Chin, Vietnamese and Farsi. When information letters were in a language other than English, an interpreter and the researcher were present in person at the time of letter distribution to allow parents to ask questions. Letters in English were distributed by the classroom teachers. Some caregivers/parents contacted the researcher by phone to receive additional information before returning signed consent forms. Signed consent forms were returned to schools and collected by the teachers then reviewed by the researcher before testing.

Participants were recruited from two primary schools with ESL programs, two Intensive English Centres and one Language Development Centre. Additional participants were recruited by word of mouth through bilingual parent groups in Perth. The coordinator of "Bilingual Families Perth" was contacted and the project was presented by the researcher at a workshop for parents of bilingual children. Information letters were distributed directly to caregivers/parents who expressed interest and signed consent forms were collected by the researcher. Consent forms were reviewed by the researcher before testing commenced.

Fifty one typically developing DLLs returned signed consent forms. However, nine TD DLLs were excluded from the study due to failing the hearing screening test (n = 6), uncooperative behaviour (n = 2) or electively dropping out (n = 1). Twenty two DLLs with PLI returned signed consent forms. Three were excluded due to failing the hearing test (n = 1), severe sound-speech impairment (n = 1) and cleft palate (n = 1).

DLLs with PLI (n = 19) were primarily recruited from a Language Development Centre (n = 14). Speech and Language Pathologists (SLPs) are part of the educational team at the centre. The children recruited from the LDC had been referred to the centre by SLPs in the community who believed they would benefit from intensive, daily intervention. These children underwent a comprehensive speech-language assessment in English to determine eligibility to attend the centre and have since been treated and monitored by SLPs who were consulted as part of the recruitment process. The additional DLLs with PLI who were recruited by word of mouth through bilingual parent groups were receiving SLP intervention due to language difficulties (n = 3) or were on a waiting list due to such difficulties (n = 2). Assessment protocols and tools were not available to the researcher. However, professional opinion about the presence of language impairment from an independent certified SLP was collected when possible (n = 17).

In summary participants were identified as DLLs with PLI based on:

- placement in speech-language therapy, a language development centre or on a waiting list for speech-language therapy due to language difficulties
- no pre-existing cognitive, sensory, psychological or neurological impairment according to parent report

The selection of TD DLLs was based on there being no reported history of language impairment/ language delay and no reported teacher/parent concern about language development.

All had typical hearing on the day of testing, were 5 and 6 year olds and functioned daily in more than one language including English.

4.4.2. Assessment process

The assessment process (see Figure 1) included two sessions for each participant: a session with the child where language processing tasks (Nonword Repetition [NWR] and Recalling Sentences [RS]) were administered in English and a meeting or phone call with the child's parents or caregivers to complete questionnaires (ALEQ and ALDeQ). See materials (Section 4.3) above for details of the method of assessments.

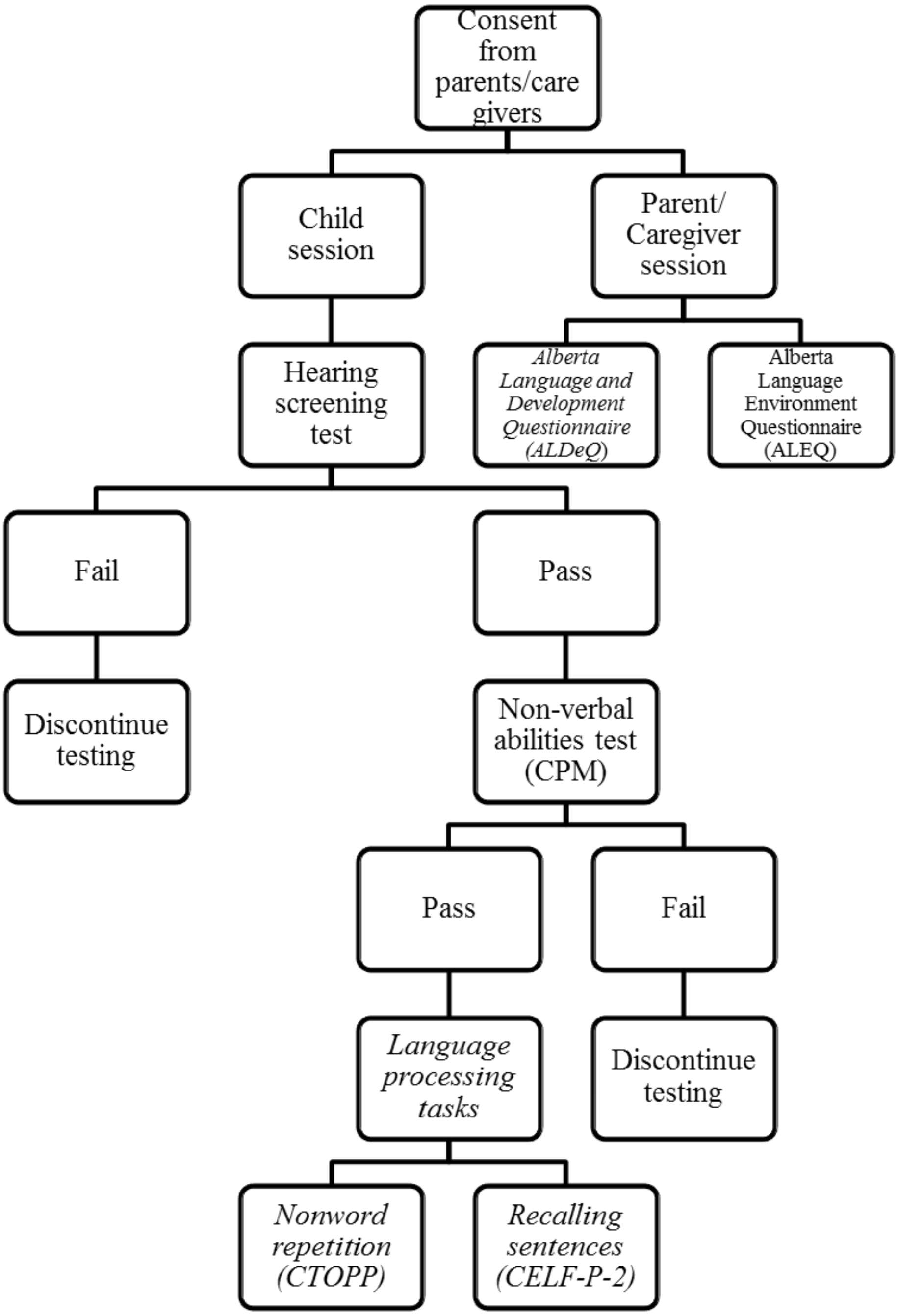

Figure 1. Assessment process

Note. *Italic*= assessment tools; non-italic= tools used for background information or eligibility (inclusion/exclusion criteria)

Data from the parents were collected by the researcher over a 25-45 minute session. Eighty four percent (n = 52) of interviews were conducted in English and 16% (n = 10) in other languages using professional translation services. The other languages were Chin, Vietnamese, Farsi and Middle Eastern Arabic. Ninety one percent (n = 56) of interviews were conducted face to face and five percent (n = 5) were conducted over the phone. Of the phone interviews, three required an interpreter.

Each child participant completed the assessments administered by the researcher over one session in a quiet room, either at home, at school or at Curtin University. All testing commenced with a screening hearing test followed by a non-verbal IQ subtest (Raven's Coloured Progressive Matrices; Raven, Raven & Court, 1998). Language processing tasks were then administered in accordance with procedures set down in the test manuals: NWR was presented by listening to an audio- recording using Laser Kids headphones and the RS task was administered orally by the researcher. Breaks were provided when needed and participants received a sticker at the end of the session.

The order of the language processing tasks was counterbalanced to control the effects of test order. The sequences were as follows: 1) screening hearing test, CPM, CTOPP and CELF-P-2; and 2) screening hearing test, CPM, CELF-P-2 and CTOPP. Each sequence was completed by 50% of the child participants.

All responses were recorded on a digital Zoom H1 Handy Recorder then copied to a laptop computer. The researcher transcribed and scored the responses in quiet following the guidelines in each respective test manual. Inter-rater reliability measures were taken and reported in section 5.3.

4.5 Rationale for Statistical Methods

Chapter 2 described the diversity of the multilingual population in Australia. For this reason, it is important to establish in detail the characteristics of the current sample. Standard descriptive statistics have been used to summarize and present the data.

The main objective of the present study was to explore an assessment approach for identification of PLI among DLLs in hopes of improving diagnostic accuracy. This is addressed in Research Question 1: in DLLs aged 5;0 to 6;11 (years; months) can the following assessment tools discriminate children with primary language impairments (PLI) from typically developing (TD) children, and which combination of tools is optimal?

- Assessment Tool 1 - Parent/caregiver questionnaire obtaining information about the child's first language
- Assessment Tool 2 - nonword repetition (NWR) task in English
- Assessment Tool 3 - recalling sentences (RS) task in English

A variety of analytical measures have been used to address this question: (a) sensitivity, specificity and overall accuracy, (b) receiver operating characteristic (ROC) curve analysis and (c) likelihood ratios.

Sensitivity and specificity rely on cut-off points to determine group classification. A score below the cut-off point is classified as impaired. Much discussion regarding the optimal cut-off point for identification of PLI among monolingual children has been published for different diagnostic tests, with cut-off points ranging between 2 standard deviations (*SD*) and 1 *SD* below the mean. Among publications which consider this issue, the following two are most relevant to the current study, since one uses the same language processingtasks while the other samples a similar population. Both publications consider scores more than 1 *SD* below the mean as an indication of PLI. The first study was conducted by Conti-Ramsden, Botting and Faragher (2001), who used the same type of language processing task (recalling sentences) as the current study. They found that the best sensitivity and specificity on NWR and RS tasks for monolingual children was produced when the cut-off point was 1 *SD* below the norm. The second study was carried out by Paradis, Schneider and Duncan (2013) in Canada. They researched a group of DLLs from linguistically diverse backgrounds using assessment tools in English, as did the present study. A cut-off point of 1 *SD* below the mean was also used. Following these leads, a cut-off score of 1 *SD* below the mean was used in the current study to compare scores to monolingual norms and thus address Research Question 2: What proportion of DLLs reach monolingual norms on NWR and RS tasks in English? Can the use of revised cut-off scores, derived using ROC curve analysis, improve diagnostic accuracy on these tasks?

Although the use of monolingual cut-off points was needed to allow comparison to monolingual norms, it has been suggested that monolingual cut-off points do not provide the most accurate discrimination between TD DLLs and DLLs with PLI (see Section 3.3). Therefore the receiver operating characteristic (ROC) curve analysis was used to test a variety of potential cut-off scores based on the study data (Portney & Watkins, 2009; Zhou, McClish, & Obuchowski, 2011). This analysis produces a graphic representation (a curve) of the distribution of scores on each assessment according to different cut-off scores. Sensitivity and specificity can be calculated along the graph to determine at which cut-off score the most participants are correctly identified. In other words, this analysis provides a set of potential combinations of sensitivity and specificity for each assessment tool, allowing the selection of the best cut-off point.

Beyond determining cut-off scores, the area under the curve reflects diagnostic accuracy. The ROC curve is plotted on a square with values of 1.0 for sensitivity (y axis) and 1-specificity (x axis) at the upper left and lower right corners, respectively. Therefore a perfect test would have 1.0 sensitivity and 0.0 specificity and the 'curve' would travel up the y axis and along the top of the graph. The area under the curve (AUC) would be the entire square (or 1.0). The larger the area under the curve, the more accurate the diagnostic material (Portney & Watkins, 2009) (see Figure 2 in section 5.2.3).

In the context of this study, the AUC is the probability that a randomly selected TD DLL will achieve a higher score than a randomly selected DLL with PLI. An area larger than 0.9 of the graph reflects high accuracy; an area between 0.7 and 0.9 reflects moderate accuracy; while between 0.5-0.7 reflects low accuracy (Fischer, Bachmann, & Jaeschke, 2003).

It is important to note that ROC curves provide a natural common scale for comparing different assessment tools. This is so even when they are measured using different units, thus allowing for comparison of assessment tools. The current study went beyond the calculation of the AUC and also calculated likelihood ratios. Likelihood ratios provide a value that indicates how strongly the score of an assessment tool can rule-in or rule-out the probability of an impairment. They exceed the dichotomous division of above or below the cut-off line by providing information on the probability of true positive, false positive, true negative and false negative results (see section 3.1.3). Therefore they permit the best use of results to establish diagnoses for an individual patient

The final part of the analysis focuses on the influence of internal and external factors on performance on language processing tasks (NWR and RS) to address Research Questions 3 and 4: (3) What is the relationship between child internal and external factors and performance on language processing tasks? (4) What is the relative contribution of each factor and which combinations of factors best predict performance on language processing tasks? Internal factors include age at time of testing, age of first exposure to English, and length of exposure to English. External factors include socioeconomic status (based on SEIFA), parent's self- rated English proficiency, parent's education (measured in years), language use in the home and language richness (in English and mother tongue).

Spearman's rho was deemed the most appropriate measure of the relationship between internal/external factors and scores on language processing tasks as it is a non-parametric measure of association and therefore does not require variables to be normally distributed. Internal relationships between the factors were also analysed to test for multicollinearity effects. Lastly, multiple regression analysis was used to identify the combinations of factors that explained optimal variance in the dependent variable.

4.6 Use of Raw and Scaled Scores

Both direct and indirect assessment tools were used in this study. The direct assessments (language processing tasks) were sub-tests of larger standardized assessments normed on a sample of predominantly monolingual children, making possible conversion to scaled scores. Scaled scores are based on converting raw scores according to age with a mean of 10 and a range of 7-13 within 1 SD of the mean. The use of scaled scores in this study was useful for two main reasons. First, it showed how DLLs performed in comparison to monolingual norms that are used in clinical settings (Williams & McLeod, 2012). Second, previous studies have shown age to have a significant effect on performance among DLLs (Goldberg, Paradis, & Crago, 2008; J. Paradis, 2011). The use of scaled scores helps overcome this effect. However, a major disadvantage is that scaled scores are based on monolingual norms and may not correctly reflect DLLs abilities. In this study, age was not a significant predictor of performance on language processing tasks (see Section 5.4.1). Therefore, both raw and scaled scores were analysed to address diagnostic accuracy queries (questions 1 and 2). As age was not found to impact language processing scores, only raw scores were used to explore the impact of internal/external factors on language processing tasks (questions 3 and 4).

CHAPTER 5 - RESULTS

5.1 Participant Characteristics

Multilingual communities in Australia differ greatly from one another. The next section will detail the main characteristics of the participants of the present study. This information (sections 5.1.1-.5.1.5) was collected using the Alberta Language Environment Questionnaire (ALEQ:J. Paradis, 2011) during parent interviews (see Section 4.3.1 for detailed description).

5.1.1 Age of First Exposure to English

Most of the participants (77%, n = 47) were consistently and sustainably first exposed to English when they entered day-care or kindergarten. Hence most of the participants were primarily exposed to their mother tongue with occasional exposure to English at the supermarket or watching television until they entered an English speaking educational setting. Therefore most participants in this study can be classified as sequential DLLs.

5.1.2 Socioeconomic Status

The sample closely reflects the average socioeconomic status in Western Australia. Socioeconomic status was calculated according to a measure produced by the Australian Bureau of Statistics (ABS); SEIFA (Socio-Economic Index For Areas). A SEIFA score is created using information about people and households in a particular area based on a five year census. Some examples of variables included are the proportion of people (1) unemployed, (2) employed and classified as labourers, (3) aged 15 years and over with no post-school qualifications, and (4) where households have no internet connection. For the purposes of the current study postal code areas were used to define geographic areas. The SEIFA score is standardized against a mean of 1000 with a standard deviation of 100. The average rank in Western Australia is 1,021 and the average rank in the sample was 1,020. Hence, the sample is highly representative of the West Australian population.

5.1.3 Parent Variables

Maternal education, measured in years, was relatively high in this sample with an average of 13.91 total years of education which is beyond high-school education. Parents' self-reported proficiency in English was on average functional. English proficiency was recorded using a 0-4 point scale: "no understanding/speaking ability" = 0, "some understanding and can say short, simple sentences" = 1, "good understanding and can express myself on many topics" = 2, "can understand and use English adequately for work and most other situations" = 3, and "can understand almost everything/very comfortable expressing myself in English in all situations" = 4. Among mothers the mean response was 2.75 (Range 0-4, *Median* 3) and among fathers the mean response was 2.95 (Range 1-4, *Median* 4). These means reflect similar proficiency among mothers and fathers, although mothers' medians and lower end of the range were poorer. In summary, parent education was relatively high and English proficiency was functional to high.

5.1.4 Language Use in the Home

Language use in the home varied between households of TD DLLs and DLLs with PLI. This factor was calculated by collecting detailed information about how the people who live in the home interact with the child on a daily basis. Parents reported the languages they speak to the child, in what languages the child replies, and how often each language is used. Scales were: "Mother tongue always/English never" = 0, "Mother tongue usually/English seldom" = 1, "Mother tongue 50%/English 50%" = 2, "Mother tongue seldom/English usually" = 3, "Mother tongue almost never/English almost always" = 4. Responses were scored and totalled according to the method used in the tool to reflect the number of individuals in the household and divided by four (the highest possible answer) to reveal a proportion of English use in the home. A higher proportion reflected a shift towards English use in the home while a lower proportion reflected relatively more use of the mother tongue at home at the time of testing.

Overall the mean proportion was .497 with a standard deviation of .312 (Range: 0-.1, *Median* = .458) reflecting equal use of English and mother tongue in the home. When analysing each group separately the mean is different between the groups. The mean proportion for TD DLLs was .381 (*SD* = .261, Range 0-.94, *Median* = .339) indicating that the mother tongue is used more often at home, while the mean proportion for the DLLs with PLI showed the opposite trend with most households shifting to English (*M* = .7543, *SD* = .258, Range .13-1, *Median* = .875). This difference is significant ($p < .001$). This result reflects that at the time of testing the DLLs with PLI were receiving on average more English input than the TD DLLs at home.

5.1.5 Language Richness

Group comparison reveals similar English richness among TD DLLs and DLLs with PLI, however mother tongue richness was lower among DLLs with PLI. Language richness was calculated by compiling literacy and language activities and the variety of speakers the child was exposed to on a weekly basis in each language. A higher score in one language suggested a richer level of exposure in that language. Overall the mean English richness was .654 (*SD* = .133, Range .31 –.88, *Median* = .687) while the mean mother tongue richness was .216 (*SD* = .168, Range 0 - .60, *Median* = .20) suggesting that overall the participants in this sample were exposed to more literacy and language activities in English. Further analysis reveals that the mean English richness among TD DLLs was lower (*M* = .632, *SD* = .139, Range .31-.88, *Median* = .625) than DLLs with PLI (*M* = .701, *SD* = .108, Range .50-.88, *Median* = .688), although this was not quite statistically significant (p = 0.0586). Conversely mother tongue richness was significantly (p < .001) higher among TD DLLs (*M* = .212, SD = .165, Range .0-.60, *Median* = .275) compared to DLLs with PLI (*M* = .092, SD = .092, Range .0-.25, *Median* = .05). This result indicates that although English richness was higher overall, TD DLLs were exposed to more language and literacy activities in their mother tongue than DLLs with PLI.

In summary, participants in the current study were representative of the mean SES status in Australia. Most parents had completed high-school and self-rated their English proficiency as functional to high. The households of TD DLLs tended to use more mother tongue than English, while households of DLLs with PLI were more likely to shift to English. Furthermore, mother tongue richness was higher among TD DLLs than DLLs with PLI.

5.2 Diagnostic Accuracy

The main objective of this study was to explore the diagnostic accuracy of three assessment tools (parent questionnaire, NWR task and RS task) independently and combined. The following sections present results describing the diagnostic capacity of each assessment tool independently and in combination with the other tools.

5.2.1 Group Comparisons

First, the mean scores for each assessment tool were used to compare TD DLLs and DLLs with PLI using independent samples *t*-tests (see Table 4). A large and significant difference between TD DLLs and DLLs with PLI was found for all three assessment materials.

Table 4

Between- Group Analysis for Assessment Tools

Assessment tool	TD *M (SD)*	PLI *M (SD)*	*t*	*p*	*d*
ALDeQ	0.8554 (0.10)	0.5179 (0.11)	*t*(59)=12.183	.000	3.307 *
NWR scaled	9.5952 (2.56)	5.6842 (2.00)	*t*(59)= 5.891	.000	1.703 *
NWR raw	7.95 (2.84)	3.21 (1.93)	*t*(59)= 5.20	.000	1.952 *
RS scaled	8.0952 (4.18)	5.0526 (2.15)	*t*(59)= 2.988	.004	0.915 *
RS raw	19.57 (9.69)	11.53 (6.58)	t(59)= 3.86	.002	0.971 *

Note. TD= typically developing; PLI= primary language impairment; ALDeQ= Alberta Language and Development Questionnaire Perth addition 2014; NWR= Nonword repetition task from the Comprehensive Test of Phonological Processing (CTOPP); RS= Recalling Sentences task from the Clinical Evaluation of Language Fundamentals Preschool-Second Edition- Australian Standardised Edition (CELF-P2); TD= typically developing; PLI= primary language impairment

*d > 0.8=large effect size

5.2.2 Comparison to Monolingual Norms and Revised Cut-Off Scores

Scaled scores were used to determine how the participant's scores in each group compared to monolingual norms. Scores below or equal to 7 were labelled as 'below monolingual norms', scores from 7 to 13 were labelled as 'within monolingual norms' and scores above or equal to 13 were labelled as 'above monolingual norms. As demonstrated in Table 5, a high percentage of DLLs with PLI fell below the monolingual norm on both language processing tasks. However, substantial proportion of TD DLLs also fell below the monolingual norm. A higher percentage of TD DLLs scored below the monolingual norm on the RS task (45%) than the NWR task (19%). As expected, only TD DLLs scored above monolingual norms on either task.

In terms of diagnostic accuracy, the monolingual cut-off scaled score (≤ 7) for NWR yielded 81% (34/42) specificity and 89% (17/19) sensitivity with overall good accuracy (84%). NWR likelihood ratios reflected suggestive accuracy in ruling in and out PLI (LR+ = 4.70, LR- = 0.13). The monolingual cut-off scaled score (≤ 7) for RS yielded 57.9% (24/42) specificity and 81% (15/19) sensitivity with overall poor accuracy (64%). Likelihood ratios indicate insufficient diagnostic power to rule in or rule out PLI with confidence (LR+ = 3.04, LR- = 0.52).

Table 5

Percentage of DLLs Below/Within/Above Monolingual Norms

Below/Within/Above Monolingual Norms	TD		PLI	
	RS	NWR	RS	NWR
	% (*n*)	% (*n*)	% (*n*)	% (*n*)
Below	45 (19)	19 (8)	84 (16)	89 (17)
Within	43 (18)	62 (26)	16 (3)	11 (2)
Above	12 (5)	19 (8)	0	0

Note. TD= typically developing; PLI= primary language impairment; NWR= Nonword repetition task; RS= Recalling Sentences task

Due to the high percentage of TD DLLs that fell below the monolingual norm on the RS task, and in an attempt to increase sensitivity and specificity values, receiver operating characteristic (ROC) curve (Zhou et al., 2011) analysis was used to explore revised cut-off scores. To increase the probability that DLLs with PLI score below the cut-off point and TD DLLs score above the cut-off point (i.e., optimal sensitivity and specificity), four cut-off scores were compared (see Table 6) to present a range of possible outcomes.

Out of the four cut-off scores, a revised scaled score of ≤ 4 (instead of ≤ 7) provided the best overall accuracy. Overall accuracy improved from 63.9% to 73.8%. However, the revised cut-off score showed a trade-off between sensitivity and specificity, improving specificity (from 54.8% to 81%), but reducing sensitivity (from 84.2% to 58%). Likelihood ratios reflect this trade off showing increased ability to rule in PLI (from LR+ = 1.86 to LR+ = 3.04), but reduced ability to rule out PLI (from LR- = 0.29 to LR- = 0.52). The remaining cut-off scores did not improve the diagnostic accuracy of the scaled scored substantially.

Interestingly the cut-off point with the best diagnostic accuracy for NWR was the monolingual cut-off point (≤ 7). Although 19% of TD DLLs scored below the monolingual normal range, a change in cut-off score did not improve diagnostic accuracy.

Table 6

Revised Scaled Cut-Off Scores for Language Processing Tasks

Language processing task	Revised scaled cut-off score	Sensitivity % (*n*)	Specificity % (*n*)	Overall accuracy % (*n*)	LR+	LR-
NWR	≤ 7	89.5 (17/19)	81 (34/42)	83.6 (51/61)	4.70	0.13
	≤ 6	68.4 (13/19)	88.1 (37/42)	81.9 (50/61)	5.75	0.36
	≤ 5	31.6 (6/19)	97.6 (41/42)	77 (47/61)	13.26	0.70
	≤ 4	15.8 (3/19)	100 (42/42)	73.8 (45/61)	Undefined	0.84
RS	≤ 7	84.2 (16/19)	54.8 (23/42)	63.9 (39/61)	1.86	0.29
	≤ 6	68.4 (13/19)	59.5 (25/42)	62.3 (38/61)	1.69	0.53
	≤ 5	57.9 (11/19)	69 (29/42)	65.6 (40/61)	1.87	0.61
	≤ 4	57.9 (11/19)	81 (34/42)	73.8 (45/61)	3.04	0.52

Note. Undefined=When the specificity is 100% the likelihood ratio is undefined

LR+ = Positive likelihood ratio; LR- = Negative likelihood ratio; NWR= Nonword repetition task; RS= Recalling Sentence task

5.2.3 Diagnostic Metrics for Assessment Tools

ROC curve analysis was further used to address diagnostic accuracy by determining which assessment tool best predicted group classification (TD DLLs as TD DLLs and DLLs with PLI as DLLs with PLI). ROC curves can be seen in Figure 2 and diagnostic accuracy, based on the best cut-off score for each assessment tool, is summarised in Table 7. The ALDeQ has the largest area under the curve (AUC =.991) indicating that it has the highest diagnostic accuracy, followed by NWR (AUC =.909 raw, 897 scaled) with moderate-high accuracy and RS with moderate accuracy (AUC =.740 raw,.724 scaled). Pairwise comparison of the ROC curves of the different assessment tools deemed them significantly different from one another (ALDeQ - NWR/Raw AUC = 0.0827, 95% CI [0.0103, 0.155], p = 0.025; ALDeQ – RS/Raw AUC = 0.268, 95% CI [0.125-0.378], $p < .001$; NWR/Raw – RS/Raw AUC = 0.169, 95% CI [0.0612, 0.277], p = 0.002). As expected, pairwise comparison confirmed no significant difference between raw and scaled scores of the NWR and RS tasks respectively (NWR/Raw – NWR/Scaled AUC = 0.0125, 95% CI [-0.0431, 0.0681], p = 0.659; RS/Raw – RS/Scaled AUC = 0.0157, 95% CI [-0.0267, 0.0581], p = 0.469).

Overall accuracy values and likelihood ratios support these results (Table 7). The ALDeQ produced the highest overall accuracy (96.7%) followed by the NWR task (82-83.6%). The RS task produced the lowest overall accuracy (73.8%). Likelihood ratios indicated that the ALDeQ ruled in and ruled out PLI with high diagnostic accuracy (LR+ 21.00= LR- =0.00). Raw and scaled scores of the language processing tasks produced similar likelihood ratios. Likelihood ratios indicated that the NWR task was more accurate than the RS task, but neither task produced high diagnostic power.

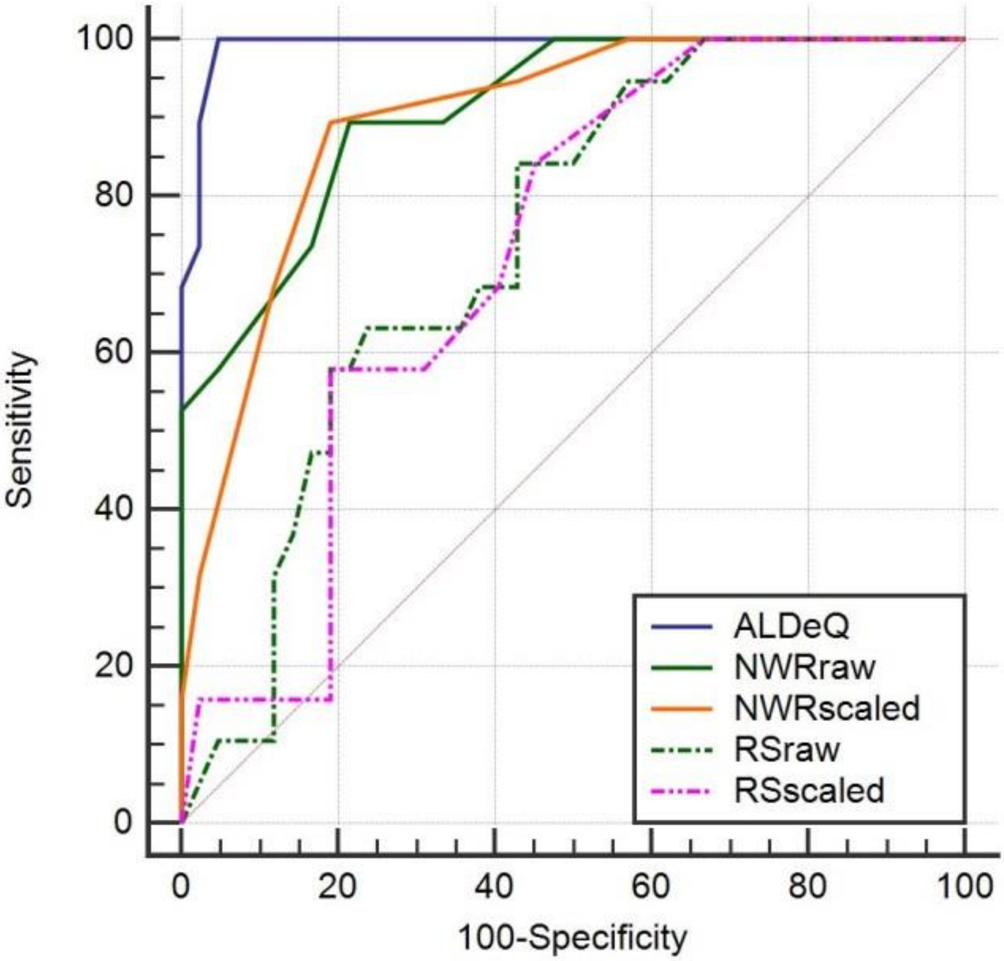

Figure 2. ROC curve of assessment tools for predicting group classification (TD DLL or DLL with PLI)

Note.TD = typically developing; DLL = dual language learner; PLI = primary language impairment; ALDeQ = Alberta Language Development; NWR = Nonword repetition task; RS = Recalling sentence: raw = raw scores; scaled = scaled scores

Table 7

Diagnostic Metrics for Assessment Tools

Assessment Tool	AUC	95% CI	Revised Cut-Off Scores	Sensitivity (%)	Specificity (%)	Overall (%)	LR+	LR-
ALDeQ	0.991	0.926-1.000	≤0.66	100.0 (19/19)	95.2 (40/42)	96.7 (59/61)	21.00	0.00
NWR-raw	0.909	0.808-0.968	≤5	89.5 (17/19)	78.6 (33/42)	82.0 (50/61)	4.18	0.13
NWR-scaled	0.897	0.792-0.960	≤7	89.5 (17/19)	81 (34/42)	83.6 (51/61)	4.70	0.13
RS-raw	0.740	0.612-0.844	≤18	57.9 (11/19)	81 (34/42)	73.8 (45/61)	3.04	0.52
RS-scaled	0.724	0.595-0.831	≤7	57.9 (11/19)	81 (34/42)	73.8 (45/61)	3.04	0.52

Note. LR+ = Positive likelihood ratio; LR- = Negative likelihood ratio; Sensitivity=% correct PLI; Specificity=% correct; ALDeQ= Alberta Language Development; NWR= Nonword repetition task; RS= Recalling sentences

5.2.4 Combinations of Assessment Tools

The revised cut-off scores were used to analyse the diagnostic accuracy of a combination of assessment tools as this has increased diagnostic accuracy in previous studies (Armon-Lotem & Meir, 2016; Conti-Ramsden et al., 2001; J. Paradis, Schneider, & Sorenson Duncan, 2013). All test combinations can be seen in Table 8.

Table 8

Diagnostic Accuracy on a Combination of Assessment Tools Using Revised Cut-Off Scores

	Combinations of assessment tools	Sensitivity (%)	Specificity (%)	Overall (%)
1	ALDeQ + NWR	89.5 (17/19)	100 (42/42)	96.7 (59/61)
2	ALDeQ + RS	57.9 (11/19)	100 (42/42)	86.9 (53/61)
3	NWR + RS	52.6 (10/19)	88.1 (37/42)	77 (47/61)
4	ALDeQ + NWR + RS	52.6 (10/19)	100 (42/42)	85.2 (52/61)
5	Scores below the cut-off line on any combination of assessment tools (ALDeQ + NWR, or ALDeQ + RS, or NWR + RS)	94.1 (18/19)	88.1 (37/42)	90.2 (55/61)

Note. Revised cut-off scores as shown in Table 7

Sensitivity=% correct PLI; Specificity=% correct TD; ALDeQ= Alberta Language and Development Questionnaire; NWR= Nonword repetition task; RS= Recalling Sentences task

The combination of the ALDeQ and NWR (Row 1) produced the highest overall accuracy (96.7%) with excellent specificity (100%) and good sensitivity (89.5%). However this combination did not improve the overall accuracy compared to the ALDeQ alone which also provided 96.7% overall correct identification of TD DLLs and DLLs with PLI. Scores below the cut-off line on any combination of assessment tools (Row 5) also provided excellent overall accuracy (90.2%) and the highest sensitivity value among the different combinations (94.1%). But this value was still lower than the sensitivity of the ALDeQ alone (100%). The lowest overall accuracy (77%) was produced when combining NWR and RS.

In summary, TD DLLs scored on average higher scores than DLLs with PLI on all three assessment tools (ALDeQ, NWR and RS). The ALDeQ provided the highest diagnostic accuracy followed by the NWR task. The RS task alone did not provide sufficient diagnostic information to correctly identify PLI among DLLs. Combinations of tools did not improve the high diagnostic accuracy produced by the ALDeQ alone, but two combinations reached the minimal sensitivity and specificity recommended by Plante and Vance (1994) for identification of PLI: ALDeQ + NWR and scoring below the cut-off point on any combination of tools.

5.3. Inter-rater Reliability

Two additional speech pathologists listened to recordings of a random selection of 10% of the language processing tasks to test inter-rater reliability. Intra-class correlation coefficient (ICC) reflected good to excellent reliability. The RS average measure ICC was .996 (95% CI [.974, .999]) and the NWR average measure ICC was .725 (95% CI [-.964, .962]). The results were transcribed by the author and used in the analysis.

5.4. Internal/external Factors and Performance

After exploring the potential diagnostic accuracy of an indirect assessment tool (parent questionnaire) in combination with direct assessment tools in English (language processing tasks), this section will further explore performance on language processing tasks. The initial interest was to see how internal and external factors impacted performance among both TD DLLs and DLLs with PLI. The hypothesis was that the scores of DLLs with PLI would be lower than TD DLLs regardless of internal and external factors since it would reflect an underlying deficit in the processing system. Hence it was thought that internal and external factors would only impact TD DLLs. Which specific factors and combinations of factors contribute to better performance on language processing tasks among TD DLLs was also of interest. It was hypothesised that both internal and external factors would impact scores on language processing tasks. Therefore correlations within internal and external factors were analysed.

5.4.1 Correlations within Internal and External factors.

Spearman’s rho correlations within internal and external factors were conducted to examine the nature and the strength of the relationship within each group of factors across the entire sample. Table 9 presents the correlations between the internal factors. Results indicated the presence of a strong negative correlation between age of first exposure to English (AoE) and length of exposure to English (LoE), $\rho(N = 61) = -.85$, $p < .01$, two –tailed. As expected, this confirms that participants with longer exposure to English were first exposed to English at a younger age. In addition, the correlation between AoE and age at the time of testing (AGE) was weak and positive, $\rho(N = 61) =-.33$, $p<.05$, two –tailed. This result indicates that in this sample older participants were first exposed to English at a later age.

Table 9

Correlations among Internal Factors

Internal factors	LoE	AoE
AoE	-.848**	
AGE	.147	.325*

Note. AoE=age of first exposure to English; LoE= length of exposure to English; AGE=age at time of testing.

***p*<.01, * *p*<.05.

Many correlations were found between external variables (Table 10). Mother and father years of education were moderately and positively correlated to socioeconomic status. Mother's self-rated proficiency was moderately to strongly positively correlated to mother and father years of education. As expected, mother tongue richness was moderately and negatively correlated to language use in the home. Hence, in homes where the mother tongue was the main language spoken by family members (represented by a relatively low 'language use in the home' score), children were exposed to more diverse activities in the mother tongue and a larger number of native mother tongue speakers with whom to communicate.

In addition a moderate and positive correlation was found between mother years of education and father years of education. Principal components analysis confirmed that both variables loaded strongly (.9+) on the same factor. The factor was subsequently named 'parental years of education' (pEDU). For each participant, scores on the factor were estimated by an unweighted average of mother and father years of education. The pEDU variable will be used in sections 5.4.1 and 5.4.2.

Table 10

Correlations among External Factors

External Factors	mSRP	fSRP	mEDU	fEDU	ER	MTR	LuH
SES	.431**	.385**	.571**	.567**	.277*	.138	-.013
mSRP		.222	.736**	.683**	.465**	.207	.218
fSRP			.164	.224	.103	.047	.169
mEDU				.682**	.289*	.201	.094
fEDU					.272*	.269*	.112
ER						-.306*	.487**
MTR							-.555**

Note. SES=socioeconomic status (measured by SEIFA [socioeconomic indexes for areas]); mSRP=mother's self-rated proficiency: fSRP=father's self-rated proficiency; mEDU=mother years of education; fEDU=father years of education; ER=English richness; MTR=mother tongue richness; LuH=language use in the home.

**p<.01, * p<.05.

5.4.2 Performance/Factors Correlations.

Ideally, a moderated regression model should be used to test the hypothesis that TD DLLs scores were more impacted by internal/external factors than DLLs with PLI. Due to the small sample size and large number of internal/external factors, however, the moderated regression model was insufficiently powered. Therefore, a more powerful alternative approach was used involving the comparison of non-parametric Spearman's rho correlations between the two groups (TD DLLS and DLLs with PLI) (see Tables 11 and 12).

It is clear that the TD DLLs had many more significant correlations between internal/external factors and performance on language processing tasks than the DLLs with PLI. In fact the DLLs with PLI only had one significant correlation. This could be explained by the number of participants in each group: 42 in the TD group and 19 in the PLI group. To control the discrepancy in group size, z-tests were used to compare the correlations in one group to the corresponding correlation in the other group. Ten correlations (23%) for the TD DLLs were significantly stronger than the corresponding correlation for the DLLs with PLI, suggesting that, in line with the hypothesis stated above, internal and external factors played a greater role in performance of TD DLLs compared to DLLs with PLI. Results suggest that the scores of DLLs with PLI reflect a deficit in their language processing system that is not related to internal or external factors. In other words variance across the full sample of TD DLLs and DLLs with PLI is primarily due to group classification. Further analysis was conducted on the TD DLLs alone to examine the contribution of internal and external factors to variance in performance.

Table 11

Correlations between Language Processing Tasks and Internal/External Factors among DLLs with PLI (n =19)

Internal/external factors		RS raw	NWR raw
internal	LoE	.313	.098
	AGE	.326	.049
	AoE	-.082	-.143
External	SES	.247	-.061
	mSRP	.048	.198
	fSRP	.385	.357
	pEDU	.005	.093
	ER	-.298	-.053
	MTR	.472*	.350
	LuH	-.121	.093

Note. AoE=age of first exposure to English; LoE= length of exposure to English; AGE=age at time of testing; SES=socioeconomic status (measured by SEIFA [socioeconomic indexes for areas]); mSRP=mother's self-rated proficiency; fSRP=father's self-rated proficiency; mEDU=mother's years of education; fEDU=father's years of education; ER=English richness, MTR=mother tongue richness; LuH=language use in the home .

* *p*<.05

Table 12

Correlations between Language Processing Tasks and Internal/External Factors among Typically Developing DLLs (n = 42)

Internal/external factors		RS raw	NWR raw
internal	LoE	.590***	.275
	AGE	.137	.022
	AoE	-.478**	-.222
External	SES	.473**	.414**
	mSRP	.722**	.481**
	fSRP	.283	.251
	pEDU	.738**	.579**
	ER	.510**	.251
	MTR	.118	-.013
	LuH	.481**	.351*

Note. AoE=age of first exposure to English, LoE= length of exposure to English, AGE=age at time of testing, SES=socioeconomic status (measured by SEIFA [socioeconomic indexes for areas]), mSRP=mother's self-rated proficiency, fSRP=father's self-rated proficiency, pEDU= mean parent years of education, ER=English richness, MTR=mother tongue richness, LuH=language use in the home , ***p<.001, **p<.01, * p<.05

5.4.3 Correlations between Language Processing Tasks and Internal/External Factors among Typically Developing DLLs.

As shown in Table 12, RS was significantly correlated to both internal and external factors for TD DLLs. Surprisingly, NWR was only significantly correlated to external variables. Better performance on both tasks was positively correlated to socioeconomic status, mother's self-rated proficiency, parents' education, and language use in the home. Strong correlations were only found between RS scores and mother's self-rated proficiency and parent years of education. The same variables were weakly to moderately positively correlated to NWR scores. Language use in the home was weakly and positively correlated to both tasks. English richness was moderately and positively correlated to performance on RS, but no correlation was found between NWR and English richness.

When considering internal variables, results showed a moderate and positive correlation between RS scores and length of exposure and a weak and negative correlation to age of exposure. No correlations to age were found.

5.4.4 Multiple Regression Analysis.

A multiple regression analysis using the stepwise entry procedure was conducted to explore which internal/external variables best predicted performance on language processing tasks. Stepwise-elimination regression analysis was used with the raw scores of the RS and NWR tasks as outcome variables. This analysis was chosen since some predictor variables (internal and external factors) were correlated with each other and therefore may account for the same portion of performance variance (see Table 13). If entered simultaneously into the regression analysis, multicollinearity effects may impact results. Only variables that were significantly correlated to language processing task scores were entered into the model.

After excluding redundant variables, 22.4% of the variability in NWR scores among TD DLLs could be attributed to parent education (Table 13), R^2 = .223, adjusted R^2 = .224, F (1, 40), p = .001. By Cohen's (1988) conventions, this effect is "medium".

With regard to RS scores, a combination of mother's self-rated proficiency and length of exposure to English accounted for 56.8-58.9% of variability (Table 13; R^2 = .589., adjusted R^2 = .568, F (2, 39), p = .001), which, according to Cohen's conventions, is rated as a "large" effect size. The multiple regression showed that length of exposure accounts for 15.8% of variance in RS scaled scores and mother's self-rated proficiency for 13.2%.

Table 13

Multiple Regression Analysis for Language Processing Tasks

Language processing tasks	Internal/external factors	*B* [95% CI]	St.error	β	sr^2
NWR	pEDU	0.277 [0.121, 0.433]	0.77	.493**	.243
RS	mSRP	12.686 [6.059, 19.313]	3.278	.461***	.158
	LoE	0.217 [0.093, 0.340]	0.061	.423*	.132

Note. *n*=42

CI=confidence interval, ***p<.001, ***p*<.01, * *p*<.05

pEDU= mean years of parent education , mSRP=mother's self-rated proficiency, LoE= length of exposure to English, RS=recalling sentences, NWR=nonword repetition

In summary, results from analysis of the TD DLL group show that the RS task was more impacted by internal and external factors than the NWR task. However neither task was neutral. Multiple regression analysis showed that the NWR task was most impacted by parent education alone, while the RS task was impacted by a combination of length of exposure to English (an internal factor) and mothers' self-rated English proficiency (an external factor).

CHAPTER 6- DISCUSSION

The main objective of this research was to further the inquiry initiated by Paradis and colleagues' (2013) Canadian study and explore the diagnostic potential of a combination of indirect (parent questionnaire) and direct (language processing) assessment tools in English to differentiate between TD DLLs and DLLs with PLI.

First the findings on diagnostic accuracy of all assessment tools and combinations of tools will be discussed, including a detailed comparison to the study conducted by Paradis and colleagues (2013). Next the discussion will focus on language processing tasks. This section will discuss the performance of DLLs in comparison to monolingual norms, the use of revised cut-off scores and the impact of internal/external factors on task scores. Finally theoretical implications will be considered.

6.1. Overall Diagnostic Accuracy

Overall, the results of this study corroborate the Canadian findings by Paradis and colleagues (2013). In both studies a combination of direct and indirect assessment tools in English differentiated between TD DLLs and DLLs with PLI. Participants in both studies were 5 and 6 year olds with and mean age of 70 months from diverse linguistic backgrounds. The number of DLLs with PLI in both studies was similar (19 in the current study and 17 in the Canadian study). However the total number of participants was much larger in the Canadian study (126 vs 61). In both studies children's mother tongue was divided into two groups based on typological characteristics (grammatical tense and verb inflection). The proportion of non-tense marking languages in the Canadian study was larger than the current study (44% vs. 21%) and the total number of mother tongues was larger in the current study (27 vs. 12). These results suggest that regardless of location and variation in participants' linguistic background, a combination of parent questionnaire and additional tasks in English may have the potential to correctly identify PLI among DLLs at the group level. Consistent with Paradis and colleagues, (2013) the results of the current study show that the parent questionnaire (ALDeQ) produced the highest diagnostic accuracy (AUC=0.991, LR+=20.83, LR-=0.05), followed by the NWR task (AUC= 0.897-0.909).

It is important to note the considerable differences between the diagnostic accuracy of the ALDeQ in the current study and the first study that used the ALDeQ for diagnostic purposes in Canada (J. Paradis et al., 2010). In the current study, all participants with PLI scored below the cut-off point on the ALDeQ (100% sensitivity) while sensitivity in the Canadian study was 66%. This may be due to the method of recruitmentMost of the DLLs with PLI (74%) in the current study were recruited from a Language Development Centre (LDC). As mentioned in chapter 4, students who attend the LDC are referred by SLPs who believe they will benefit from intensive daily intervention. In order to attain eligibility to service, all children must meet criteria defined by performance in English alone. Hence children who attend LDCs may lie at the more severe end of the PLI spectrum and consequently are more easily identified by the ALDeQ. In the Canadian study all children underwent speech-language assessments and were diagnosed as having a language impairment. Based on assessment outcomes, some were assigned to special programs while others underwent individual speech-language intervention by a speech pathologist. Additional details regarding the proportion of children in each program was not available. However, the possibly larger variation in assessment outcomes and form of intervention, suggests that the Canadian sample included a wider range of severity than the current study and therefore the ALDeQ may not have been as accurate at correctly identifying children with PLI. In contrast to the differences in sensitivity, specificity values in both studies were similar: 95.2% in the current study and 96% in the Canadian study.

Results of the current study showed that TD DLLs scored significantly higher ($p < .005$) than DLLs with PLI on all assessment tools: ALDeQ, NWR and RS. This result was obtained while at the time of testing the DLLs with PLI were receiving significantly more English exposure at home than the TD DLLs (0.754 vs 0.381 composite scores, $p < .001$), and had been exposed to English for a longer period of time (39.32 vs 30.64 months). Although the difference in length of exposure was not statistically significant (p=.091), the combined impact of both these factors could be seen as an advantage for the DLLs with PLI. Cumulative length of exposure and current exposure to the language of assessment have been recorded as significant factors positively impacting performance on different tasks, including language processing tasks (Chiat, 2015; Gibson A. et al., 2015; Kohnert et al., 2006; Sharp & V. C. M. Gathercole, 2013; Summers et al., 2010; Thordardottir & Brandeker, 2013). Despite the apparent advantage in exposure to English, the DLLs with PLI scored significantly lower than the TD DLLs. This difference adds to the body of research that supports the diagnostic potential of NWR and RS tasks in English for DLLs (Chiat et al., 2013; Paradis et al. 2013; Sorenson Duncan & Paradis, 2016; Windsor et al., 2010.

Similar results have been reported in cross-linguistic NWR studies that found that TD DLLs outperformed DLLs with PLI in both languages they spoke (Armon-Lotem & Meir, 2016; Windsor et al., 2010), as well as studies that used a NWR task in the child's second language alone (Boerma et al., 2015; Thordardottir & Brandeker, 2013; Tuller et al., 2013). In addition, previous research has shown group differences between DLLs with and without PLI on RS tasks in their second language (Thordardottir & Brandeker, 2013; Tuller et al., 2013; Ziethe et al., 2013). This result is particularly interesting in light of previous research that has shown that better knowledge of and experience with the test language contributes to repetition accuracy on both NWR tasks (Chiat, 2015; Ieclercq et al., 2013; Jones et al., 2010; Messer et al., 2010; Sharp & V. C. M. Gathercole, 2013) and RS tasks (Polišenská et al., 2015). The impact of factors associated with language knowledge and experience will be discussed in section 6.2.3, yet the results of the current study suggested that the influence of compounding factors does not interfere with the tasks' ability to differentiate between TD DLLs and DLLs with PLI at the group level.

In terms of diagnostic accuracy, the NWR task met the minimal criteria for diagnostic accuracy according to Plante and Vance (1994) of over 80% sensitivity and 80% specificity, when scaled scores were applied (sensitivity 89.5%, specificity 81%, AUC= 0.897, LR+=4.71, LR-=0.13). But diagnostic accuracy was marginally below the minimal requirement using raw scores (sensitivity 89.5%, specificity 78.6%, AUC= 0.909, LR+=4.18, LR-=0.13). Scaled scores are based on converting raw scores according to age. They were produced as part of the norming process of the entire assessment that included the NWR task used in this study (CTOPP: Wagner et al., 1999), hence they reflect predominantly monolingual performance in relation to age. The current study tested the correlation between NWR performance and age directly and found no significant correlation (p=.474). Therefore the use of scaled scores on the present sample is less appropriate despite the slightly higher diagnostic accuracy. Regardless of this discrepancy, likelihood ratios for both scaled and raw scores indicate only suggestive accuracy in ruling in and ruling out PLI. Previous studies have found the NWR to provide slightly better capacity to rule out PLI. NWR tasks in English among Spanish-English school aged DLLs (Windsor et al., 2010: LR+ =2.20 LR- =0.09) and in French among preschool DLLs from a range of linguistic backgrounds (Thordardottir & Brandeker, 2013: LR+ =2.20 LR- =0.09) produced high accuracy in ruling out PLI.

The RS task in the current study produced insufficient diagnostic accuracy (scaled scores: sensitivity 57.9%, specificity 73.8%, AUC= 0.740, LR+=3.05, LR-=0.52; raw scores: 57.9%, specificity 73.8%, AUC= 0.724, LR+=3.05, LR-=0.58). This result is similar to the findings of Thordardottir and Brandeker (2013) who also reported insufficient diagnostic power to rule in or rule out PLI (LR+ =2.14; LR- =0.14). Conversely, Armon-Lotem and Meir (2016) reported higher diagnostic accuracy than both the current study and that of Thordardottir and Brandeker (2013). They found that the RS task produced sufficient diagnostic power to rule in and rule out PLI with confidence (LR+ =8.92 LR- =0.00) among preschool Russian-Hebrew DLLs. The enhanced accuracy reported by Armon-Lotem and Meir could be a result of the specific RS task they used. Their RS task was designed specifically for Russian-Hebrew DLLs, while the present study used a RS task designed for monolingual English children (Wiig et al., 2006) and Thordardottir and Brandeker used a RS task adapted for monolingual French children (Thordardottir et al., 2011).

The current study also looked at the impact of different combinations of the assessment tools on diagnostic accuracy. No combination of tools produced higher diagnostic accuracy than the ALDeQ alone. Paradis and colleagues (2013) also investigated the diagnostic accuracy of combinations of assessment tools. They tested different combinations of a range of assessment tools in English, including the ALDeQ and the same NWR task used in the current study (see Table 14). Results showed that a combination of a parent questionnaire, the CTOPP NWR task, a grammatical task and a narrative task (ALDeQ + NWR + TEGI + ENNI) produced the highest diagnostic accuracy (92% overall accuracy). Other combinations provided lower specificity and sensitivity values and in turn lower overall accuracy. In all combinations reported by Paradis and colleagues (2013) the ALDeQ was the strongest predictor according to standardized coefficient values. However they did not include a comparison of the ALDeQ alone to different combinations. Therefore it is difficult to compare the results of the current study to those of Paradis and colleagues. However, previous research using different diagnostic tools has reported similar findings to those of the current study. Armon-Lotem and Meir (2016) found that no combination of tools improved the diagnostic accuracy of the RS task alone when using assessment tools in Hebrew (see Table 14). Hence, when one tool has very high sensitivity and specificity values there is little capacity for improvement when combined with additional tools.

Table 14

Diagnostic Accuracy of Assessment Tools and Combinations of Tools in Current Study and Previous Studies

Study	Combinations of assessment tools*	Overall accuracy (%)
Current study	ALDeQ	97
	ALDeQ + NWR (CTOPP)	97
	ALDeQ + NWR or ALDeQ + RS or NWR + RS**	90
	ALDeQ + RS	87
	ALDeQ + NWR (CTOPP) + RS	85
Paradis et al., 2013	ALDeQ + NWR (CTOPP) + TEGI + ENNI	92
	ALDeQ + TEGI +NWR (CTOPP) + ENNI + PPVT	91
	ALDeQ + NWR (CTOPP) + TEGI	84
Armon-Lotem & Meir, 2016	RS	91
	RS + NWR	91
	RS + FWD	90
	RS + NWR + FWD	90

Note. Overall accuracy = calculated according to sensitivity and specificity values

*Order reflects prediction strength/ individual diagnostic accuracy of assessment tool

** Scores below the cut-off line on any combination of assessment tools

ALDeQ= Alberta language and development Questionnaire; NWR= Nonword repetition task; CTOPP = Comprehensive Test of Phonological Processing; TEGI =Test of Early Grammatical Impairment; ENNI = Edmonton Narrative Norms Instrument; RS= Recalling Sentences, FWD = Forward digit span

In the current study a combination of the ALDeQ with the English NWR task produced the same overall accuracy as the ALDeQ alone (97%). In addition when calculating scores below the cut-off line on any combination of the three assessment tools (ALDeQ/ NWR/ RS) overall accuracy stayed relatively high at 90%. Other studies of DLLs using similar assessment tools have shown that a combination of tools upholds the minimal accuracy recommendations

(Armon-Lotem & Meir, 2016; J. Paradis, Schneider, & Sorenson Duncan, 2013; Thordardottir & Brandeker, 2013; Tuller et al., 2013).It is important to note that the use of more than one assessment tool is of clinical value (see section 7.1). Therefore the use of a combination of tools that upholds the minimal standards for diagnostic accuracy is of superior value than a single tool, even if that tool has high diagnostic accuracy alone For this reason results of the current study supports the use of a combination of assessment tools for identification of PLI among DLLs.

In summary, this study corroborates Paradis and colleagues study (2013) suggesting that a combination of indirect assessment of the child's mother tongue via parent questionnaire with direct assessment in English has the potential to provide invaluable information to disentangle PLI from multilingualism. Likelihood ratios indicate that only the ALDeQ can provide high diagnostic accuracy. A combination of tools does not improve diagnostic accuracy beyond that of the ALDeQ alone, but two combinations (ALDeQ + NWR; scores below the cut-off line on any combination of assessment tools) upheld minimal diagnostic criterion.

6.2. Performance on Language Processing Tasks

The proceeding sections will focus on the direct assessment tools used in the current study, namely the NWR and RS tasks (language processing tasks). The current data will be discussed in relation to performance of monolingual peers. In addition the use of revised cut-off scores will be critiqued as well as the impact of internal/external factors on performance.

6.2.1. Comparison to the Monolingual Normative Range

As expected most participants with PLI (84-89%) scored below the monolingual normative range or cut-off point, which is indicative of impairment, on language processing tasks (NWR and RS). A small proportion of DLLs with PLI scored within the normal range despite being a part of the PLI group. This may be due to initial misdiagnosis. Hence these participants may not have PLI despite meeting the study's criteria for PLI and scoring below the cut-off point on the ALDeQ. Alternatively, the language processing tasks did not reflect their classification as DLLs with PLI correctly. Rate of misdiagnosis on the RS was slightly higher than the NWR task with 16% (3/19) of DLLs with PLI scoring above the cut-off line compared to 11% (2/19) scoring above the cut-off line on the NWR task.

The characteristics of the misdiagnosed DLLs with PLI differed greatly in regards to length of exposure to English, age of first exposure to English, age and language use in the home. Therefore these factors most likely did not contribute to their enhanced performance. Interestingly, maternal self-rated proficiency was similar across these participants; all mothers rating themselves as having the highest level of English proficiency. The impact of maternal self-rated proficiency on the performance of TD DLLs will be discussed in section 6.2.3, however the positive impact of this external factor on performance is interesting to note at this point. It implies that although DLLs with PLI are expected to perform below the cut-off line due to impairment, some variance in scores may be attributed to compounding factors. These participants may have had relatively stronger language skills than other DLLs with PLI in the sample and therefore the impact of internal/external factors was more similar to that of TD DLLs. The very small figures of the current sample do not allow us to draw any conclusions, but this may be of interest in future research.

The results of the current study clearly show that TD DLLs perform differently than their English speaking monolingual peers on tasks in English. A considerable proportion scored below the monolingual mean (<1SD) on the language processing tasks (19% on NWR and 45% on RS). This result is consistent with previous studies that show a clear difference in performance between TD DLLs and monolingual peers. Paradis et al. (2013) showed that 24 - 78% of TD DLLs scored below the monolingual mean on a range of language tasks in English and Chiat and colleagues (2013) found that 9-88% of TD DLLs scored below the monolingual mean on standardized RS tasks in English, German and Hebrew. In addition, Sorenson Duncan and Paradis (2016) reported that 29% of TD DLLs scored below the monolingual normative range on a NWR task in English. It should be noted that the range and highest percentage of misdiagnosis differs greatly from study to study and within studies from task to task. This is most likely due to the methodical and sample differences between studies in regards to age of participants, assessment tools and additional internal/external factors. Nevertheless, a pattern of performance can be generalized from these findings. Despite the differences between the studies, a substantial proportion of TD DLLs are misdiagnosed when tasks are administered in English and monolingual norms used.

The differences between TD DLL performance and monolingual norms in the current study are predominantly apparent on the RS task. Previous research using RS tasks among DLLs have resulted in similar findings (Armon-Lotem & Meir, 2016; Komeili & Marshall, 2013; Wiig et al., 2006). Of particular interest are the findings of Wiig, Secord and Semel since the present study used the same RS task. Both studies show that as a group TD DLLs score below monolingual peers. However, Wiig, Secord and Semel found that TD DLLs scores remained within the monolingual norm, whereas in the current study 45% scored below the monolingual norm. Both samples were similar in size. The current sample consisted of 42 TD DLLs while Wiig, Secord and Semel included 47 TD DLLs in their sample of children from diverse cultural and linguistic backgrounds. Wiig, Secord and Semel did not provide information regarding language dominance or length of exposure to English. Therefore the discrepancy in RS results is difficult to explain, however it may be a reflection of the differences in inclusion criteria between the studies. The inclusion criteria in Wiig, Secord and Semel demanded that at least one parent was born overseas, while the current study required that a DLL function in two or more languages on a daily basis. The former criterion is broader and could have allowed for children who only rarely access their mother tongue to be included in the sample and therefore minimise the difference in RS scores from monolingual peers.

With respect to the NWR task, in the present study TD DLLs outperformed DLLs with PLI on the NWR task as a group; however a substantial percentage (19%) did not meet monolingual norms. This pattern of performance, with TD DLLs outperforming DLLs with PLI, but not performing as well as monolingual peers, reflects previous studies using different NWR tasks among DLLs (Armon-Lotem & Meir, 2016; Boerma et al., 2015; Kohnert et al., 2006; Sorenson Duncan & Paradis, 2016; Windsor et al., 2010).

To date, two Canadian studies (J. Paradis, Schneider, & Sorenson Duncan, 2013; Sorenson Duncan & Paradis, 2016) used the same NWR task in English as the current study (CTOPP: Wagner et al., 1999) with DLLs from a range of linguistic backgrounds. Canadian results showed that the proportion of TD DLLs that did not reach the monolingual normal range (29% and 42% respectively) was larger than the corresponding proportion in the present sample (19%). The difference in proportion of potentially misdiagnosed TD DLLs may be due to length of exposure to English. The TD DLLs in the present study were exposed to English for longer than the TD DLLs in the Canadian studies (on average 31months, compared to 16 and 21 months).

Another difference between all three studies is mother tongue typology, and specifically language phonology, that may have influenced the ability to reach monolingual norms. The large range of languages in both the current study and Paradis and colleague's study (2013) does not allow for direct comparison, however the results of Sorenson Duncan and Paradis (2016) demonstrate that language typology could be a viable factor that impacts results. This study divided the sample into two groups based on mother tongue typology: Chinese languages (Mandarin and Cantonese) and South Asian languages (Hindi, Punjabi and Urdu). These languages differ on a group level in terms of syllable structure. A brief generalization of these differences is that among Chinese languages the frequency of one syllable words is relatively high and the frequency of consonants in the coda position is relatively low. In contrast, in South Asian languages the syllable structure is broadly more similar to English consisting of multi-syllable words and frequently positioning consonants in the coda position. Sorenson Duncan and Paradis (2016) demonstrated that when taking language typology and length of exposure to English together, rate of misdiagnosis changes. They found that when they divided their sample and focused on DLLs from the Chinese group with less than 18 months exposure, 48% fell below the monolingual norm. In contrast when focusing on the South Asian group with more than 18 months exposure, 19% fell below the monolingual norm. These results suggest that when DLLs are exposed to English for longer and their mother tongue phonology is broadly more similar to English phonology, performance improves. Additional research, focusing on specific language groups is needed to further test this hypothesis.

Yet another difference between the present study and the Canadian studies is evident in size. The present study analysed a sub-group of 42 TD participants as compared with 109 TD DLLs in Paradis et al.'s study (2013) and 75 in Sorenson Duncan and Paradis's study (2016). This difference limits the relative statistical vigour of the current findings.

In sum, the findings of the current study support previous research that demonstrates that TD DLLs perform differently to monolingual peers on language processing tasks. The next section will discuss the potential of a multilingual specific scoring method in an attempt to minimize misidentification.

6.2.2. Revised Cut-Off Scores

Previous research has shown that TD DLLs with less accumulated language knowledge or who receive less input in the test language, score below monolingual means (Armon-Lotem et al., 2011; J. Paradis, 2010; Thordardottir, 2011). Logically, if the cut-off point of tasks were lowered, the proportion of TD DLLs below the cut-off line would decline. Explicit evidence supporting this hypothesis was demonstrated by Armon-Lotem and Meir (2016). They found that revised cut-off scores for NWR and RS tasks improved specificity levels from poor (48%, 63%) to acceptable levels (79%, 89%) while maintaining sensitivity above 80%.

The impact of revised cut-off scores on diagnostic accuracy in the current study was inconclusive, for two main reasons. First, ROC curve analysis showed that the best sensitivity and specificity values (89% and 81% respectively) for the NWR task were produced using the monolingual cut-off point. Hence alternative cut-off points did not improve diagnostic accuracy. This may be due the relatively high accuracy of the NWR task using monolingual cut-off scores, leaving little room for improvement in the application of revised scores.

Second, while a revised cut-off score did improve overall accuracy of the RS task (from 64% to 74%), there was a trade-off between specificity and sensitivity levels. Monolingual cut-off scores produced better sensitivity (84%) than specificity (55%); and revised cut-off scores produced better specificity (81%) than sensitivity (58%). These results reflect similar diagnostic accuracy with both cut-off scores, neither result reached the minimal accuracy recommended by Plante and Vance (1994) of above 80% sensitivity and 80% specificity. Therefore revised cut-off scores were of no diagnostic value for either measure in the current study.

Gillam, Peña and Perez (2013) reported a similar outcome exploring bilingual specific cut-off points for assessment of DLLs in English. Their study used vocabulary and grammar subtests of the Test of Language Development – Primary (TOLD-P; Newcomer & Hammill), and the Test of Narrative Language (TNL; Gillam & Pearson, 2004). First, monolingual cut-off scores were used and analysed using the EpiSLI model (Tomblin et al., 1996). According to this model two or more composite scores of 6.75 or lower are indicative of PLI. This monolingual criterion produced 95% sensitivity and 45% specificity. The 95% CI for the likelihood ratios (LR+[1.45, 2.06], LR- [0.02, 0.73]) reflected insufficient diagnostic power to rule in PLI. Use of revised criteria, with new composite cut-off scores for each subtest, improved specificity from 45% to 68%, but reduced sensitivity from 95% to 86% and did not improve the power to rule in PLI (95% CIs LR+ [1.97, 3.57], LR- [0.07, 0.60]). The sensitivity and specificity of these tests, using both monolingual and revised cut-off points, did not reach minimal accuracy, and the difference between cut-off values was insignificant (overlapping LR+ and LR- 95% CIs).

These findings are surprising since the strong correlation between accumulated quantity of language input and performance on language tasks implies that adjusting cut-off scores in accordance to language exposure (i.e. lower cut-off scores for DLLs) logically should improve diagnostic accuracy (Thordardottir, 2015a). In the current sample, the large variability in factors that influence performance on tasks, including age of first exposure and length of exposure to the task language, could have impacted the ability to make such predictions. The next section will discuss these factors in more detail.

6.2.3. Internal and External Factors

An additional aim of the current study was to explore the impact of internal and external factors on the language processing scores of TD DLLs. Internal factors include age at time of testing, age of first exposure to English and length of exposure to English. External factors include socioeconomic status (based on SEIFA), parent's self- rated proficiency in English, parent’s education (measured in years), language use in the home and language richness (in English and mother tongue).

Results show that both language processing tasks, NWR and RS, were positively correlated to the following external factors: socioeconomic status, parent education, mother’s self-rated English proficiency and language use in the home for TD DLLs. However NWR was not correlated to internal factors, while RS was negatively correlated to age of first exposure to English and positively correlated to length of exposure to English. Hence the RS task was more impacted by internal/external factors than the NWR task. This pattern was supported by the multiple regression analysis. Twenty-two percent of variability in NWR scores among TD DLLs could be attributed to an external factor: parent education. In comparison, 57% of variability on the RS task could be attributed to a combination of an external factor, maternal self-rated English proficiency, and an internal factor, length of exposure to English.

Similar results have been found among both monolingual children and DLLs. Ballandares, Marshall and Griffiths (2016) found that RS was impacted by socioeconomic status, while NWR was not, among 126 TD Spanish monolingual children aged 5;6 in Chile. Among DLLs, Thordardottir and Brandeker (2013) found that among 5-year-old simultaneous French-English DLLs from Canada, amount of language input was correlated to performance on RS tasks. Hence DLLs who received relatively more French input scored higher on the RS task in French and DLLs who received relatively more English input scored higher on the RS task in English. However, no correlation between amount of input and French NWR scores was found and English NWR scores were only weakly correlated to input in English.

Previous research exploring the impact of external factors on language processing tasks is scarce, however similarities to the current study can be found. Chiat & Polišenská (2016) found that a NWR task in English was marginally correlated to socioeconomic status, measured by geographical location of residence, among DLLs from diverse linguistic backgrounds. In the current study a weak correlation was found between socioeconomic status, measured in a similar fashion, and both NWR and RS scores. In addition maternal education was found to be correlated to performance on RS tasks among Russian-German DLLs using a RS task in German (the children's second language) (Chiat et al., 2013). In the current study parent education was also found to be correlated to performance on the RS task, however it was not selected for the stepwise regression analysis. This suggests that while a correlation was found, other factors, specifically mother's self-rated English proficiency and length of exposure, had a stronger correlation.

When addressed in a broad manner, these results are consistent with studies that showed that a range of external factors impact outcome measures among DLLs. Rojas and colleagues (2016) found correlations between maternal educational level and language outcomes in English among Spanish-English DLLs in kindergarten in the US. In addition, Golberg, Paradis and Crago (2008) reported a correlation between maternal education and lexical development in English among DLLs from various linguistic backgrounds.

Mother's self-rated proficiency in English has been correlated to DLLs success in tasks in English (Chondrogianni & Marinis, 2011; Scheffner Hammer et al., 2012). In the current study this factor was correlated to both tasks, but more strongly to the RS task. In addition Paradis

(2011) showed that English richness accounted for variance in language outcome measures in English. In the present study English richness was correlated to performance on the RS task. However this factor was not selected in the stepwise regression analysis, suggesting that it had less influence on performance in the current study than in Paradis' study. In the current study, language use in the home was calculated as the proportion of English spoken among family members at home. This factor was weakly correlated to both language processing tasks and was not selected for the stepwise regression analysis. Similarly, Paradis (2011) found that language use in the home was relatively less influential than other factors, such as English richness and mother tongue typology, on vocabulary and grammar tasks in English among DLLs aged 4;10 to 7;0 from a range of linguistic backgrounds.

The fact that both language processing tasks were correlated to external factors may explain the findings discussed in section 6.2.1. showing differences in DLLs' and monolinguals' performance on both language processing tasks. These findings support the hypothesis that language processing tasks are not language neutral, although NWR is more so than RS.

While both tasks were correlated to external factors, the RS task was also impacted by internal factors: length of exposure to English and age of first exposure to English. These factors have been argued to be the strongest predictors of language development among DLLs (Thordardottir, 2015a). It is important to mention that these two factors were, unsurprisingly, internally correlated to one another. Hence DLLs who were first exposed to English at a younger age were more likely to have relatively longer length of exposure to English.

Similar results were found for two Hebrew cohorts (English- Hebrew and Russian-Hebrew) reported in two different publications that included a RS task (Armon-Lotem et al., 2011; Chiat et al., 2013). Participants in these studies and the current study were similar in age (average of 70 months in both studies), length of exposure to the task language (average of 35 months in both studies) and sample size (samples ranging between 35-78 participants). Performance on a standardized RS task in Hebrew was correlated to age of first exposure with a trend towards length of exposure. In the current study length of exposure and age of first exposure were also correlated to performance, however of the two only length of exposure was selected for the stepwise regression analysis. Although both factors were correlated to RS in the two studies, the strength of correlation differed. This may be due to variance in external factors that could not be directly compared due to differences in methodology. Similarly, Thordardottir and Brandeker (2013) found that among 5-year-old simultaneous French-English DLLs from Canada those who received relatively more French input scored higher on the RS task in French and DLLs who received relatively more English input scored higher on the RS task in English.

With regard to the NWR task, previous research has produced mixed results. Summer and colleagues (2010) found that Spanish-English DLLs who had more cumulative experience in Spanish repeated longer nonwords with higher accuracy on a NWR task in Spanish than those with more cumulative experience in English. Similarly Sorenson Duncan and Paradis (2016) showed that longer cumulative English exposure positively impacted performance on a NWR task in English among DLLs with Chinese and South Asian mother tongues. Yet Thordardottir & Brandeker (2013) found that, similarly to the present study, cumulative French exposure was not correlated to performance on a NWR task in French and performance on a NWR task in English was only weakly correlated to English exposure.

The deviation of the current study from Sorenson Duncan and Paradis' (2016) study may be due to differences in participant's length of exposure to English, language typology and sample size. Participants in the current study were exposed to English for longer (31 vs 21 months), which may have brought them to a saturation point, in which length of exposure does not have an effect on NWR results. Another possible explanation may be language typology. The mother tongue of most participants in the present study was broadly more similar to English than Sorenson Duncan and Paradis' study (for more details see section 6.2). Therefore, positive transfer of linguistic knowledge from the child's mother tongue to English may have supported imitation of nonwords in English. In addition the current sample was smaller than Sorenson Duncan and Paradis's study (42 vs 75 TD DLLs), limiting the statistical significance of the current study. A comparison to Summer and colleagues study (2010) was not possible due to many methodological differences.

In sum, contrary to initial assumptions made in regard to language processing tasks in general and NWR in particular, the findings of the current study show that compounding factors impact performance on both NWR and RS tasks. Stronger correlations were found to the RS task, suggesting that more linguistic knowledge is required to correctly imitate a sentence than a nonword.

6.3. Theoretical implications

Overall, results of the current study support the use of psycholinguistic processing theories to analyse language and language impairment among DLLs. The results of the language processing tasks showed clear group differences between TD DLLs and DLLs with PLI. The difference in performance between the two groups reflects that language deficit is correlated to limited capacity and poorer function of underlying mechanisms. In other words, it is reasonable to conclude that the underlying reason DLLs with PLI did not imitate nonwords and sentences as accurately as TD DLLs in the present sample, is correlated to limited short term phonological memory capacity and poorer abilities to draw and integrate information from language sub-systems. Hence current results support theoretical accounts (M. Paradis, 2004) that the monolingual hypothesis of limited information processing capacity (Leonard, 2014) can be adapted when addressing DLLs with PLI.

Turning to the results of TD DLLs, the impact of external/internal factors on language processing task scores contribute to a growing body of knowledge that questions the original assumption that NWR tasks are linguistically and culturally neutral. According to the early perspective on NWR tasks, repetition of nonwords does not require knowledge of semantics and morphosytax (Dollaghan & Campbell, 1998; Law et al., 2011), as it was thought that hearing, retaining and repeating new auditory information does not require access to knowledge stored in the long term memory. Yet data from the current study and other recently published research (Engel et al., 2008; Gibson A. et al., 2015; Kohnert et al., 2006; S. A. S. Lee & Gorman, 2013; Sharp & V. C. M. Gathercole, 2013; Summers et al., 2010), that show differences between TD DLLs and monolingual peers, suggest that language experience contributes to performance and that repetition is supported by information stored in the long term memory (Rönnberg et al., 2013).

These results support the main ideas of social interactivism and usage based perspectives. These theories maintain that language acquisition is embedded in one's environment and impacted by individual linguistic experiences (Bloom and Tinker, 2001; Bybee, 2010). Hence, to acquire new linguistic knowledge a critical mass of input must be attained through meaningful communicative interactions with skilled partners (V. C. M. Gathercole, 2007). These theories predict that DLLs will lag behind monolingual peers as they have fewer opportunities to build up the amount of knowledge needed to reach a critical mass, and thus produce lower scores than monolingual peers on language tasks. The present data showed that a substantial proportion of TD DLLs fell below the monolingual normative range on both tasks and found correlations between performance and length of exposure to English. In addition scores were influenced by factors related to quality of input (i.e. parent education and mothers' self-rated English proficiency). Therefore these findings support the fundamental assumptions of social interactivism and usage bases theories and demonstrate how DLLs perform differently to monolingual children.

A comparison of the two language processing tasks to one another further strengthen the basic principles of the three theoretical perspectives that form the framework of this thesis**.** Results revealed that relatively more knowledge is required to successfully repeat a sentence than a nonword. While repetition of both tasks require activation of short and long term memory (Baddeley, 2000; Leonard, 2014; Polišenská et al., 2015), repetition of nonwords requires relatively less retrieval from the long term memory. To correctly repeat a nonword one needs to retain information in the short term memory while accessing phonological and semantic information in the long term memory (Ieclercq et al., 2013). However, to successfully repeat a sentence one must access information from a wider range of linguistic domains (morphology, syntax, semantics and pragmatics) (Polišenská et al., 2015). Hence more language experience is needed to attain a level of language knowledge across key domains in order to succeed on RS tasks. These results provide further support to the theoretical importance of input factors on language acquisition and support the psycholinguistic hypothesis that internal organization of items is determined by thresholds. As overall experience in each language is reduced, retrieval of items is more difficult for DLLs than monolinguals due to higher thresholds.

In sum, the findings of the current study support the complementary and integrated use of psycholinguistic, usage-based and social interactivism theories through which to view multilingual language development and assessment. In addition results support previous adaptations of these theories to address multilingual phenomena.

6.4. Summary of Discussion

The current project shows that a combination of indirect assessment of a DLL's mother tongue via parent questionnaire with direct assessment in English using language processing tasks has the potential to disentangle PLI from multilingualism at the group level. Results corroborate the findings of Paradis and colleagues (2013) who tested a similar approach in Canada on which this study was based. In addition, the current findings support previous research that TD DLLs perform differently to monolingual peers on language processing tasks and that performance on these tasks are impacted by internal/external factors. Contrary to expectations, use of novel cut-off scores did not improve the diagnostic accuracy of individual language processing tasks. Lastly, current findings support the integrated use of usage-based, psycholinguistic and social interactivism perspectives as an overarching framework through which to address multilingual language and language development matters.

The proceeding chapter will discuss conclusions and clinical implications of the current findings as well as limitations and scope for future research.

CHAPTER 7- CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

This study, the first of its kind in the Australian context, contributes important information to better understand how to disentangle PLI from multilingualism and thus to allow equal speech pathology service provision to all children, regardless of linguistic background. The main focus was to explore an all English approach, building on findings from Paradis el al. (2013), with the practical prospect of providing Australian SLPs with clinical tools.

7.1. Conclusions and Clinical Implications

Evidence from this project adds to the limited body of research specific to Australia. As detailed in the chapters above, the main findings of the study suggest that an assessment approach in English alone, that includes a combination of direct (NWR and RS in English) and indirect (parent questionnaire) assessment tools, has the potential to differentiate between TD DLLs and DLLs with PLI. The fact that the study produced similar outcomes to the Canadian study on which it was based (Paradis et al. 2013) further supports use of this approach. Hence, the current approach may assist clinicians to determine whether a child presents with an underlying language deficit due to impairment or a language difference due to multilingualism regardless of linguistic background or geographical location.

The results also contribute to the growing body of evidence demonstrating that language processing tasks (NWR and RS) are not culture and content neutral. In other words, although both tasks reflect a child's underlying language processing capacity, additional factors still impact their scores. Hence, a child's performance on these tasks cannot be seen as a 'clean' representation of their language processing abilities independent of internal and external variables.

Good individual diagnostic accuracy of the parent questionnaire (ALDeQ) and the NWR task support the use of these tools in clinical practice. Findings from the present study contribute to past findings that support the use of a parent questionnaire to gain information on the child's mother tongue when direct assessment is not possible (Paradis et al. 2013; Tuller et al., 2013; Tuller, 2015). In addition, similarly to previous research (Armon-Lotem & Meir, 2016; Buil-Legaz, Aguilar-Mediavilla & Adrover-Roig, 2016; S. A. S. Lee & Gorman, 2013; Thordardottier & Brandeker, 2013), results suggest that the NWR task may have the potential to minimize the disadvantage of minimal or reduced exposure to English in the assessment process. However the impact of compounding factors should always be considered when interpreting results.

In addition, results showed that the RS task did not provide sufficient diagnostic accuracy to identity PLI among DLLs alone and was more influenced by factors related to language experience than the NWR task. These findings suggest that children who have not accumulated knowledge in a particular language may lack sufficient knowledge to imitate sentences correctly, but may have adequate knowledge to complete NWR tasks. Hence the findings of the current study demonstrate the paradox of assessment of DLLs in their second language. As the RS task reflects performance across a range of language domains (semantics, syntax, morphology and pragmatics) it can provide relatively more diagnostic information than the NWR task, yet at the same time it is more impacted by language experience that may be lacking. Nevertheless the information gained from the RS task alone is by no means sufficient information to answer a diagnostic query or plan intervention.

Although the RS task did not reach minimal diagnostic criteria, a combination of any two assessment tools below the cut off line did (90% overall accuracy). Hence the RS task is of diagnostic value only in combination with both the ALDeQ and the NWR task. However, a combination of the ALDeQ and the NWR task without the RS task produced higher overall diagnostic accuracy (97%). This did not exceed the overall accuracy of the ALDeQ singularly (also 97%). Therefore one could conclude that the contribution of the language processing tasks to the ALDeQ is marginal.

Yet it is important to note that the need for a combination of assessment tools in clinical practice goes beyond sensitivity and specificity values. Kohnert (2010) recommends using a combination of direct and indirect measures to allow SLPs to integrate each child's individual circumstances with language outcomes when faced with a diagnostic question. The combination of tools in the current study provides indirect information on how the child functions in her mother tongue and direct information on language processing in English. A combination of the ALDeQ with at least one language processing task reflects performance in both languages, albeit the mother tongue is assessed indirectly through parent questionnaire. Thus the combination is of more clinical value than a single measure which provides information on one language. This is by no means satisfactory information for a comprehensive language assessment, neither does it provide sufficient data for identifying intervention goals, yet it does provide invaluable information to determine whether the child's language is different due to multilingualism or impaired due to an underlying disorder.

Therefore, based on the current data, SLPs are best advised to add the ALDeQ and the NWR task to their assessment battery when direct assessment of the DLL's mother tongue is not possible. These tools should be administered as part of a wider language evaluation that includes evaluation of all language domains and includes data collection from multiple sources to create a detailed profile that reflects the child's linguist abilities across different contexts.

7.2. Limitations

The current study addressed the complex issue of identification of PLI among DLLs in Australia. In line with the complexity of this matter, there are a number of expansions and alterations that would benefit the project, but were not possible within the constraints of the current study. These are considered below.

First, the classification of participants as typically developing or with primary language impairment was mainly based on the child's education setting or whether the child was receiving speech-language intervention. Although this technique is considered 'gold standard' when evaluating a new assessment approach (Leonard, 2014), the premise on which the need for this study was based was lack of appropriate assessment tools and high risk of misdiagnosis using common practises. Ideally all participants should have been tested in all the languages in which they function.

The recruitment approach did not allow for appraisal of severity of language impairment. No access to previous medical or educational records was sought by the researcher and consultation with current SLPs treating participants was minimal. However, most participants with PLI were recruited from language development centres (LDCs) that typically capture children at the more severe end of the PLI spectrum. Hence an additional limitation of the current study may be that it is more indicative for DLLs with pronounced language impairments and is less relevant to those with mild impairments.

A second limitation to the present study is the small number of participants. The findings should be seen as a contribution to the larger scale Canadian studies (J. Paradis et al., 2010; J. Paradis, Schneider, & Sorenson Duncan, 2013; Sorenson Duncan & Paradis, 2016) and as impetus for further multilingual research within Australia, particularly in the field of assessment. The modest sample size, while providing statistically significant results, also reduces the ability to draw strong conclusions in regard to the relationship between language processing tasks and language experience variables.

7.3. Scope for Future Research

Research in the field of multilingual language development and particularly assessment of language impairment has dramatically increased in the last few years. However, many questions still remain unanswered and little research has been conducted in Australia.

Future research elaborating on the current study would ideally plan for participation of a larger group of children. This would not only strengthen statistical power, but possibly provide assessment criteria for specific language sub-groups. For example, if a large group of Arabic-English DLLs were recruited, their data could be used to establish Australian specific Arabic-English criteria. This information would help clinicians determine what is expected from an Arabic-English DLL and thus minimize bias related to comparison to inappropriate reference groups.

Another direction for future research is to explore the impact of internal/external factors on additional assessment tools in English. Ideally a large group of participants would allow the research to systematically take into account specific factors, such as length of exposure to English and linguistic background. This could enable the establishment of specific recommendations that take into account individual child differences. For example, a different test battery may be recommended for DLLs from a Chinese background with less than 18 months exposure to English than DLLs from a European background with more than 30 months exposure to English.

In addition future research could use a similar design with a larger number of assessment tools (such as grammar and narrative assessments) to prove a more comprehensive reflection of linguistic abilities. This avenue could potentially lead to the creation of a full independent language evaluation tailored for DLLs in Australia.

To enhance generalization capacity, future research should examine DLLs with a wider range of language impairment severity (from mild to severe). To do so group identification would have to go beyond classification according to educational setting and include either assessment in the child's mother tongue or dynamic assessment.

Finally, future research could build on the potential found in the use of parent questionnaires and study a younger group of participants. This is of particular importance among DLLs as recent studies have shown that DLLs are identified later than monolingual children (Jordaan, 2008; Morgan et al., 2015) and therefore could miss timely intervention. The general framework of the ALDeQ could be used to build a new questionnaire, fit for this population.

In sum, results of the current study suggest that correct identification of PLI among DLLs using an all English approach within the complex multilingual context of Australia is possible. Novel data from the present study, the first of its kind in the Australian context, shows that the approach explored in this project has the potential to minimize the disadvantage of minimal or reduced exposure to English in the assessment process of DLLs while providing a practical and evidence based solution for SLPs. Use of the ALDeQ parent questionnaire and language processing tasks (NWR and RS) can assist in answering diagnostic questions as part of a wider test battery. SLPs are encouraged to use this approach when direct assessment of the DLL's mother tongue is not possible.

REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481-509. doi: 10.1017/S027226310808073X

Abutbul-Oz, H., Armon-Lotem, S., & Walters, J. (2012). *Bilingual Parents Questionnaire (BIPAQ).* Ramat-Gan: Bar Ilan University.

American Speech-Language-Hearing Association. (2004). Knowledge and skills needed by speech-language pathologists and audiologists to provide culturally and linguistically appropriate services. Retrieved from http://www.asha.org/docs/pdf/KS2004-00215.pdf.

Armon-Lotem, S. (2010). Instructive bilingualism: Can bilingual children with specific language impairment rely on one language in learning a second one? *Applied Linguistics, 31*, 251-343. doi: 10.1017/S0142716409990385

Armon-Lotem, S. (2012). Between L2 and SLI: Inflections and prepositions in the Hebrew of bilingual children with TLD and monolingual children with SLI. *Journal of Child Language, 41*(1), 3-33. doi:10.1017/S0305000912000487

Armon-Lotem, S., de Jong, J., & Meir, N. (2015). *Assessing multilingual children: Disentangling bilingualism from language impairment*. Bristol, UK: Multilingual Matters.

Armon-Lotem, S., Joffe, S., Abutbul-Oz, H., Altman, C., & Walters, J. (2015). Language exposure, ethnolinguistic identity and attitudes in the acquisition of Hebrew as a second language among bilingual preschool children from Russian- and English- speaking backgrounds. In T. Gruter & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 77-98). Amsterdam, The Netherlands: John Benjamins.

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

Armon-Lotem, S., & Ohana, O. (2016). A CDI study of bilingual English-Hebrew children – frequency of exposure as a major source of variation. *International Journal of Bilingual Education and Bilingualism*, 1-17. doi: 10.1080/13670050.2016.1179257

Armon-Lotem, S., Walters, J., & Gagarina, N. (2011). The impact of internal and external factors on linguistic performance in the home language and in L1 among Russian-Hebrew and Russian-German preschool children. *Linguistic Approaches to Bilingualism, 1*(3), 291-317. doi: 10.1075/lab.1.3.04armissn 1879–9264

Australian Bureau of Statistics. (2011). Languages spoken at home; Top 20 - Australia. Retrieved from www.abs.gov.au/census

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*(11), 417-423.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.

Balladares, J., Marshall, C. R., & Griffiths, Y. (2016). Socio-economic status affects sentence repetition, but not non-word repetition, in Chilean preschoolers. *First Language, 36*(3), 338-351. doi: 10.1177/0142723715626067

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*(1), 1-29.

Bedore, L. M., Peña, E. D., Garcia, M., & Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference? *Language, Speech & Hearing Services in Schools, 36*(3), 188-200.

Ben-Zeev, S. (1977). The infuence of bilingualism on cognitive strategy and cognitive development. *Child Development, 48*(3), 1009-1018.

Bishop, D. V. M., & Leonard, L. B. (2000). *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. East Sussex, UK: Psychology Press.

Bishop, D. V. M. (2003). *The Children's Communicaiton Checklist, Version 2 (CCC-2).* London, UK: Psychological Corporation.

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Communtion Disorders, 44*(5), 600-615. doi: 10.1080/13682820802259662

Blom, E., & Baayen, H. R. (2013). The impact of verb form, sentence position, home language, and second language proficiency on subject–verb agreement in child second language Dutch. *Applied Psycholinguistics, 34*, 777-811. doi: 10.1017/S0142716412000021

Blom, E., Chondrogianni, V., Marinis, T., & Vasic, N. (2016). The acquisition of verbal paradigms in Dutch and Greek L2 children: Cross-linguistic differences and inflectional defaults. *The International Journal of Bilingualism, 20*(4), 387-398.

Blom, E., & Paradis, J. (2013). Past tense production by English second language learners with and without language impairment. *Journal of Speech, Language & Hearing Research, 56*, 281-294.

Blom, E., Paradis, J., & Duncan, T. S. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular-s in child L2 English. *Language Learning, 62*(3), 965-994.

Blom, E., & Vasic, N. (2011). The production and processing of determiner-noun agreement in child L2 Dutch. *Linguistic Approaches to Bilingualism, 1*(3), 265-290.

Bloom, L., & Tinker, E. (2001). *The intentionality model and language acquisition: Engagement, effort, and the essential tension in development.* Oxford, United Kingdom: Wiley.

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

Bosch, L., Costa, A., & Sebastian-Galles, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12*(2), 189-221.

Bosch, L., & Sebastian-Galles, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy, 2*(1), 29-49.

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: Caseload characteristics. *International Journal of Language & Communication Disorders, 39*(3), 330-324.

Buil-Legaz, L., Aguilar-Mediavilla, E., & Adrover-Roig, D. (2016). Longitudinal trajectories of the representation and access to phonological information in bilingual children with specific language impairment. *International Journal of Speech-Language Pathology, 18*(5), 473-482. doi: 10.3109/17549507.2015.1126638

Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216-236). New York, NY: Routledge.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.

Caesar, L. G., & Kohler, P. D. (2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language, Speech & Hearing Services in Schools, 38*, 190-200.

Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, s., Pecini, C., . . . 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.

Chiat, S. (2015). Non-word repetition. In S. Armon-Lotem, J. de Jomg & N. Meir (Eds.), *Assessing multilingual children- Disentangling bilingualism from language impairment.* Croydon, UK: Multilingual Matters.

Chiat, S., Armon-Lotem, S., Marinis, T., Polisenska, K., Roy, P., & Seeff-Gabriel, B. (2013). The potential of sentence imitation tasks for assessment of language abilities in sequential bilingual children. In V. C. M Gathercole (Ed.), *Bilinguals and assessment: State of the art guide to Issues and solutions from around the world*. Bristol, UK: Multilingual Matters.

Chiat, S., & Polišenská, K. (2016). A framework for crosslinguistic nonword repetition tests: Effects of bilingualism and socioeconomic status on children's performance. *Journal of Speech, Language & Hearing Research, 59*, 1179-1189. doi: 10.1044/2016_JSLHR-L-15-0293

Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism, 1*(3), 318–345. doi: 10.1075/lab.1.3.05cho

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without language impairments (SLI). *International Journal of Language and Communication Disorders, 43*(1), 1-40.

Cohen, J. (1988). *Statisical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 52*(6), 407-428.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language & Hearing Research, 46*, 1029-1037. doi: 1092-4388/03/4605-1029

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741-748. doi: 10.1111/1469-7610.00770

Conti-Ramsden, G., Simkin, Z., & Pickles, A. (2006). Estimating familial loading in SLI: A comparison of direct assessment versus parental interview. *Journal of Speech, Language & Hearing Research, 49*, 88-101. doi: 1092-4388/06/4901-0088

Core, C., Hoff, E., Rumiche, R., & Senor, M. (2013). Total and conceptual vocabulary in Spanish-English bilinguals from 22 to 30 months: Implications for assessment. *Journal of Speech and Hearing Research, 56*, 1637-1647.

de Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Appled Linguistics, 13*(1), 1-24.

de Bree, E., Wijnen, F., & Gerrits, E. (2010). Non-word repetition and literacy in Dutch children at-risk of dyslexia and children with SLI: Results of the follow-up study. *Dyslexia, 16*(1), 36-44. doi: 10.1002/dys.395

De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol, UK: Multilingual Matters.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders, 46*(6), 613-627. doi: 10.1111/j.1460-6984.2011.00049.x

Department of Education. (2015). Language Development Centre 2016 referral guidelines. from: http://northeastldc.wa.edu.au/wp-content/uploads/2015/05/LDC-Referral-Guideline-Document-2016.pdf.

Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairmant: A study of diagnostic accuracy. *Journal of Speech, Language & Hearing Research, 56*, 323-336.

Dollaghan, C. A, & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language & Hearing Research, 41*(5), 1136-1146.

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Paul H. Brookes.

Dromi, E., Leonard, L. B., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language & Hearing Research*, *42*(6).

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test (3rd ed.).* Circle Pines, MN: Pearson Education.

Durkin, K., Conti-Ramsden, G., & Simkin, Z. (2011). Functional outcomes of adolescents with a history of specific language impairment (SLI) with and without autistic symptomatology. *Journal of Autism and Developmental Disorders, 42*, 123-138. doi: 10.1007/s10803-011-1224-y

Ellis Weismer, S., & Hesketh, L. J. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying seeking rates. *Journal of Speech and Hearing Research*, *39*, 177-190.

Engel de Abreu, P. M. J., Cruz-Santos, A., & Puglisi, M. L. (2014). Specific language impairment in language-minority children from low-income families. *International Journal of Language & Communication Disorders, 49*(6), 736-747.

Engel, P. M. J., Santos, F. H., & Gathercole, S. E. (2008). Are working memory measures free of socioeconomic influence? *Journal of Speech, Language & Hearing Research, 51*(6).

Fielding, R. (2015). *Multilingualism in the Australian suburbs: A framework for exploring bilingual identity*. Singapore: Springer.

Fischer, J. E., Bachmann, L. M., & Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: Clinical example of sepsis. *Intensive Care Medicine, 29*, 1043-1051. doi: 10.1007/s00134-003-1761-8

Gathercole, S. E. (2006). Nonword repetiotion and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morphosyntactic development in bilingual children. *The International Journal of Bilingual Education and Bilingualism, 10*(3), 224-247.

Gawlitzek-Maiwald, I., & Tracy, R. (1996). Bilingual bootstrapping. *Linguistics, 34*(5), 901-926.

Gibson A., T., Summers, C., Peña, E. D., Bedore M., L., Gillam, R., B., & Bohman, T. M. (2015). The role of phonological structure and experience in bilingual children's nonword repetition performance. *Bilingualism: Language and Cognition, 18*, 551-560. doi: :10.1017/S1366728914000248

Gillam, R. B., & Pearson, N. (2004). *Test of Narrative Language.* Austin, TX: Pro-Ed.

Gillam, R. B., Peña, E. D., & Perez, L. M. (2013). Identification of specific language impairment in bilingual children: l. Assessment in English. *Journal of Speech, Language, and Hearing Research, 56*, 1813-1823.

Girbau, D., & Schwartz, R. G. (2007). Non-word repetition in Spanish-speaking children with specific language impairment (SLI). *International Journal of Language and Communication Disorders, 42*(1), 59-75.

Goldberg, H., Paradis, J., & Crago, M. B. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Journal of Applied Psycholinguistics, 29*(1), 41-65. doi: 10.1017/S014271640808003X

Goverment of Western Australia. (2013). Hearing assessment: Tympanometry. Retrieved from http://www.pmh.health.wa.gov.au/general/CACH/docs/manual/6%20Procedures/6.3/6.3.3_Tympanometry.pdf

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3-15.

Gudmundsson, E., & Gretarsson, S. J. (2013). Mothers' questionnaire of preschoolers' language and motor skills: A validation study. *Child: Care, Health and Development, 39*(2), 246-252. doi: 10.1111/j.1365-2214.2011.01362.x

Guiberson, M., Rodríguez, B. L., Nippold, M., & Troia, G. (2013). Classification accuracy of nonword repetition when used with preschool-age Spanish-speaking children. *Language, Speech & Hearing Services in Schools, 44*(2), 121-132.

Gutierrez-Clellen, V. F., & Simon-Cereijido, G. (2009). Using language sampling in clinical assessments with bilingual children: Challenges and future directions. *Seminars in Speech and Language, 30*(4), 234-245. doi: 10.1055/s-0029-1241722

Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition, 4*(1), 85-99.

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language and Communication Disordorders, 48*(1), 1-24. doi: 10.1111/j.1460-6984.2012.00178.x

Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish-English bilingual children's language abilities. *Journal of Speech, Language & Hearing Research, 55*, 1251-1264.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes.

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy, 29*(1), 57-75.

Hintze, J. M., Ryan, A., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy skills and the Comprehensive Test of Phonological Processing. *School Psychology Review, 32*(4), 541-556.

Hoang, T. V., Schelstraete, M.-A., Tran, Q. D., & Bragard, A. (2014). Sentence repetition in Vietnamese : A marker of oral language and behavioral difficulties. *Canadian Journal of Speech-Language Pathology and Audiology, 37*(4), 280-298.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary developmnet via maternal speech. *Child Development*, *74*, 1368-1378.

Hoff, E., Core, C., Place, S., Rumiche, R., Senor, M., & Parra, M. (2011). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1-27.

Hopp, H. (2011). Internal and external factors in the child L2 development of German determiner phrase. *Linguistic Approaches to Bilingualism, 1*(3). doi: 10.1075/lab.1.3.02hopissn 1879–9264

Ieclercq, A.-L., Maillart, C., & Majerus, S. (2013). Nonword repetition problems in children with specific language impairment: A deficit in accessing long-term linguistic representations? *Topics in Language Disorders, 33*(3), 238-254. doi: 10.1097/TLD.0b013e31829dd8c9

Jackson, E., Leitao, S., & Classen, M. (2016). The relationship between phonological short-term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International Journal of Language & Communication Disorders, 51*(1), 61-73.

Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: Accuracy and error from two nonword repetition tests. *Journal of Speech, Language & Hearing Research, 53*, 1642-1655.

Jordaan, H. (2008). Clinical intervention for bilingual children: An international survey. *Folia Phoniatrica et Logopaedica, 60*, 97-105. doi: 10.1159/000114652

Kail, R., & Leonard, L. B. (1986). *Word-finding abilities in language-impaired children*, West Lafayette, IN.

Kalnak, N., Peyrard-Janvid, M., Forssberg, H., & Sahlen, B. (2014). Nonword repetition - A clinical marker for specific language impairment in Swedish associated with parents' language-related problems. *PLoS ONE, 9*(2), 1-10. doi: e89544. doi:10.1371/journal.pone.0089544

Kamieniecki, G. W., & Lynd-Stevenson, R. M. (2002). Is it appropriate to use United States norms to assess the “intelligence” of Australian children? *Australian Journal of Psychology, 54*(2), 67-78. doi: 10.1080/00049530210001706523

Kan, P. F., & Kohnert, K. (2005). Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. *Journal of Speech and Hearing Research, 48*, 372-378.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech & Hearing Services in Schools, 43*, 81-96.

Kimble, C. (2013). Speech-Language Pathologists' comfort levels in English language learner service delivery. *Communication Disorders Quarterly, 35*(1), 21-27. doi: 10.1177/1525740113487404

Klinger, J. k., & Artiles, A. J. (2003). When should bilingual students be in special education? *Educational Leadership*, *61*, 66-71.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*(6), 456–473. doi: 10.1016/j.jcomdis.2010.02.002

Kohnert, K. (2013). *Language disorders in bilingual children and adults* (2 ed.). San Diego, CA: Plural Publishing.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with primary language impairment from typical bilinguals? *Journal of Learning Disabilities Research and Practice, 21*, 19-29.

Komeili, M., & Marshall, C. R. (2013). Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. *Clinical Linguistics & Phonetics, 27*(2), 152-162. doi: 751625

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *American Journal of Speech-Language Pathology, 12*(1), 73-91.

Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech & Hearing Services in Schools, 34*(1), 44-55.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and languae delay: Findings from a systematic review of the literature. *International Journal of Language & Communication Disorders, 35*(2), 165-188.

Law, J., McBean, K., & Rush, R. (2011). Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage. *International Journal of Language & Communication Disorders, 46*(6), 657-664. doi: 10.1111/j.1460-6984.2011.00036.x

Lee, K., Chiu, S. N., van Hasselt, C. A., & Tong, M. (2009). The accuracy of parent and teacher reports in assessing the vocabulary knowledge of Chinese children with hearing impairment. *Language, Speech & Hearing Services in Schools, 40*(1), 35-45.

Lee, S. A. S., & Gorman, B. K. (2013). Nonword repetition performance and related factors in children representing four linguistic groups. *The International Journal of Bilingualism, 17*(4), 479-495.

Leonard, L. B. (2014). Specific language impairment across languages. *Child Development Perspectives, 8*(1), 1-5.

Leonard, L. B., Eyer, J., Bedore L, M., & Bernard, G. (1997). Three accounts of the grammatical morpheme difficulties of Englsih-speaking children with specific langugae impairment. *Journal of Speech, Language & Hearing Research, 40*(4), 741-753.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Massachusetts: MIT Press.

Liu, P. D., McBride-Chang, C., Wong, A. M.-Y., Tardif, T., Stokes, S. F., Fletcher, P., & Shu, H. (2010). Early oral language markers of poor reading performance in Hong Kong Chinese children. *Journal of Learning Disabilities, 43*(4), 322-331.

Lo Bianco, J. (2014). Domesticating the foreign: Globalization's effects on the place/s of languages. *The Modern Language Journal, 98*(1), 312-325. doi: 10.1111/j.1540-4781.2014.12063.x

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language, 37*(4), 817-840.

Marinis, T., & Chondrogianni, V. (2010). Production of tense marking in successive bilingual children: When do they converge with their monolingual peers? *International Journal of Speech-Language Pathology, 12*(1), 19-28. doi: 10.3109/17549500903434125

Matthews, S., & Yip, V. (2003). Relative clauses in early bilingual development: Transfer and universals. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 39-81). Berlin: Walter de Gruyter Gmbh & Co.

May, A., & Williams, C. J. (2012). Using parent report for assessment of the first language of English language learners. *Journal of Clinical Practice in Speech-Language Pathology, 14*(1), 25-30.

McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *Journal of Speech, Language & Hearing Research, 52*, 1213-1229.

Medwetsky, L. (2011). Spoken language processing model: Bridging auditory and language processing to guide assessment and intervention. *Language, Speech and Hearing Services in Schools, 42*, 286-296.

Meir, N., Walters, J., & Armon-Lotem, S. (2015). Disentangling SLI and bilingualism using sentence repetition tasks: The impact of L1 and L2 properties. *International Journal of Bilingualism*. doi: 10.1177/1367006915609240

Meisel, J. M. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development grammars. In J. Cenoz & F. Geness (Eds.), *Trends in Bilingaul Acquisition* (Vol. 1, pp.11-41). Philadelphia, PA: John Benjamins.

Meisel, J. M. (2011). *First and second language aquisition: Parallels and defferences*. Cambridge, United Kingdom: Cambridge University Press.

Mennen, I., & Stansfield, J. (2006). Speech and language therapy service delivery for bilingual children: A survey of three cites in Great Britain. *International Journal of Communtiona Disorders, 41*(6), 635-652.

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*, 306-323. doi: 10.1016/j.jecp.2009.12.006

Morgan, P. L., Farkas, G., Hillemeier, H. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special eduction: Longitudial evidence across five disability conditions. *Educational Researcher, 44*(5), 278-292. doi: 10.3102/0013189X15591157

Newcomer, P. L., & Hammill, D. D. (1997). *Test of Language Development – Primary* (3rd ed.). Austin, TX: Pro-Ed.

Oliveri, M. E., Ercikan, K., & Simon, M. (2015). Framework for developing comparable multilingual assessments for minority populations: Why context matters. *International Journal of Testing, 15*(2), 94-113. doi: 10.1080/15305058.2014.986271

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*, 191-230.

Owens, R. E. (2014). *Language disorders: A functional approach to assessment and intervention*. Upper Saddle River, NJ: Pearson.

Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning, 60*(3), 651-680. doi: 10.1111/j.1467-9922.2010.00567.x

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43*, 474-497.

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

Paradis, J., & Gruter, T. (2015). Inroduction to "Input and experience in bilingual development" In J. Paradis & T. Gruter (Eds.), *Input and experience in bilingual development* (pp. 1-15). Amsterdan: The Netherlands: John Bengamins.

Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of Child Language, 38*, 554-578.

Paradis, J., Nicoladis, E., & Genesee, F. (2000). Early emergence of structural contraints on code mixing: Evidence from French-English bilingual children. *Bilingualism: Language and Cognition, 3*(3), 245-261.

Paradis, J., Schneider, P., & Sorenson Duncan, T. (2013). Discriminating children with language impairment among English-language learners from diverse first-language backgrounds. *Journal of Speech and Hearing Research*, *56*(3), 971-981.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Vol. 18). Amsterdan: John Benjamins

Paul, R. (2012). *Language disorders from infancy through adolescence: Assesment and intervention*. St. Louis, MO: Elsevier Mosby.

Pearson, B. Z. (1998). Assessing lexical development in bilingual babies and toddlers. *International Journal of Bilingualism, 2*(3), 347-372. doi: 10.1177/136700699800200305

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*(1), 93-120. doi: 10.1111/j.1467-1770.1993.tb00174.x

Peña, E. D. (1993). *Dynamic assessment: A non-biased approach for assessing the language of young children*. Philadelphia, PA: Temple University.

Peña, E. D., Bedore, L. M., & Zlatic-Giunta, R. (2002). Category-generation performance of bilingual children: The influence of condition, category, and language. *Journal of Speech, Language & Hearing Research, 45*(5), 938-947.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology, 15*(3), 247-254.

Petersen, D. B., & Gillam, R. B. (2015). Predicting reading ability for bilingual latino children using dynamic assessment. *Journal of Learning Disabilities, 48*(1).

Petitto, L. A., Katerelos, M., Levi, B., Gauna, K., Tetrault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language, 28*(2), 1-44.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech & Hearing Services in Schools, 25*, 15-24. doi: 10.1044/1058-0360.0301.25

Polišenská, K. (2011). *The inlfuence of linguistic stucture on memory span: Repetition tasks as a measure of language ability.* (Doctor of Philosophy), City University London, Unpublished.

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language & Communication Disorders, 50*(1), 106-118.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3 ed.). Sydney, Australia: Pearson.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *38*(3), 265–282. doi: 10.1006/jmla.1997.2546

Poulin-Dubois, D., & Goodz, N. (2001). Language differentiation in bilingual infants: Evidence from babbling. In J. Cenoz & F. Geness (Eds.), *Trends in Bilingual Acquisition* (Vol. 1). Philadelphia: John Benfamins Publishing.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven's Coloured Progressive Matrices (CPM)* (1998 ed.). Oxford, UK: Oxford Psychologists Press.

Restrepo, M. A., & Gutierrez-Clellen, V. (2001). Article use in Spanish-speaking children with specific language impairment. *Journal of Child Language, 28*, 433-452.

Rice, M. L. (2004). Growth models of developmental language disorders. In M. L. Rice & S. F. Waren (Eds.), *Developmental language disorders: From phenotypes to etiologies* (pp. 207-240). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L. (2013). Language growth and genetics of specific language impairment. *American Journal of Speech-Language Pathology, 15*(3), 223-233. doi: 783113

Rice, M. L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2;6 to 21 years of age. *Journal of Speech, Language & Hearing Research, 58*, 345-359. doi: 10.1044/2015_JSLHR-L-14-0150

Rice, M. L., & Wexler, k. (2001). *Test of Early Grammatical Impairment.* New York, NY: Psychological Corporation.

Roesch, A. D., & Chondrogianni, V. (2016). *"Which mouse kissed the frog?"* Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of *wh*-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language, 43*(3), 635-661. doi: 10.1017/S0305000916000015

Rojas, R., Iglesias, A., Bunta, F., Goldstein, B., Goldenberg, C., & Reese, L. (2016). Interlocutor differential effects on the expressive language skills of Spanish-speaking English learners. *International Journal of Speech-Language Pathology, 18*(2), 1-12. doi: 10.3109/17549507.2015.1081290

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7,* 1-17. doi: 10.3389/fnsys.2013.00031

Roseberry-McKibbin, C. (1995). *Multilingual students with special needs*. Oceanside, CA: Academic Communication Associates.

Roseberry-McKibbin, C., Brice, A., & O'Hanlon, L. (2005). Serving English language learners in public school settings: A national survey. *Language, Speech & Hearing Services in Schools, 36*(1), 48-61.

Roy, P., & Chiat, S. (2013). Teasing apart disadvantage from disorder: The case of poor language. In C. R. Marshall (Ed.), *Current issues in developmental disorders* (pp. 125-150): Hove, UK: Psychology Press.

Runnqvist, E., Fitzpatrick, I., Strijkers, K., & Costa, A. (2012). An appraisal of the bilingual language production system: Quantitaively or qualitatiely different from monolinguals?

In T. K. Bhatia & R. C. William (Eds.), *The handbook of bilingualism and multilingualism* (2 ed, pp. 244-265). West Sussex, UK: Blackwell.

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education- Rates and predictors of identifications for services. *Journal of Learning Disabilities, 42*(2). doi: 10.1177/0022219408326221

Sánchez, L. (2006). Bilingualism/second-language research and the assessment of oral proficiency in minority bilingual children. *Language Assessment Quarterly, 3*(2), 117-149. doi: 10.1207/s15434311laq0302_3

Scheele, A., Leseman, p., & Mayo, A. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*, 117-140. doi: doi:10.1017/S0142716409990191

Scheffner Hammer, C., Komaroff, E., Rodríguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish-English bilingual children's language abilities. *Journal of Speech, Language & Hearing Research, 55*. doi: 10.1044/1092-4388(2012/11-0016)

Schneider, P., Dube, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instument (ENNI).* Retrieved from www.rehabmed/ualberta.ca/spa/enni

Schulz, P., & Friedmann, N. (2011). Specific language impairment (SLI) across languages: Properties and possible loci. *Lingua, 12*(3), 333-338.

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *WIREs Cognitive Science*, *7*(4), 264-275. doi: 10.1002/wcs.1393

Sebastian-Galles, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phenemic contrasts in early bilinguals. *Cognition, 72*(2), 111-123.

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language & Communication Disorders, 45*(6), 691-702.

Seme, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Development (CELF)- 4.* San Antonio, TX: The Psychological Corporation.

Semel, E., Wiig, E. H., & Secord, W. A. (2006). *Clinical Evaluation of Language Development (CELF)- 4* (Australian ed.). Sydney, Australia: Pearson Clinical and Talent Assessment.

Serratrice, L. (2012). The bilingual child. In R. C. William & T. K. Bhatia (Eds.), *The handbook of bilingualism and multilingualism* (2 ed, pp. 87-108). Chichester, West Sussex, UK.

Shaalan, S. (2010). *Investigating grammatical complexity in Gulf Arabic speaking children with specific language impairment (SLI).* (Docotor of Philosophy), University College London, London, UK.

Sharp, K. M., & Gathercole, V. C. M. (2013). Can a novel word repetition task be a language-neutral assessment tool? Evidence from Welsh–English bilingual children. *Child Language Teaching and Therapy, 29*(1), 77-89.

Sorenson Duncan, T., & Paradis, J. (2016). English language learners' nonword repetition performance: The influence of age, L2 vocabulary size, length of L2 exposure, and L1 phonology. *Journal of Speech, Language & Hearing Research, 59*, 39-48. doi: 10.1044/2015_JSLHR-L-14-0020

Speech Pathology Australia (SPA). (2009). Working in a culturally and linguistically diverse society [position paper] Retrieved fromhttp://www.speechpathologyaustralia.org.au /library/ Clinical_Guidelines/Working_in_a_CALD_Society.pdf

Speech Pathology Australia (SPA). (2011). Competency-Based Occupational Standards for Speech Pathologists. Retrieved from https://www.speechpathologyaustralia.org.au

/SPAweb/Resources_For_Speech_Pathologists/CBOS/SPAweb/Resources_for_Speech_Pathologists/CBOS/CBOS.aspx?hkey=c1509605-c754-4aa8-bc10-b099c1211d4d

Speech Pathology Australia (SPA). (2015). A snapshot of Australia's speechies. *2015 annual report*. Melbourne: Australia.

Squires, J. (1996). Parent-completed developmental questionnaires: A low-cost strategy for child-find and screening. *Infants and young children, 9*(1), 16-28.

Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholiguistic framework*. London: UK: Whurr.

Statistics Canada. (2011). Census of population. from http://www12.statcan.gc.ca/

Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49*(2), 219-236.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language & Hearing Research, 44*(6), 1321-1337.

Stow, C., & Dodd, B. (2003). Providing an equitable service to bilingual children in the UK: a review. *International Journal of Language & Communication Disorders, 38*(4), 351-377. doi: 10.1080/1368282031000156888

Summers, C., Bohman, T., Gillam, R., Peña, E. D., & Bedore, L. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders, 45* (4), 480-493.

Thal, D. J., O'Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language & Hearing Research, 42*, 482–496.

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism, 14*(4), 426-445.

Thordardottir, E. (2015a). Proposed diagnostic procedures for use in bilingual and cross-linguistic contexts. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 331-358). Croydon, UK: Multilingual Matters.

Thordardottir, E. (2015b). The relationship between bilingual exposure and morphosytactic development. *International Journal of Speech-Language Pathology, 17*(2), 97-114.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46,* 1-16.

Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., . . . Chilingaryan, G. (2011). Sensitivity and specificity of French language and processing measures for the identification of primary language impairment at age 5. *Journal of Speech, Language & Hearing Research, 54*(2), 580-597.

Tomblin, J. B., Records, N., & Zhang, X. (1996). A system for the diagnosis of specific langauge impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*(6), 1284-1294.

Tuller, L. (2015). Clinical use of parent questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Assessing Multilingual Children: Disentangling Bilingualism form Language Impairment*.(pp.301-330). Croydon, UK: Multilingual Matters

Tuller, L., Abboud, L., Ferré, S., Fleckstein, A., Prévost, P., dos Santos, C., . . . Zebib, R. (2013). Specific language impairment and bilingualism: Assembling the pieces. In C. Rasha Hamann & E. Ruigendijk (Eds.), *Language Acquisition and Development Proceedings of GALA 2013* (pp. 533-568). Newcastle, UK: Cambridge Scholars.

Ullrich, D., Ullrich, K., & Marten, M. (2014). A longitudinal assessment of early childhood education with integrated speech therapy for children with significant language impairment in Germany. *International Journal of Language & Communication Disorders, 49*(5), 558-566.

United Kingdom Office for National Statistics. (2013). *Languages in England and Wales 2011*. Kew, London: Crown Retrieved from www.ons.gov.uk.

United States Census Bureau. (2015). Detailed languages spoken at home and ability to speak English for the population 5 Years and over for United States: 2009-2013. Retrived from http://www.census.gov/en.html

Unsworth, S. (2015). Comparing the role of input in bilingual acquisition across domains In J. Paradis & T. Gruter (Eds.), *Input and experience in bilingual development* (pp.181-202). Amsterdam: The Netherlands: John Benjamins.

Uziel-Karl, S., Kanaan, F., Yifat, R., Meir, I., Abugov, N., & Ravid, D. (2014). Hebrew and Palestinian Arabic in Israel: Linguistic frameworks and speech-language pathology services. *Topics in Language Disorders, 34*(2), 133-154. doi: 10.1097/TLD.0000000000000013

van Agt, H., Verhoeven, L., van den Brink, G., & de Koning, H. (2011). The impact on socio-emotional development and quality of life of language impairment in 8-year-old children. *Journal of Developmental Medicine and Child Neurology, 53*(1), 81-88.

Vender, M., Garraffa, M., Sorace, A., & Guasti, M. T. (2016). How early L2 children perform on Italian clinical markers of SLI: A study of clitic production and nonword repetition. *Clinical Linguistics & Phonetics, 30*(2), 150-169. doi: 10.3109/02699206.2015.1120346

Verdon, S. (2015). Enhancing practice with culturally and linguistically diverse families. *Journal of Clinical Practice in Speech-Language Pathology, 17*, 1-6.

Verdon, S., McLeod, S., & McDonald, S. (2014). A geographical analysis of speech-language pathology services to support multilingual children. *International Journal of Speech-Language Pathology, 16*(3), 304-316.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4*(4), 313-330.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language, 5*, 311-326.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, Mass: MIT Press.

Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing (CTOPP).* Austin, TX: Pro-Ed.

Washbrook, E., Waldfogel, J., Brabury, B., Corak, M., & Ghanghro, A. A. (2012). The development of young children of immigrants in Australia, Canada, the United Kingdom and the United States. *Child Development, 83*(5), 1591-1607. doi: 10.1111/j.1467-8624.2012.01796.x

Werker, J. F., & Weikum, W. M. (2006). Bilingual speech processing in infants and adults. In P. McCardle & E. Hoff (Eds.), *Childhood bilingualism : Research on infancy through school age* (pp. 1-19). Clevedon, UK: Multilingual Matters.

Wiig, E. H., Secord, W. A., & Semel, E. (2006). *Clinical Evaluation of Language Fundamentals Preschool (CELF-P)-2* (Australian and New Zealand ed.). Sydney, Australia: Pearson Clinical and Talent Assessment.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). Determining the severity of a langauge disorder. *Clinical Evaluation of Language Fundamentals (CELF) 5.* Retrieved from http://www.speechandlanguage.com/wp-content/uploads/2013/06/celf-5_determining_severity_lang_disorder.pdf

Williams, C. J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology, 14*(3), 292-305. doi: 636071

Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology, 19*(4), 298-310.

Yip, V., & Matthews, S. (2007). Theoretical framework. *The bilingual child* (pp. 22-55). Cambridge: Cambridge University Press.

Zentella, A. C. (1999). *Growing up bilingual*. Malden, MA: Blackwell.

Zhou, X.-H., McClish, D. K., & Obuchowski, N. A. (2011). *Statistical methods in diagnostic medicine* (Vol. 2). Hoboken, NJ: Wiley.

Ziethe, A., Eysholdt, U., & Doellinger, M. (2013). Sentence repetiton and digit span: Potential markers of bilingual children with suspected SLI? *Logopedics Phoniatrics Vocology, 38*, 1-10.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale, Fourth Edition (PLS-4)*. San Antonio, TX: Harcourt.

APPENDICES

Appendix A

Curtin University Human Research Ethics Committee memorandum (HR 225/2014)

Curtin University

Memorandum

To	Associate Professor Corinne Williams, Psychology
From	Professor Peter O'Leary, Chair Human Research Ethics Committee
Subject	Protocol Approval **HR 225/2014**
Date	18 December 2014
Copy	Naomi Yiftachel Psychology Dr Brody Heritage Psychology

Office of Research and Development
Human Research Ethics Committee

TELEPHONE 9266 2784
FACSIMILE 9266 3793
EMAIL hrec@curtin.edu.au

Thank you for providing the additional information for the project titled "*Assessing Dual Language Learners with Primary Language Impairment: Developing a assessment approach*". The information you have provided has satisfactorily addressed the queries raised by the Committee. Your application is now **approved**.

- You have ethics clearance to undertake the research as stated in your proposal.
- The approval number for your project is **HR 225/2014.** *Please quote this number in any future correspondence*.
- Approval of this project is for a period of four years **18-12-2014** to **18-12-2018**.
- Your approval has the following conditions:
  i) Annual progress reports on the project must be submitted to the Ethics Office.
- **It is your responsibility, as the researcher, to meet the conditions outlined above and to retain the necessary records demonstrating that these have been completed. See:** Western Australian University Sector Disposal Authority (WAUSDA).

Applicants should note the following:

It is the policy of the HREC to conduct random audits on a percentage of approved projects. These audits may be conducted at any time after the project starts. In cases where the HREC considers that there may be a risk of adverse events, or where participants may be especially vulnerable, the HREC may request the chief investigator to provide an outcomes report, including information on follow-up of participants.

The attached **Progress Report** should be completed and returned to the Secretary, HREC, C/- Office of Research & Development annually.

Our website https://research.curtin.edu.au/guides/ethics/non_low_risk_hrec_forms.cfm contains all other relevant forms including:

- Completion Report (to be completed when a project has ceased)
- Amendment Request (to be completed at any time changes/amendments occur)
- Adverse Event Notification Form (If a serious or unexpected adverse event occurs)
- Western Australian University Sector Disposal Authority (WAUSDA)

Yours sincerely

Professor Peter O'Leary
Chair Human Research Ethics Committee

Appendix B

Western Australian Department of Education approval letter (D15/0015395)

Government of **Western Australia**
Department of **Education**

Your ref :
Our ref : D15/0015395
Enquiries :

Ms Naomi Yiftachel
7/112 Hensman Road
SUBIACO WA 6008

Dear Ms Yiftachel

Thank you for your application received 13 October 2014 to conduct research on Department of Education sites.

The focus and outcomes of your research project, *Assessing Dual Language Learners with Primary Language Impairment - Developing an Assessment Approach,* are of interest to the Department. I give permission for you to approach principals to invite their participation in the project as outlined in your application. It is a condition of approval, however, that upon conclusion the results of this study are forwarded to the Department at the email address below.

Consistent with Department policy, participation in your research project will be the decision of the schools invited to participate, individual staff members, the children in those schools and their parents. A copy of this letter must be provided to principals when requesting their participation in the research. Researchers are required to sign a confidential declaration and provide a current Working with Children Check upon arrival at Department of Education schools.

Responsibility for quality control of ethics and methodology of the proposed research resides with the institution supervising the research. The Department notes a copy of a letter confirming that you have received ethical approval of your research protocol from the Curtin University Human Research Ethics Committee.

Any proposed changes to the research project will need to be submitted for Department approval prior to implementation.

Please contact Dr Adriaan Wolvaardt, Research and Evaluation Officer, on (08) 9264 5512 or researchandpolicy@education.wa.edu.au if you have further enquiries.

Very best wishes for the successful completion of your project.

Yours sincerely

ALAN DODSON
DIRECTOR
EVALUATION AND ACCOUNTABILITY

19 January 2015

151 Royal Street, East Perth Western Australia 6004

Appendix C

Alberta Language Environment Questionnaire (ALEQ)- Perth Edition 2014

Target child's initials:
Date of birth of target child:
Gender of target child:
Date:
Interpreter/ Research Assistant:
What is the target child's mother tongue?
What is the family's total weekly income on average (in dollars)?
What suburb do you live in?

Age at Test					Age of Arrival					*Months (length) of Exposure* *(copy from page 10)*
	Year	Month	Day			Year	Month	Day		
Date of testing					*Date of Arrival*					
Date of Birth					*Date of Birth*					
Chronological age					*Chronological age*					

Questions to the target child's **MOTHER:**

1a. How many years have you been in Australia?		*Converts to months:*
b. Approximate date of arrival (month/year)		

c. Was the target child born in Australia?	Yes (go to 2)	No (go to d.)		*Date of Arrival*
d. If not, did the target child come to Australia at the same time?	Yes (go to 2)	No (go to e.)		
e. If not, when did the target child come to Australia?	Year	Month	Day	

2. How much English do you speak? (Parent/Caregiver self-rating)

0	1	2	3	4
Not Fluent in English	Limited Fluency in English	Somewhat Fluent in English	Quite Fluent in English	Very Fluent in English
No understanding or speaking ability	Some understanding and can say short, simple sentences	Good understanding and can express myself on many topics	Can understand and use English adequately for work and most other situations	Understand almost everything. Very comfortable expressing myself in English in all situations
	e.g. can answer the phone in English	e.g. can go to the doctor and explain what is wrong	e.g. can communicate effectively with teachers at parent teacher interviews; could work in the service-industry ; can follow movies or television shows	

Comments/descriptions of the abilities in English:

ENG= English, MT=Mother Tongue

3. What language(s) do you speak with the target child?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

4. What language(s) does the target child speak with you (his/her mother)					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

5. What language do you speak *most often* with the other people in your home?		Score /4 Enter as a variable
0	4	
Mostly MT	Mostly ENG	

6a. Do you work outside the home?	Yes (go to c.)	No (go to b.)			Score /4
b. Are you a student?	Yes (go to c.)	No (go to 8)			
c. If yes, is the language of the workplace/school English?					
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

7. How many years of education do you have (including home country and Australia)?

Education	**Completed?**		**Years of School**
Primary	Yes	No	6/7
Secondary	Yes	No	5/6
College	Yes	No	2/3/4
University-Degree	Yes	No	3/4

Note other educational experiences here:

University- Master	Yes	No	2	
University- Phd	Yes	No	4/5	

Questions to the target child's **FATHER:**

8a. How many years have you been in Australia?		*Converts to months:*
b. Approximate date of arrival (month/year)		

9. How much English do you speak? (Parent/Caregiver self-rating)

0	1	2	3	4
Not Fluent in English	Limited Fluency in English	Somewhat Fluent in English	Quite Fluent in English	Very Fluent in English
No understanding or speaking ability	Some understanding and can say short, simple sentences	Good understanding and can express myself on many topics	Can understand and use English adequately for work and most other situations	Understand almost everything. Very comfortable expressing myself in English in all situations
	e.g. can answer the phone in English	e.g. can go to the doctor and explain what is wrong	e.g. can communicate effectively with teachers at parent teacher interviews; could work in the service-industry ; can follow movies or television shows	

Comments/descriptions of the abilities in English:

ENG= English, MT=Mother Tongue

10. What language(s) do you speak with the target child?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

11. What language(s) does the target child speak with you (his/her father)					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

12. What language do you speak *most often* with the other people in your home?		Score /4 Enter as a variable
0	4	
Mostly MT	Mostly ENG	

13a. Do you work outside the home?			Yes (go to c.)	No (go to b.)	Score /4
b. Are you a student?			Yes (go to c.)	No (go to 8)	
c. If yes, is the language of the workplace/school English?					
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

14. How many years of education do you have (including home country and Australia)?

Education	Completed?		Years of School
Primary	Yes	No	6/7
Secondary	Yes	No	5/6
College	Yes	No	2/3/4
University-Degree	Yes	No	3/4
University- Master	Yes	No	2

Note other educational experiences here:

University- Phd	Yes	No	4/5	

Questions to parents/caregivers about **OTHER** adults and children in the home

15a. Are there other adults in the home? For example, a grandparent?	Yes (go to b)	No (go to 21)
15b. If yes, how many?	(go to 16)	

16. If yes, is one of these adults the child's primary caregiver?	Yes (go to 17)	No (go to 19)

17. If yes, what language(s) does the primary caregiver speak with the target child?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

18. If applicable, what language(s) does the target child speak with the primary caregiver?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

19a. If there are other adults in the home (who are not the primary caregiver), do they regularly interact with the target child?			Yes (go to b)	No (go to 21)	Score /4 Include in Language Use Score *If there is more than one adult in this category, record a value for each adult.*
19b. If yes, what language(s) does the adult relative(s) speak with the target child?					
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

<table>
<tr><td colspan="5">20. If applicable, what language(s) does the target child speak with the adult relative(s) who are not the primary caregiver?</td><td rowspan="3">Score /4

Include in Language Use Score

If there is more than one adult in this category, record a value for each adult.</td></tr>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>ENG never
MT always</td><td>ENG seldom
MT usually</td><td>ENG 50%
MT 50%</td><td>ENG usually
MT seldom</td><td>ENG almost always
MT almost never</td></tr>
</table>

21a. Does the target child have brothers or sisters?	Yes	No (go to 28)
How many?		
If they are over the age of 2, answer the following questions		

22. **Sibling 1**:		Older		Younger
	Gender	Male		Female
	Date of Birth	Day	Month	Year

<table>
<tr><td colspan="5">23. What language(s) does Sibling 1 speak with the target child?</td><td rowspan="3">Score /4

Include in Language Use Score</td></tr>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>ENG never
MT always</td><td>ENG seldom
MT usually</td><td>ENG 50%
MT 50%</td><td>ENG usually
MT seldom</td><td>ENG almost always
MT almost never</td></tr>
</table>

<table>
<tr><td colspan="5">24. What language(s) does the target child speak with Sibling 1?</td><td rowspan="3">Score /4

Include in Language Use Score</td></tr>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>ENG never
MT always</td><td>ENG seldom
MT usually</td><td>ENG 50%
MT 50%</td><td>ENG usually
MT seldom</td><td>ENG almost always
MT almost never</td></tr>
</table>

25. **Sibling 2:**		Older		Younger
	Gender	Male		Female
	Date of Birth	Day	Month	Year

26. What language(s) does **Sibling 2** speak with the target child?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

27. What language(s) does the target child speak with **Sibling 2**?					Score /4 Include in Language Use Score
0	1	2	3	4	
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never	

Language Use in the Home- Score sheet
Higher scored (greater than 0.5) indicate more of a shift towards English use in the home. Lower scores (less than 0.5) indicate maintenance of the Mother Tongue.

<table>
<tr><td></td><td>Score</td><td></td><td>Score</td><td></td><td></td></tr>
<tr><td>Mother to Child (Q3)</td><td></td><td>Child to Mother (Q4)</td><td></td><td></td><td></td></tr>
<tr><td>Father to Child (Q10)</td><td></td><td>Child to Father (Q11)</td><td></td><td></td><td></td></tr>
<tr><td>Other Adult to Child – Primary Caregiver (Q17)</td><td></td><td>Child to Other Adult –Primary Caregiver (Q18)</td><td></td><td></td><td></td></tr>
<tr><td>Other Adult to Child- not Primary Caregiver (Q19b)</td><td></td><td>Child to Other Adult- not Primary Caregiver (Q20)</td><td></td><td></td><td></td></tr>
<tr><td>Sibling 1 to Child (Q23)</td><td></td><td>Child to sibling 1 (Q24)</td><td></td><td></td><td></td></tr>
<tr><td>Sibling 2 to Child (Q26)</td><td></td><td>Child to sibling 2 (Q27)</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">TOTAL:

<u>Sum of scores</u>
Number of scores x4</td><td></td><td rowspan="2">TOTAL:

<u>Sum of scores</u>
Number of scores x4</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Language Use in the Home (fraction from above)</td><td>↓</td><td>+</td><td>↓</td><td rowspan="2">=</td><td></td></tr>
<tr><td></td><td>+</td><td></td><td></td></tr>
</table>

Questions to parents about the TARGET CHILD

28a. Does the target child attend an **ONLY ENGLISH** speaking preschool/ school?	Yes (go to 29)	No (go to b)

28b. If not, what part of the day is in English?

0	0.25	0.5	0.75	1
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never

29. At what age did the target child start receiving <u>consistent and significant</u> exposure to English?
Consistent and significant= English-language day-care or babysitter or at least three days per week or equivalent part-time. English-language school of any kind counts as consistent and significant exposure.

	Day	Month	Year		Day	Month	Year
Age				Date of entry into program			

Age of Exposure

	Day	Month	Year
Date of Exposure			
Date of Birth			
Age of Exposure			
Additional information (eg. 6 months or more in non-English environment)			

Length (Months) of Exposure

Age at test in months	
	-
Age of Exposure	
	=

30. What literacy and other language activities does the target child do <u>each week?</u>

Reading: includes having books read to them/looking at books.
Computer/Tablet: includes internet, games, online storybooks that involve language.
Movies: on computer or TV.
Extra-curricular: outside of school.

	ENGLISH			MOTHER TONGUE		
Activities	everyday	At least once a week	Almost never/never	everyday	At least once a week	Almost never/never
Reads books or magazines	2	1	0	2	1	0
Uses a computer	2	1	0	2	1	0
Watches TV or movies	2	1	0	2	1	0
Storytelling (tells about his/her day or made-up stories)	2	1	0	2	1	0
Singing Songs	2	1	0	2	1	0
TOTAL						
GRAND TOTAL BY LANGUAGE	/10			/10		

31a. What literacy and other language activities in the MOTHER TONGUE does the target child do each week? For example, weekend language school, bilingual playgroup.	1)
	2)
	3)
	4)

How often?					Score /4
0	1	2	3	4	
Child receives little or no formal instruction in his/her mother tongue	Child attends mother tongue classes outside of school (once a week)	Child attends mother tongue classes outside of school (more than once a week)	Child is registered in a part-time bilingual program at school or preschool	Child registered in a full-time bilingual program at school or preschool	

31b. Does your child attend any extra-curricular activities?

	Every day	At least once a week	Almost never/ never
English	2	1	0
Mother Tongue	2	1	0

ENG /2

MT /2

32. What are the languages spoken between your child and the friends he/she plays with regularly?

0	1	2	3	4
ENG never MT always	ENG seldom MT usually	ENG 50% MT 50%	ENG usually MT seldom	ENG almost always MT almost never

ENG score /4

4	-	ENG score	=	MT score

Calculating Richness Scores:

English Richness Scores	
Q30	/10
Q31b	/2
Q32	/4
Total	/16

Mother Tongue Richness Scores	
Q30	/10
Q31a	/4
Q31b	/2
Q32	/4
Total	/20

Appendix D

Alberta Language and Development Questionnaire (ALDeQ)- Perth Edition 2014

Early Milestones

<table>
<tr><td>1</td><td>When did your child first begin to walk?</td><td>months</td><td>3=≤15 months
0=>16 months
Score: /3</td></tr>
<tr><td rowspan="5">2</td><td>How old was your child when he/she first spoke a word?</td><td>months</td><td>6=≤15 months
4=16-24 months
0=>25 months
Maybe 0=≥26
Score: /6</td></tr>
<tr><td colspan="3">Examples of the child’s first words (with translations):</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"></td></tr>
<tr><td rowspan="5">3</td><td>How old was your child when he/she began to put words together to make short sentences? (for example: “more milk”, “bye mum”</td><td>months</td><td>6=≤24 months
4=25-30 months (2 to 2;6)
0=≥31 (closer to age 3 or older)
Score: /6</td></tr>
<tr><td colspan="3">Examples of the child’s first sentences (with translations):</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"></td></tr>
<tr><td>4</td><td colspan="2">When you think about other children you know at that age, do you think your child was different when he/she started to use language?
<table><tr><td>3</td><td>2</td><td>1</td><td>0</td></tr><tr><td>Not different at all</td><td>A little different</td><td>Quite different</td><td>Very different</td></tr></table></td><td>If child is said to be better or quicker= 3.

Score: /3</td></tr>
<tr><td></td><td colspan="2">To calculate the subtotal for section A, add the total possible score for all questions answered as the denominator. Then add the scores for the parent’s responses as the numerator. If all questions were answered, the denominator would be 18.</td><td>SUB TOTAL A

/</td></tr>
</table>

Current Abilities in the First Language

*Compare the child to other duel language learners/bilingual children, except for question 10.

<table>
<tr><td>5</td><td>Compared to other children of the same age, how do you think that your child expresses him/herself?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Not very well</td><td>A little less</td><td>The same</td><td>Very good/better/one of the best</td></tr>
</table>
</td><td>Score:
/3</td></tr>
<tr><td>6</td><td>Compared with other children of the same age, how do you think your child pronounces words?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Not very clearly</td><td>Sometimes not clear</td><td>The same</td><td>Very clear/one of the best</td></tr>
</table>
</td><td>Score:
/3</td></tr>
<tr><td>7</td><td>Is it easy for your family or friends to have conversation with your child?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>No, very hard</td><td>Sometimes not easy</td><td>Easy enough</td><td>Very easy</td></tr>
</table>
</td><td>Score:
/3</td></tr>
<tr><td>8</td><td>Compared with other children of the same age, does your child have difficulty producing correct sentences? (have appropriate vocabulary, correct grammar, long enough sentences to get ideas across)
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>A lot of difficulties</td><td>Some difficulties</td><td>same</td><td>No difficulties, maybe better</td></tr>
</table>
</td><td>Score:
/3</td></tr>
<tr><td>9</td><td>Are you satisfied with how your child speaks your mother tongue?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Not satisfied at all</td><td>Maybe no satisfied</td><td>satisfied</td><td>Completely satisfied</td></tr>
</table>
</td><td>Score:
/3</td></tr>
<tr><td>10a</td><td>Do you think your child speaks your mother tongue like the children in the home country?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Not as good as home country</td><td>Sort of like home country with some differences</td><td>Mostly yes-close to home country</td><td>Yes-better or just like home country</td></tr>
</table>
</td><td>Score:
/3</td></tr>
</table>

10b	Why are you not satisfied? Why do you think your child is different from children in the home country? Do you think he/she may be losing the mother tongue in favour of English?	
To calculate the subtotal for section B, add the total possible score for all questions answered as the denominator. Then add the scores for the parent's responses as the numerator. If all questions were answered, the denominator would be: 18		**SUB TOTAL B** **/**

Behaviour Patterns and Activity Preferences

<table>
<tr>
<td>11</td>
<td>Does your child like to read books or have books read to him/her?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Never</td><td>Rarely</td><td>Sometimes</td><td>Very much</td></tr>
</table>
</td>
<td>Score:
/3</td>
</tr>
<tr>
<td>12</td>
<td>How does your child read and write in the MOTHER TONGUE compared with other children his/her age? (If young, does he/she have interest in numbers and alphabet/characters and some word recognition)
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Noticeably worse than other children</td><td>Not as well as other children</td><td>Same as other children</td><td>Very well, maybe better</td></tr>
</table>
</td>
<td>If never been taught, omit this question.

Score:
/3</td>
</tr>
<tr>
<td>13</td>
<td>What kind of activities does he/she like to do?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Other (ex. Television, video games, dress-up or childish games for age).</td><td>Physical games (ex. Ball games, swimming)</td><td>Cognitive games (ex. Puzzles, drawing, card games, cares, computer games-not movies)</td><td>Language games (ex. Reading, writing, playing school)</td></tr>
</table>
</td>
<td>If more than one category comes up, add total and divide by number of scores.

Score:
/3</td>
</tr>
<tr>
<td rowspan="2">14</td>
<td>How quickly/easily does your child learn new things? For example: sports, puzzles, rules of board games.
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Long time, sometimes never learns</td><td>Needs help and time to learn new things</td><td>A few tries</td><td>Same day/immediately</td></tr>
</table>
</td>
<td>Score:
/3</td>
</tr>
<tr>
<td colspan="2">Examples of child learning new things:</td>
</tr>
</table>

<table>
<tr><td>15</td><td>What are the activity patterns shown by your child? An activity can be any of the following: games, eating, dancing, watching TV etc.
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>More than one/many activities at a time and seldom finishes any of them</td><td>Two to four activities at a time and finishes one</td><td>One or two activities at a time and finishes one</td><td>One activity at a time and finishes it.</td></tr>
</table></td><td>Score:
/3</td></tr>
</table>

<table>
<tr><td>16</td><td>Does your child get frustrated when he/she cannot communicate his/her ideas?
<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Frequently</td><td>Often</td><td>Sometimes</td><td>Not at all</td></tr>
</table></td><td>Score:
/3</td></tr>
<tr><td colspan="2">To calculate the subtotal for Section C, add the total possible score for all questions answered as the denominator. Then add the scores for the parent’s responses as the numerator. If all questions were answered, the denominator would be: 18</td><td>SUB TOTOAL
/</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

Family History

<table>
<tr><td>17a</td><td colspan="2">Can you tell us about your relatives? What kind of education and professions do they have in the home country? (This will provide context for the next question).</td></tr>
<tr><td>17b</td><td>Did all the members of your family finish school?
<table><tr><td>0</td><td>3</td></tr><tr><td>No</td><td>Yes</td></tr></table>
If not, why?</td><td>If parent gives a reason that is environmental or externat (war, funds) DO NOT score this question

Score: /3</td></tr>
<tr><td>18</td><td>Is there anyone among the child’s immediate family or other relative who had difficulties learning to read and write, in speaking and pronunciation, slow to learn to talk?<table><tr><td>0</td><td>3</td><td>6</td></tr><tr><td>Yes, definitely (go to question 19)</td><td>Yes, possibly (go to question 19)</td><td>No indication</td></tr></table></td><td>If uncomfortable answering or does not know the answer, DO NOT score the question.

If reason is environmental or externat (war, funds) DO NOT score this question

Score: /6</td></tr>
<tr><td colspan="2">To calculate the subtotal for Section C, add the total possible score for all questions answered as the denominator. Then add the scores for the parent’s responses as the numerator. If all questions were answered, the denominator would be: 9</td><td>SUB TOTOAL

/</td></tr>
</table>

19	Fill in this section if parents have a family history. These questions are a guide to prompt more detailed information.

	Brother or sister		Father		Mother		Relative of the father		Relatives of the mother	
Difficulties in school or learning	Y	N	Y	N	Y	N	Y	N	Y	N
Language or pronunciation problems (including grammar and stuttering)	Y	N	Y	N	Y	N	Y	N	Y	N
Special education classes	Y	N	Y	N	Y	N	Y	N	Y	N
Speech and language therapy	Y	N	Y	N	Y	N	Y	N	Y	N
Problems following directions or understanding questions	Y	N	Y	N	Y	N	Y	N	Y	N
Problems reading or learning to read	Y	N	Y	N	Y	N	Y	N	Y	N
Difficulty learning English	Y	N	Y	N	Y	N	Y	N	Y	N
Repeated one or more grades in school	Y	N	Y	N	Y	N	Y	N	Y	N

Calculating the ALDeQ Total Score:

Section A	
Section B	
Section C	
Section D	
Total	

	Total score mean	**SD**	**1 SD range**	**-1.25 SD**	**-1.5 SD**	**-2 SD**	**-2.5 SD**	**-3 SD**
Canada	**.81**	**.12**	**.69-.93**	**.66**	**.63**	**.57**	**.51**	**.45**
Australia	**.81**	**.11**	**.70-.92**	**.67**	**.64**	**.59**	**.53**	**.48**

If a child's score is <u>-1.25</u> SD or lower, it is advisable to calculate and check the section scores, and the parents' answer to question 10b. If the only score NOT in normative range is in section B, and parents indicated that first language loss could be taking place, then it is possible the low ALDeQ score is NOT suggestive of PLI.

Section	**A**			**B**			**C**			**D**		
	Mean	SD	-1 SD	Mean	SD	-1 SD	Mean	SD	-1 SD	Mean	SD	-1 SD
Canada	.90	19	.72	.69	.26	.44	.82	.13	.69	.83	.30	.57
Australia	.79	.25	.54	.87	.14	.73	.67	.20	.47	.83	.22	.72

Norming sample in Canada: 179 DLLs with a mean age of 70 months (5;8 years) with a range of 63-77 months (5;2 to 6;4 years). The average exposure to English in school or preschool was 17 months (range of 7 to 27 months).

Pilot study in Australia: 14 DLLs with a mean age of 80.43 months (6;7 years) with a range of 5;3 years to 8;7 years. Children had an average of 46% use of mother tongue a week. All participants were in year 1 at English speaking schools.

Appendix E

Information letter for principals of mainstream schools

Naomi Yiftachel

Masters Student

School of Psychology & Speech Pathology

Curtin University of Technology

GPO Box U 1987,

Perth, WA 6845

Dear Principal,

<u>Assessment of language impairment amongst multilingual children</u>

My name is Naomi Yiftachel. I am a qualified Speech Language Pathologist undertaking a research project under the supervision of Associate Professor Cori Williams, at Curtin University. My project aims to better evaluate language problems in multilingual children aged 5-7 years old.

I am currently recruiting typically developing multilingual children (children who function in two or more languages on a daily basis) for my study. I would like to invite your school to take part in the project.

What does participation in the research project involve? I seek to access preschool and Year 1 multilingual pupils at your school who have given consent by their caregivers. I will take each participating child out of class (to a quiet room or space in your school) for a 45 min session. This session involves a hearing screening test (including tympanometry), a nonverbal abilities test (Raven’s Coloured Progressive Matrices and two language processing tasks (nonword repetition and sentence imitation). The assessment will be administered in a fun, playful way and will be audio recorded. If a child’s hearing or assessment results raise concern, his/her parents/caregiver will be notified and encouraged to consult with me or the child's teacher.

We will keep your school's involvement in the administration of the research procedures to a minimum. However, we ask for your school's assistance in sending home and collecting information packs (letters and consent forms) for pupils who fit the profile: typically developing children who function in two or more languages on a daily basis. We also ask schools to provide a quiet space or room to conduct assessment sessions.

Since the parents/caregivers of the participants may not be fluent in English, you will be consulted regarding the need to translate information packs to the main language subgroups of children attending your school before recruitment commences.

In addition, I will interview the child's parent or caregiver about language development and language environment. This will take approximately 45 minutes. It can be conducted in English or another language of their choice (with the help of a translator) and take place at home or over the phone. This interview may be audio recorded.

How can this project benefit the participants, your school and the wider community? This study will benefit the growing number of multilingual children in Perth schools and the professionals that work with them by exploring a new assessment approach that sensitively addresses multilingual language development. The results may assist in correctly distinguishing between multilingual children with different, but typical, language to those with a deficit who may need intervention. Current assessment tools lack this tailored approach and have been found to result in misidentification, leaving some children without timely intervention and others wrongly assigned to special education programmes.

To what extent is participation voluntary and what are the implications of withdrawing that participation?

Participation in this research project is entirely voluntary. We require written consent from both the caregiver and the pupil. Participants are able to withdraw their involvement at any time without adverse consequences.

What will happen to the information collected and is privacy and confidentiality assured?

If you consent to take part in this project and a participating student requires a full intelligence assessment within the next 2 years, it is important to inform the person conducting the assessment that the student has already been assessed on the Raven's Coloured Progressive Matrices. Therefore I will provide the school with a list of students who received this test. This information will help professionals to determine the most appropriate assessment to use in the future if needed.

The research team will ensure that any information that identifies participants or their schools will be removed from the data collected. The data will be kept in a secure cabinet at Curtin University and on a password protected external hard drive for a minimum of 7 years and will only be accessible to those directly involved in the project. All data will then be destroyed by shredding hard documents and deleting computer files.

The identity of participants and the school will not be disclosed at any time, except in circumstances that require reporting under the Department of Education *Child Protection* policy, or where the research team is legally required to disclose information. Participant privacy, and the confidentiality of information disclosed by participants, is assured at all other times.

I will write this project's results in my master's thesis and it might be published in a journal or presented at a conference. The data will be used for this project, and will not be used in any extended or future research without first obtaining explicit written consent from participants. Consistent with Department of Education policy, your school will receive a summary of the research findings.

Is this research approved?

This study has been approved by the Curtin University Human Research Ethics Committee (Approval Number HR 225/2014). The Committee is comprised of members of the public, academics, lawyers, doctors and pastoral carers. Its main role is to protect participants. If needed, verification of approval can be obtained either by writing to the Curtin University Human Research Ethics Committee, c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth, 6845 or by telephoning 9266 2784 or by emailing hrec@curtin.edu.au. The research has met the policy requirements of the Department of Education (copies of these letters are attached).

Do all members of the research team who will have contact with the pupils have a Working with Children Check? Yes. Please find evidence of my Working with Children Check attached to this letter.

Who do I contact if I wish to discuss the project further? If you would like to discuss any aspect of this study with a member of the research team, please contact my supervisor from the details provided below. If you wish to speak with an independent person for verification of ethics approval, please contact the Curtin University Ethics Committee (see details below).

How do I indicate my willingness for the school to be involved? If you have had all questions about the project answered to your satisfaction, and are willing for the school to participate, please complete the **Consent Form** on the following page.

This information letter is for you to keep.

Thank you

Naomi Yiftachel

naomi.yiftachel@postgrad.curtin.edu.au

Associate Professor Cori Williams
Curtin University
Tel: 9266 7865
Email: c.j.williams@curtin.edu.au

Human Research Ethics Committee
Curtin University
Tel: 9266 2784
Email:hrec@curtin.edu.a

CONSENT FORM

(For Principal)

- I have read this document and understand the project, as described within it.
- I am satisfied with the answers I received for any questions I may have had.
- I am willing for this school to become involved in the research project, as described.
- I understand that participation in the project is entirely voluntarily.
- I understand that the school is free to withdraw its participation at any time, without consequence.
- I understand that this research may be published in a journal or presented at a conference, provided that the participants or the school are not identified in any way.
- I understand that the school will be provided with a copy of the findings from this research upon its completion.

Name of Site Manager (printed):	**Name of School:**
Signature:	**Date: / /**

Appendix F

Information letter for principals of language development centres (LDCs)

Naomi Yiftachel

Masters Student

School of Psychology & Speech Pathology

Curtin University of Technology

GPO Box U 1987,

Perth, WA 6845

Dear Principal,

<u>Assessment of language impairment amongst multilingual children</u>

My name is Naomi Yiftachel. I am a qualified Speech Language Pathologist undertaking a research project under the supervision of Associate Professor Cori Williams, at Curtin University. My project aims to better evaluate language problems in multilingual children aged 5-7 years old.

I am currently recruiting multilingual children (children who function in two or more languages on a daily basis) with language difficulties. I would like to invite your school to take part in the project.

What does participation in the research project involve? I seek to access preschool and Year 1 multilingual pupils at your school who have given consent by their caregivers. I will take each participating child out of class (to a quiet room or space in your school) for a 45 min session. This session involves a hearing screening test (including tympanometry), a nonverbal abilities test (Raven's Coloured Progressive Matrices) and two language processing tasks (nonword repetition and sentence imitation). The assessment will be administered in a fun, playful way and will be audio recorded. If a child's hearing is not within the normal range, his/her parents/caregiver will be notified and referred to the suitable support services. If needed, this information will be conveyed in the parent's/caregiver's dominant language with the assistance of translators and interpreters.

We will keep your school's involvement in the administration of the research procedures to a minimum. However, we ask for your school's assistance in sending home and collecting information packs (letters and consent forms) for pupils who fit the profile: children who function in two or more languages on a daily basis **and** have language difficulties. We also ask schools to provide a quiet space or room to conduct assessment sessions.

Since the parents/caregivers of the participants may not be fluent in English, you will be consulted regarding the need to translate information packs to the main language subgroups of children attending your school before recruitment commences.

In addition, I will interview the child's parent or caregiver about language development and language environment. This will take approximately 45 minutes. It can be conducted in English or another language of their choice (with the help of a translator) and take place at home, at school or over the phone. This interview may be audio recorded.

How can this project benefit the participants, your school and the wider community? This study will benefit the growing number of multilingual children in Perth schools and the professionals that work with them by exploring a new assessment approach that sensitively addresses multilingual language development. The results may assist in correctly distinguishing between multilingual children with different, but typical, language to those with a deficit who may need intervention. Current assessment tools lack this tailored approach and have been found to result in misidentification, leaving some children without timely intervention and others wrongly assigned to special education programmes.

To what extent is participation voluntary and what are the implications of withdrawing that participation?

Participation in this research project is entirely voluntary. We require written consent from both the caregiver and the pupil. Participants are able to withdraw their involvement at any time without adverse consequences.

What will happen to the information collected and is privacy and confidentiality assured?

Following parent/caregiver permission, child assessment results will be provided to their school. This information will be useful to the child's teacher and speech pathologist to better understand the nature of the child's language difficulties. In addition, if a participating student requires a full intelligence assessment within the next 2 years it is important to inform the person conducting the assessment that the student has already been assessed on the Raven's Coloured Progressive Matrices. Therefore I will provide the school with a list of students who received this test. This information will help professionals to determine the most appropriate assessment to use in the future if needed.

The research team will ensure that any information that identifies participants or their schools will be removed from the data collected. The data will be kept in a secure cabinet at Curtin University and on a password protected external hard drive for a minimum of 7 years and will only be accessible to those directly involved in the project. All data will then be destroyed by shredding hard documents and deleting computer files.

The identity of participants and the school will not be disclosed at any time, except in circumstances that require reporting under the Department of Education *Child Protection* policy, or where the research team is legally required to disclose information. Participant privacy, and the confidentiality of information disclosed by participants, is assured at all other times.

I will write this project's results in my master's thesis and it might be published in a journal or presented at a conference. The data will be used for this project, and will not be used in any extended or future research without first obtaining explicit written consent from participants. Consistent with Department of Education policy, your school will receive a summary of the research findings.

Is this research approved?

This study has been approved by the Curtin University Human Research Ethics Committee (Approval Number HR 225/2014). The Committee is comprised of members of the public, academics, lawyers, doctors and pastoral carers. Its main role is to protect participants. If needed, verification of approval can be obtained either by writing to the Curtin University Human Research Ethics Committee, c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth, 6845 or by telephoning 9266 2784 or by emailing hrec@curtin.edu.au. The research has met the policy requirements of the Department of Education (copies of these letters are attached).

Do all members of the research team who will have contact with the pupils have a Working with Children Check? Yes. Please find evidence of my Working with Children Check attached to this letter.

Who do I contact if I wish to discuss the project further? If you would like to discuss any aspect of this study with a member of the research team, please contact my supervisor from the details provided below. If you wish to speak with an independent person for verification of ethics approval, please contact the Curtin University Ethics Committee (see details below).

How do I indicate my willingness for the school to be involved? If you have had all questions about the project answered to your satisfaction, and are willing for the school to participate, please complete the **Consent Form** on the following page.

This information letter is for you to keep.

Thank you

Naomi Yiftachel

naomi.yiftachel@postgrad.curtin.edu.au

Associate Professor Cori Williams

Curtin University

Tel: 9266 7865

Email: c.j.williams@curtin.edu.au

Human Research Ethics Committee

Curtin University

Tel: 9266 2784

Email: hrec@curtin.edu.a

CONSENT FORM

(For Principal)

- I have read this document and understand the project, as described within it.
- I am satisfied with the answers I received for any questions I may have had.
- I am willing for this school to become involved in the research project, as described.
- I understand that participation in the project is entirely voluntarily.
- I understand that the school is free to withdraw its participation at any time, without consequence.
- I understand that with parents'/caregivers' consent, assessment results will be provided to the school.
- I understand that this research may be published in a journal, or presented at a conference provided that the participants or the school are not identified in any way.
- I understand that the school will be provided with a copy of the findings from this research upon its completion.

Name of Site Manager (printed):	**Name of School:**
Signature:	**Date: / /**

Appendix G

Parent information letter for participants from mainstream schools and the general community

Naomi Yiftachel

Masters Student

School of Psychology & Speech Pathology

Curtin University of Technology

GPO Box U 1987,

Perth, WA 6845

Dear Parent/Caregiver

Assessment of language impairment amongst multilingual children

My name is Naomi Yiftachel. I am a qualified Speech Language Pathologist doing a research project under the supervision of Associate Professor Cori Williams, at Curtin University. My project aims to better evaluate language problems in multilingual children aged 5-7 years old.

I am currently recruiting typically developing multilingual children (children who function in two or more languages on a daily basis) that have no pre-existing cognitive, sensory, psychological or neurological impairment. I would like to invite you and your child to take part in the project.

What does participation in the research project involve?

This study will consist of two major parts: a parent/caregiver interview and a child language assessment.

Your child is invited to participate in a session that involves a hearing screening test (with tympanometry), a nonverbal abilities test (the Raven's Coloured Progressive Matrices) and language tasks in which your child will be asked to mimic made-up words and repeat sentences. The tasks will be administered in a fun, playful way. The session will take place in a quiet room or space at his or her school over 45 minutes. The times your child participates in assessment sessions will be negotiated with their class teacher in advance, to ensure that they provide minimal disturbance to classroom activities. Sessions will be audio recorded with your approval.

You will be notified and provided with the appropriate information if there are concerns about your child's hearing or task results.

Also, you are invited to an interview about your child's language development and language environment. This will take approximately 45 minutes. It can be conducted in English or another language of your choice (with the help of a translator) and take place at home or over the phone. Please indicate your preference on the attached consent form. Interviews may be audio recorded with your approval.

Do my child and I have to take part?

No. Participation in this research project is entirely voluntary. This decision should always be made completely freely. All decisions made will be respected by members of the research team without question.

Your child has also been provided with a letter from us that we encourage you to discuss with him/her.

What if either of us was to change our mind?

If a decision is made to participate, it will need to be made by __________ for you and your child to be included in the project.

Once a decision is made to participate, either you or your child can change your mind at any time.

Your decision will not affect your family's relationship with your child's teacher or school. Even after taking part, we can destroy any information we have collected about your child, unless we have already published a paper or report on the study.

What will happen to the information collected, and is privacy and confidentiality assured?

Your privacy is very important. We will remove you and your child's name and any information that could be used to identify him/her, or you, from the information we collect. No information about you, your child or the school your child attends will be published. We will be using an audio recorder to record you and your child's speech but only those involved in this research can listen to the recording. We will safely store the information for a minimum of 7 years so that only the researchers can see it, and then it will be destroyed by shredding hard data and deleting computer files.

If you consent to take part in this project and your child requires a full intelligence assessment within the next 2 years, it is important to inform the person conducting the assessment that the student has already been assessed on the Raven's Coloured Progressive Matrices. This information will help professionals to determine the most appropriate assessment to use in the future if needed.

We assure the privacy and confidentiality of information you provide, except in circumstances that require reporting under the Department of Education Child Protection policy, or where we are legally required to disclose that information.

The data will be used only for this project, and will not be used in any extended or future research without first obtaining explicit written consent from you and your child.

I will write this project's results in my master's thesis and it might be published in a journal, or presented at a conference, but always without any identifying information. The school will be given a summary of the findings for you to see if you would like to know what the research found.

Is this research approved?

This study has been approved by the Curtin University Human Research Ethics Committee (Approval Number HR 225/2014). The Committee is comprised of members of the public, academics, lawyers, doctors and pastoral carers. Its main role is to protect participants. If needed, verification of approval can be obtained either by writing to the Curtin University Human Research Ethics Committee, c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth, 6845 or by telephoning 9266 2784 or by emailing hrec@curtin.edu.au. The research has met the policy requirements of the Department of Education.

How do I know that the people involved in this research have all the appropriate documentation to be working with children?

All persons undertaking research activities on Department sites must complete a Confidential Declaration. Under the Working with Children (Criminal Record Checking) Act 2004, people undertaking research that involves contact with children must undergo a Working with Children Check. Evidence that these checks are current has been provided to the Principal of your school your child attends.

Who do I contact if I wish to discuss the project further?

If you would like to discuss any aspect of this study with a member of the research team, please contact my supervisor using the details provided below. If you wish to speak with an independent person about how the project is being conducted or was conducted, please contact the Curtin University Human Research Ethics Committee (contact details below).

How do my child and I become involved?

Please ensure that you:

Discuss what it means to take part in the project with your child before you both make a decision; and

Take up my invitation to ask any questions you may have about the project.

Once all questions have been answered to your satisfaction, and you and your child are both willing for him/her to become involved, please complete the Consent Form on the following page (your child is also asked to complete the Consent Form attached to his/her letter) and return it to your child's teacher.

This project information letter is for you to keep.

Thank you,

Naomi Yiftachel

naomi.yiftachel@postgrad.curtin.edu.au

Associate Professor Cori Williams

Curtin University

Tel: 9266 7865

Email: c.j.williams@curtin.edu.au

Human Research Ethics Committee

Curtin University

Tel: 9266 2784

Email: hrec@curtin.edu

CONSENT FORM

(For Parent/Caregiver)

- I understand the information about the project.

- I have asked any questions I may have had and I am happy with the answers.

- I understand that it is up to me to decide whether my child and I take part.

- I understand what it means for me to participate in this project and am happy to take part in it.

- I have discussed with my child what it means to participate in this project. He/she has explicitly indicated a willingness to take part, as indicated by his/her completion of the child consent form.

- I give permission to record me and my child's speech.

- I give permission to test my child's hearing including a tympanometry test.

- I give permission for results of the Raven's Coloured Matrices test to be given to other professionals if my child needs a full intelligence assessment within two years from the date of testing.

- I understand that we can pull out of the project at any time without affecting the family's relationship with my child's teacher or school.

- I am happy for the project to be presented at a conference and possibly published in a journal. I know that my child's school will not be identified in any way.

- I understand that I can request a summary of the findings after the project is finished.

Name of Child (printed):	**Name of School the child attends:**
I would like to be interviewed in: **English** **Other (please specify):**	**I would like the interview to take place:** **Over the phone** **At my child's school** **At my home- face to face (please add address):**
Signature of Parent/Caregiver:	**Date: / /**
Phone number:	**E-mail address:**

Appendix H

Parent information letter for participants from LDCs

Naomi Yiftachel

Masters Student

School of Psychology & Speech Pathology

Curtin University of Technology

GPO Box U 1987,

Perth, WA 6845

Dear Parent/Caregiver

<u>Assessment of language impairment amongst multilingual children</u>

My name is Naomi Yiftachel. I am a qualified Speech Language Pathologist doing a research project under the supervision of Associate Professor Cori Williams, at Curtin University. My project aims to better evaluate language problems in multilingual children aged 5-7 years old.

I am currently recruiting multilingual children (children who function in two or more languages on a daily basis) with language difficulties that have **no** pre-existing cognitive, sensory, psychological or neurological impairment.

I would like to invite you and your child to take part in the project.

What does participation in the research project involve?

This study will consist of two major parts: a parent/caregiver interview and a child language assessment.

Your child is invited to participate in a session that involves a hearing screening test (with tympanometry), a nonverbal abilities test (Raven's Coloured Progressive Matrices) and language tasks in which your child will be asked to mimic made-up words and repeat sentences. The tasks will be administered in a fun, playful way. The session will take place in a quiet room or space at his or her school over 45 minutes. The times your child participates in assessment sessions will be negotiated with their class teacher in advance, to ensure that they provide minimal disturbance to classroom activities. Sessions will be audio recorded with your approval.

You will be notified and referred to suitable support services if there are concerns about your child's hearing.

Also, you are invited to an interview about your child's language development and language environment. This will take approximately 45 minutes. It can be conducted in English or another language of your choice (with the help of a translator) and take place at home, at your child's school, or over the phone. Please indicate your preference on the attached consent form. Interviews may be audio recorded with your approval.

Do my child and I have to take part?

No. Participation in this research project is entirely voluntary. This decision should always be made completely freely. All decisions made will be respected by members of the research team without question.

Your child has also been provided with a letter from us that we encourage you to discuss with him/her.

What if either of us was to change our mind?

If a decision is made to participate, it will need to be made by **__________** for you and your child to be included in the project.

Once a decision is made to participate, either you or your child can change your mind at any time.

Your decision will not affect your family's relationship with your child's teacher or school. Even after taking part, we can destroy any information we have collected about your child, unless we have already published a paper or report on the study.

What will happen to the information collected, and is privacy and confidentiality assured?

Your privacy is very important. With your permission, your child's assessment results will be provided to their school. This information will be useful for your child's teacher and speech pathologist. The results of this study will lead to a better understanding of language difficulties amongst multilingual children.

If you consent to take part in this project and your child requires a full intelligence assessment within the next 2 years, it is important to inform the person conducting the assessment that the student has already been assessed on the Raven's Coloured Progressive Matrices. This information will help professionals to determine the most appropriate assessment to use in the future if needed.

The research team will remove you and your child's name and any information that could be used to identify him/her, or you, from the information we collect. No information about you, your child or the school your child attends will be published. We will be using an audio recorder to record you and your child's speech but only those involved in this research can listen to the recording. We will safely store the information for a minimum of 7 years so that only the researchers can see it, and then it will be destroyed by shredding hard data and deleting computer files.

We assure the privacy and confidentiality of information you provide, except in circumstances that require reporting under the Department of Education *Child Protection* policy, or where we are legally required to disclose that information.

The data will be used only for this project, and will not be used in any extended or future research without first obtaining explicit written consent from you and your child.

I will write this project's results in my master's thesis and it might be published in a journal, or presented at a conference, but always without any identifying information. The school will be given a summary of the findings for you to see if you would like to know what the research found.

Is this research approved?

This study has been approved by the Curtin University Human Research Ethics Committee (Approval Number HR 225/2014). The Committee is comprised of members of the public, academics, lawyers, doctors and pastoral carers. Its main role is to protect participants. If needed, verification of approval can be obtained either by writing to the Curtin University Human Research Ethics Committee, c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth, 6845 or by telephoning 9266 2784 or by emailing hrec@curtin.edu.au. The research has met the policy requirements of the Department of Education.

How do I know that the people involved in this research have all the appropriate documentation to be working with children?

All persons undertaking research activities on Department sites must complete a Confidential Declaration. Under the *Working with Children* (Criminal Record Checking) Act 2004, people undertaking research that involves contact with children must undergo a Working with Children Check. Evidence that these checks are current has been provided to the Principal of your school your child attends.

Who do I contact if I wish to discuss the project further?

If you would like to discuss any aspect of this study with a member of the research team, please contact my supervisor using the details provided below. If you wish to speak with an independent person about how the project is being conducted or was conducted, please contact the Curtin University Human Research Ethics Committee (contact details below).

How do my child and I become involved?

Please ensure that you:

Discuss what it means to take part in the project with your child before you both make a decision; and

Take up my invitation to ask any questions you may have about the project.

Once all questions have been answered to your satisfaction, and you and your child are both willing for him/her to become involved, please complete the Consent Form on the following page (your child is also asked to complete the Consent Form attached to his/her letter) and return it to your child's teacher.

This project information letter is for you to keep.

Thank you,

Naomi Yiftachel

naomi.yiftachel@postgrad.curtin.edu.au

Associate Professor Cori Williams

Curtin University

Tel: 9266 7865

Email: c.j.williams@curtin.edu.au

Human Research Ethics Committee

Curtin University

Tel: 9266 2784

Email: hrec@curtin.ed

CONSENT FORM

(For Parent/Caregiver)

- I understand the information about the project.

- I have asked any questions I may have had and I am happy with the answers.

- I understand that it is up to me to decide whether my child and I take part.

- I understand what it means for me to participate in this project and am happy to take part in it.

- I have discussed with my child what it means to participate in this project. He/she has explicitly indicated a willingness to take part, as indicated by his/her completion of the child consent form.

- I give permission to provide the school with my child's assessment results.

- I give permission for results of the Raven's Coloured Matrices test to be given to other professionals if my child needs a full intelligence assessment within two years from the date of testing.

- I give permission to record me and my child's speech.

- I give permission to test my child's hearing including a tympanometry test.

- I understand that we can pull out of the project at any time without affecting the family's relationship with my child's teacher or school.

- I am happy for the project to be presented at a conference and possibly published in a journal. I know that my child's school will not be identified in any way.

- I understand that I can request a summary of the findings after the project is finished.

Name of Child (printed):	**Name of School the child attends:**
I would like to be interviewed in: **English** **Other (please specify):**	**I would like the interview to take place:** **At my home- face to face** **At my child's school** **Over the phone**
Signature of Parent/Caregiver:	**Date: / /**
Phone number:	**E-mail address:**