Can a Short Parent Questionnaire be Helpful for Correctly Identifying Children with and without Specific Language Impairment?

Marianne Paul, M.Sc.(A), M.Sc.

School of Communication Sciences and Disorders

McGill University, Montreal, Canada

August, 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Communication Sciences and Disorders

© Marianne Paul, 2016

Table of Contents

À toute ma belle et grande famille,

En particulier mes petits amours, Aldéric, Xavier et Lili,

Et mon grand amour Yan

Abstract

Language development is influenced by various factors related to intrinsic characteristics of the child, such as specific language impairment (SLI), or to his or her environment, such as a bilingual upbringing. However, it is in the presence of a language impairment that a child should receive language intervention. The purpose of the current thesis is to investigate the usefulness of the Milestones en français du Québec (MilBec: Paul & Elin Thordardottir, 2010), a short parent questionnaire for children between 12 and 71 months adapted from Luinge et al. (2006) to screen for SLI.

In Chapter 1, the adaptation procedures are detailed and a cross-sectional study that led to the creation of the MilBec is presented, followed by a longitudinal study. The results showed that the MilBec is developmentally sensitive with a rapid increase up to 36-42 months, followed by a slowed increase ending with a ceiling effect for children 60 months and older. The MilBec's validity is supported by correlation analyses with other language tests, and high levels of reliability are suggested for test-retest and inter-judge reliability measures.

In Chapter 2, three studies are presented: a preliminary normative study, a diagnostic accuracy study, and a study presenting the analysis of the children's performance on two items viewed as potential clinical markers, namely object clitic and subject use. Object clitics are pronouns used between the subject and the verb functioning as the direct object of the verb (e.g. "Je LE prends" ('I IT take')). The results indicated that the MilBec is developmentally sensitive and has a good diagnostic accuracy, with sensitivity of 100% and specificity of 93% at identifying children with a moderate to severe SLI. The concurrent validity of the MilBec is further supported by high correlations with various language tests. Finally, the analyses showed that when a parent reports that his or her child does not use the subject after the age of 27

months, or the object clitic after the age of 31 months, it is a strong sign of the presence of SLI (a specificity of 100% and 97%, respectively). However, if the child uses these elements, it cannot be used to rule out SLI, given the sensitivity of only 20% for subject use and 53% for object clitic use.

In Chapter 3, the usefulness of the separation of the MilBec's the items into two subscales, depending on whether the items assess a language domain particularly affected, or unaffected, by bilingualism is explored. A pilot study comparing the performance of two groups of TD bilingual children, with either a high or a low exposure to French, showed that bilingual children generally perform similarly to monolingual children on the MilBec. However, a higher proportion of bilingual children with low exposure to French failed the subscale targeting domains more affected by bilingualism. In contrast, monolingual children with SLI failed both subscales.

Together, the findings indicated that MilBec is a promising screening test to identify French-speaking children with a moderate to severe SLI, even for bilingual children. Future studies should be performed to collect norms from a population sample, to further investigate the performance of bilingual children, and to determine its predictive validity.

Keywords: French, specific language impairment (SLI), test validation, normative data, parent-questionnaire, screener, bilingual, language acquisition

Résumé

Le développement du langage est influencé par de nombreux facteurs qui peuvent être intrinsèques à l'enfant, tel un trouble primaire du langage (TPL), ou plutôt lié à son environnement, tel le bilinguisme. Par contre, ce n'est qu'en présence d'un trouble que l'enfant devrait recevoir des services en orthophonie. L'objectif de cette thèse est de documenter l'utilité du *Milestones* en français du Québec (MilBec: Paul & Elin Thordardottir, 2010), un court questionnaire aux parents pour les enfants entre 12 et 71 mois, adapté du questionnaire de Luinge et coll. (2006), pour dépister les TPL.

Au Chapitre 1, la procédure utilisée lors de l'adaptation du questionnaire et une étude transversale ayant mené à la création du MilBec sont présentées, suivies d'une étude longitudinale. Les résultats montrent que le MilBec est sensible au développement, avec une augmentation rapide de la performance jusqu'à l'âge de 36-42 mois suivie d'une période de ralentissement de l'augmentation ou un effet plafond chez les enfants de 60 mois et plus. La validité concurrente du MilBec est documentée par des corrélations avec d'autres tests de langage et un haut niveau de fiabilité test-retest et inter-juge sont obtenus.

Au Chapitre 2 une étude normative préliminaire, une étude vérifiant la précision du dépistage et une étude analysant la performance des enfants à deux items portant sur des marqueurs cliniques potentiels sont présentées, c'est-à-dire l'utilisation d'objet clitiques et l'utilisation du sujet. Les objet clitiques sont des pronoms ayant une fonction d'objet direct et placés entre le sujet et le verbe ("Je LE prends"). Les résultats indiquent que le MilBec est sensible au développement, qu'il a d'excellents niveaux de sensibilité (100 %) et de spécificité (93 %) pour l'identification des enfants ayant un TPL modéré à sévère. La validité concurrente du MilBec est documentée par les corrélations élevées entre le MilBec est d'autres tests de

langage. Finalement, l'analyse de la performance aux deux items indique que si un parent indique que son enfant n'utilise pas de sujets après l'âge de 27 mois, ou d'objets clitiques après 31 mois, c'est un signe important de la présence d'un TPL (avec une spécificité de 100 % et 97 % respectivement). Par contre, si l'enfant utilise ces éléments, cela ne permet pas d'exclure la présence d'un TPL, tel que l'indique les faibles niveaux de sensibilité de 20% pour l'utilisation des sujets et de 53% pour l'utilisation d'objets clitiques.

Au Chapitre 3, l'utilité de séparer les items du MilBec en deux échelles, selon que l'item évalue un domaine du langage particulièrement affecté ou non par le bilinguisme, est explorée. Une étude de cas multiple d'enfants bilingues au DT ayant un haut ou un bas niveau d'exposition au Français a été effectuée. Les résultats montrent que les enfants bilingues performent au MilBec généralement de manière similaire aux enfants unilingues. Par contre, une plus grande proportion d'enfants bilingues ayant une faible exposition au français échoue la sous-échelle qui inclue les éléments particulièrement affectés par le bilinguisme. De leur côté, les enfants unilingues ayant un TPL échouent aux deux échelles.

Le MilBec est un test de dépistage prometteur pour l'identification des enfants ayant un TPL modéré à sévère, même pour les enfants bilingues. Des études permettant de collecter des données normatives au niveau de la population, continuer l'investigation de la performance des enfants bilingues et déterminer la valeur prédictive du MilBec devraient être effectuées dans le futur.

Mots clés: Français, trouble primaire du langage, dysphasie, validation de test, données normatives, questionnaire aux parents, dépistage, bilingue, acquisition du langage

Acknowledgments

I want to thank Elin Thordardottir, PhD, who has accepted to be my supervisor for this doctoral work. I am very grateful for her help and support, and for all the experiences I have been able to live over the years with her as my mentor. I truly appreciate it. I also want to thank the members of my thesis committee, Aparna Nadig, PhD and Ann Sutton, PhD, who provided helpful comments and suggestions at different stages of the process, and the members of my thesis defense committee. I want to thank Dre Julie Lamoureux, statistician, for her help in choosing the appropriate analysis to determine the effect of age in the longitudinal data, and for her confirmation that I had performed and interpreted the results appropriately. You all contributed to making me a better researcher, and helped me learn many things over the years.

I wish to thank all children and their parents who have agreed to participate to my research project, as well as daycare directors who agreed to distribute the project's recruitment letter. It would not have been possible without them all. I also want to extend a special thank all the people who helped me with the recruitment for the different studies presented in this thesis. Notably, for the children with specific language impairment, I want to thank Eve Julie Rioux for her help, as well as all the health professionals who accepted that I come to present the project in one of their therapy sessions at the Institut Raymond-Dewar. For the recruitment of participants with typical development, I also which to thank Florence Paul, Véronique Paul, Aurélie Paul, Marilène Girouard, and Marie-Pier Gingras. For her help with testing materials, I am also thankful to Martine Gendron.

Je tiens à remercier les membres de ma famille et mes amis qui m'ont écouté parler (encore et encore) de mon projet de doctorat. Merci à mes parents, mes beaux-parents et mes sœurs, en particulier Aurélie, pour leur aide avec les enfants. Merci aussi à mes nouveaux

collègues à l'Université du Québec à Trois-Rivières pour vos encouragements et votre soutien.

I am grateful to the funding I have received over the course of my doctoral studies, which includes a Provost's Fellowship and a Graduate Excellence Fellowship from McGill University, stipends from my supervisor Elin, and a student travel award for a SRCLD conference.

Preface and Contribution of Authors

The work presented in this thesis is the work of me, Marianne Paul, conducted under the supervision of Elin Thordardottir, PhD. The first part of the article presented in Chapter 1 describes the adaptation of the Dutch parent-questionnaire by Luinge, Post, Wit, and Goorhuis-Brouwer (2006) that led to the creation of the Milestones en Français du Québec (MilBec: Paul and Elin Thordardottir, 2010). This work was performed during an independent study course I took under the supervision of Elin Thordardottir during my studies in the MSc program in 2008-2009 at McGill University. As this work had not been previously published, but was the basis of my doctoral work, it was included in that first article. The other studies presenting data on the MilBec were all undertaken as part of my doctoral work. All studies presented in this dissertation were designed, the data collected, analyzed and interpreted by me, with guidance offered by Elin Thordardottir at all steps. Members of the committee, Aparna Nadig and Ann Sutton, provided helpful suggestions concerning the studies' design and planned analyses, as well as on the 'almost final' version of this dissertation. I drafted and revised all three articles included in the thesis and Elin Thordardottir, the second author on all manuscripts, provided comments and suggestions on all the earlier versions.

The contribution to knowledge of the studies presented in this thesis concerned the development and validation of the MilBec and the documentation of some aspects of its validity and reliability; its usefulness with the population of bilingual children was also explored. The data, collected in the Province of Quebec, could be used as preliminary norms. From a theoretical point of view, these data could lead to a better understanding of the manifestation of specific language impairment (SLI) in Francophone children between 3- and 5-year-old and the impact of a reduced exposure due to bilingualism on the acquisition of French.

The data presented in this thesis were presented during my doctoral studies in poster format in two international conferences (IASCL, 2011; SRCLD, 2014); two additional posters based on the data collected in this thesis will be presented at the international conference of the École d'orthophonie et d'audiologie de l'Université de Montréal, to be held in November 2016. The data were also presented in one national conference (REPAR, 2013) and one regional conference (Carrefour des connaissances, 2016), as well as in two local oral communications (conférence-midi CRIR, April 28th, 2016; conférence-midi at the Institut Raymond Dewar, June 21st, 2016). At the time of the final submission of this thesis, part of the data presented here are reported in a manuscript submitted to the Canadian Journal of Speech-Language Pathology & Audiology. The submitted manuscripts are not identical to the ones included in this thesis

General Introduction

The present doctoral work has three main goals related to documenting the usefulness of a new language screening tool, the Milestones en Français du Québec (MilBec: Paul and Elin Thordardottir, 2010), a French adaptation of the Dutch parent-questionnaire by Luinge, Post, Wit, and Goorhuis-Brouwer (2006). The first goal is to document the MilBec's developmental sensitivity, concurrent validity and reliability. The second goal is to investigate the diagnostic accuracy of the MilBec as a whole and two of its items specifically, targeting potential clinical markers of specific language impairment (SLI), namely subject use and object clitic use, to identify children with a moderate to severe SLI. The third goal is to explore if the MilBec can be used to rule out SLI in bilingual children, by comparing the answer profiles of two groups of bilingual children who differ in their level of exposure to French with the performance of monolingual children with and without SLI.

The comparison of the language skills of children with or without SLI as assessed with the MilBec can have theoretical importance. Indeed, the way in which a child's language is not acquired as expected when compared to peers informs on the nature of SLI. This comparison has to take into consideration the high variability in language skills in typically developing children, both between and within children across time (e.g. Duff, Reen, Plunkett, & Nation, 2015 for vocabulary). This variability can stem from different elements, one being related to the input received by the child, notably when the child is raised in a bilingual environment. In this section, a review of the challenges related to the early identification of French-speaking monolingual and bilingual children with SLI will be presented. This review will be followed by an overview of the theories on language acquisition and language impairment and their influence on the MilBec's item selection.

Early Identification of Monolingual and Bilingual Children with SLI

The variability of the manifestations of SLI in conjunction with the variability in normal language development is probably a contributing factor to the fact that a majority of children with SLI are not identified before they begin formal schooling. In this section, the difficulties related to the early identification of French-speaking children with SLI and the role of input in language development will be briefly overviewed.

Early Identification of French-Speaking Children with SLI

It has been found in an epidemiology study of SLI in English-speaking children attending kindergarten that the parents of 71% of the children whose language scores indicated a language impairment had never been informed that their child had language difficulties (Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997). Many years later, a similar finding was reported for French-speaking children in the preliminary prevalence study conducted in Quebec, where 77% of the 39 children who failed the assessment protocol had never been previously identified as having SLI (Elin Thordardottir, 2010). The fact that such a high proportion of children with SLI were not identified early might be related to the fact that reliable early indicators of language impairment are not currently known. Traditionally, children identified early as having a high risk of language impairment are those with a slowed acquisition in vocabulary and early syntax at age 2, often referred to as "late-talkers" or "late-bloomers" if they catch-up with their peers. However, recent findings in English indicate that these children often reach normal limits between 5 and 7 years of age (Rescorla, 2011), and that the majority of the children still demonstrating language difficulties at age 8 were not having a slowed early vocabulary and grammatical development at age 2 (Poll & Miller, 2013).

In a literature review by Desmarais, Sylvestre, Meyer, Bairati, and Rouleau (2008), it

was found that the group of late-talking children was a very heterogeneous group; it was also highlighted that in addition to a delayed expressive vocabulary, other skills should be considered when describing this group. This was further supported by the results of Desmarais, Sylvestre, Meyer, Bairati, and Rouleau (2010), where the performance of 68 French-speaking late-talkers between 18 and 35 months was investigated. To be included in the study, the children had to have received a score under the 10th percentile on the Inventaire MacArthur Bates du Développement de la Communication (IMBDC: Trudeau, Frank, & Poulin-Dubois, 1999; Frank, Poulin-Dubois, & Trudeau, 1997), a French adaptation of the MacArthur Communicative Development Inventories (MCDI: Fenson et al. 1993). A cluster analysis identified three variables in addition to expressive vocabulary, namely the "comprehension", "expression and engagement in communication" and "cognitive development", used to categorize these late-talkers into three distinct subgroups. Further longitudinal analysis is required to examine whether these additional factors have predictive value.

In French, what should be considered a strong sign of SLI is still unclear, although there is an accumulation of findings supporting a general slowing in language acquisition in various language domains (Elin Thordardottir, 2016). Many studies that report on specific clinical markers focused on older children. Notably, difficulties with the use of object clitics (e.g. "Je LA veux" (I IT want) for "Je veux LA POMME" (I want THE APPLE)) (Jakubowicz, Nash, Rigaut, & Gérard, 1998; Paradis, Crago, & Genesee, 2009; Grüter, 2005) is one marker often used by speech-language pathologists. However, the evidence supporting its usefulness in identifying children with SLI is mixed (e.g. Elin Thordardottir & Namazi, 2007; Hamann et al., 2003; Hamann, 2003).

In Quebec, an additional challenge in the early identification of Francophone children is

the lack of validated tools to assess French (Gaul Bouchard, Fitzpatrick, & Olds, 2009). This is particularly true for tests assessing the language skills of children under the age of 4 years, since only two are reported as having a documented validity for French-speaking children in Quebec (Monetta et al., n.d.). These two tests are the IMBDC presented earlier and the Échelle de Vocabulaire en Image Peabody (EVIP: Dunn, Thériault-Whalen, & Dunn, 1993), a French adaptation of the Peabody Picture Vocabulary Test – Revised (PPVT-R: Dunn & Dunn, 1981). These two tests target early vocabulary or grammatical development. While they are very important tests to document the children's skills on these aspects, assessing these skills exclusively is not sufficient to identify many of the children with SLI, as previously mentioned. There is thus a need for the development of new assessment tools developed for Francophone children living in Quebec, as well as a need to increase the knowledge base on the language skills of children with SLI under the age of 5 years.

Importance of the Input

There are different environmental factors that can have an important impact on the child's language development. Notably, the child's family's socio-economic status and the mother's level of education are two factors particularly well-documented in the literature (for a review, see Prathanee, Thinkhamrop, & Dechongkit, 2007). Another important environmental factor is whether one or more languages are spoken to the child. One consequence of exposing a child to two or more languages is that it divides the child's waking time between them. This division leads to a reduced language input in each language compared to the input received by monolingual children. This reduced exposure may lead typically developing (TD) bilingual children to have lower language skills in each language than their monolingual peers during acquisition (for a review, see Elin Thordardottir, 2011; 2014; 2015a). For example, simultaneous

bilingual preschoolers were found to score significantly lower than monolingual peers if they had spent less than 40% of their time exposed to that language (Elin Thordardottir, 2011) and sequential bilinguals were found to have lower skills in their second language even after six years of schooling in that language, although they had an equally good performance in school, when compared to their monolingual peers (Hemsley, Holm, & Dodd, 2006). Nonetheless, a notable advantage of bilingualism is that the child can communicate in more than one languages, whereas their monolingual peers cannot.

A particularly complicating factor for the correct identification of TD bilingual children is that their performance in each language is sometimes, but not always reduced. Indeed, the performance of bilingual children on grammatical measures can be similar to that of their monolingual peers with a language impairment in terms of the number and type of errors they exhibit (for a review, see Kohnert & Medina, 2009). For example, in a study on French-English bilinguals in Montreal it has been found that only the children with low levels of exposure to English exhibited a stage of Extended Optional Infinitive at age 5 years, a strong clinical marker of SLI in English in monolingual children. The same children demonstrated well-developed morphosyntactic skills in French, in which they received a higher level of exposure (Elin Thordardottir, 2015a).

Language acquisition is a complex process which is further complicated when the child is exposed to more than one language. This partly explains why the early identification of SLI is such a challenging task. The importance placed on input and the underlying difficulties that children are believed to have when they have language impairment varies across the theories of language development and language impairment, as will be briefly overviewed in the next section.

Theoretical Considerations in Test Development

There are many theories about the mechanisms underlying language acquisition or language impairment, although there is no consensus on which one is the most appropriate (Kennison, 2014; Kuder, 2013). The theoretical framework used during a test creation has an impact on the selection of items. For the parent questionnaire presented in this thesis, the theoretical framework can be qualified as eclectic: because each theory has its own strengths and weaknesses and focuses on different aspects of language development, they are all considered as potentially useful in the selection of items.

Theories of Language Acquisition

One way that theories of language acquisition can be categorized is based on the role assigned to the child's experience (for a review of different theories, see Kennison, 2014; Kuder, 2013). The first camp assigns only a minimal role to experience, and contains different theories in which language is conceptualized as an innate ability. Most of these theories are based on Noam Chomsky's Universal Grammar, in which it is posited that children possess grammatical knowledge from birth. Children would only have to set the parameters of the specific language that they are being exposed to. Under these theories, there is thus not much importance placed on the quantity of exposure to the language in the language acquisition process (for a review, see Leonard, 2014; Tomasello, 2000).

In contrast, theories from the second category place a high importance on the children's experiences (for a review, see Kennison, 2014; Kuder, 2013) - this has been particularly highlighted, for example, in the work of Elena Lieven. Some theories present language learning as being similar to any other learning; other approaches focus more on the importance of the assistance provided by the adults during learning. Tomasello (2000) pointed out that the high

importance placed on exposure received by the child allowed to explain the language ability variations either from individual differences in underlying skills or from the child's environment.

Theories of Language Impairment

When a child's language is particularly slow to develop compared to peers, the child may be described as having language impairment. Specific language impairment (SLI) is a disorder characterized by language abilities lower than expected given the child's age that cannot be attributed to other conditions such as autism, intellectual disability, or deafness. In Quebec, the official term used by the Ordre des orthophonistes et audiologistes du Québec (OOAQ) to describe SLI is Trouble primaire du langage/dysphasie ('primary language impairment' (PLI)), which has been redefined in 2004. Different elements within this definition highlight the complexity of correctly identifying children with SLI. Firstly, there is a high level of heterogeneity of manifestations, both within and between individuals. Indeed, to be considered as having a language impairment a person can have difficulties in the expressive modality, or in the expressive and receptive modalities, in any combination of two or more language domains (phonology, morphology, syntax, semantic, and pragmatic). Secondly, the manifestations of PLI within the same person can change during his/her lifetime (OOAQ, 2004)[1].

Theories of language impairment have taken two main views on how they classify children with SLI: the proponents of the delay view consider them as a subgroup of the population and the proponents of the deviant view them as forming a group of their own. Each of

[1] The term SLI rather than PLI will be used throughout the thesis for consistency. Both terms refer to the same impairment, even if in the OOAQ definition of PLI the notion of persistence of the disorder and the handicapping situation it causes are required, while they are not in the SLI definition. However, in SLI a more severe impairment is also associated with a high risk of a persistent impairment.

these camps comprises accounts that have been elaborated and nuanced over the years, but the basic distinction lies in this delay vs deviance opposition. Authors supporting the delay view propose that the element differentiating children with or without language impairment is their localization within the language skills continuum. The authors of the deviant view suggest that the difference is qualitative, with the presence of deviance from typical language development (for a discussion on the different views, see Leonard, 2014; Rice, 2000, 2004).

In support of the delay view of language impairment theories, Rescorla (2011) concluded in her extensive review that language skills of children with SLI differed only quantitatively, and not qualitatively, from the skills of children with TD. Furthermore, she concluded that when they were older, the late-talkers had language skills falling in-between the skills of children with or without SLI. Language would thus develop in a similar manner for all children: some children would be more skilled (i.e. children with TD), others would have lower skills, but this would be more easily apparent only at a young age (i.e. late-talkers), and finally some children would have apparent difficulties at all ages (i.e. children with SLI). Under this view, children with SLI would thus be delayed in their language acquisition when compared to children with TD, being slower at acquiring it. A possible explanation for the children's low performance, or slower learning, could be a limitation in processing skills (e.g. Elin Thordardottir, 2011). The proponents of a limited-processing capacity theory vary on whether they posit a general deficit or a deficit specific to a particular domain (for a review, see Leonard, 2014). Following this delay viewpoint, the items should focus on behaviors of normal development that characterize a younger TD range than is targeted in the test. Children with SLI, because of their less developed skills, would not yet be at the stage where these skills have been acquired by TD children.

In contrast, the proponents of the deviant language view usually advance that children

with language impairment have difficulties with innate linguistic knowledge. The most common linguistic framework used to explain these difficulties is the Universal Grammar proposed by Chomsky, which was briefly mentioned earlier. The manifestation of this missing knowledge would be specific error patterns observed in children with SLI. These error patterns would be apparent even when children with SLI are compared to children with comparable overall language skills, usually younger children who do not have language impairment (e.g. Jakubowicz et al., 1998; Rice & Wexler, 1996; Rice, 2004). These specific linguistic elements more prone to a particularly delayed acquisition are referred to as clinical markers. These elements would be extremely difficult to acquire by children with SLI, because these children would lack the innate knowledge required for their acquisition. Many theories under this deviant view have been put forward (for a review of these theories, see Leonard, 2014), the most widely known being the extended optional infinitive proposed by Rice & Wexler (1996). Under such a view of language impairment, the children are assessed with great attention on the clinical markers of SLI; a screener test should thus include items targeting these markers.

The identification of SLI in monolingual and bilingual children touches on two important theoretical issues: the nature of SLI and the role of input. Even though many elements related to these issues render the correct identification of children with SLI a complex thing to do, it is a task that must be performed with the highest accuracy possible. Indeed, the allocation of specialized resources should be delivered to children with a language impairment, to limit the handicap associated with this disorder, and not to children who demonstrate normal language variations. To do so, children have to be correctly identified as TD or as having SLI, and this early identification might be increased if a validated and normed screening tool is available.

Statement of Purpose

The present doctoral work investigated the psychometric properties of a new parent questionnaire to screen for language impairment in French-speaking children, namely the Milestones en Français du Québec (MilBec: Paul & Elin Thordardottir, 2010). Its applicability to both monolingual and bilingual children will also be investigated. In Chapter 1, the adaptation procedures of the MilBec will be presented, as well as data from a longitudinal study documenting aspects of its validity and reliability. In Chapter 2, a normative study and a diagnostic accuracy study will be presented. These studies were performed to verify whether the MilBec can adequately screen monolingual children for language impairment, while also providing additional information regarding the test's validity and providing preliminary normative data. This chapter also verifies, using an item analysis, the usefulness of two clinical markers proposed for school-age children, namely the use of subjects and the use of object clitics, at identifying children with SLI in children between 3 and 5 years. Finally, Chapter 3 will explore the usefulness of extracting language profiles by creating two subscales, containing items that are either expected to be more affected or less affected by bilingualism, to better identify SLI in bilingual children. The links between the different studies testing the MilBec are illustrated in Figure 1.

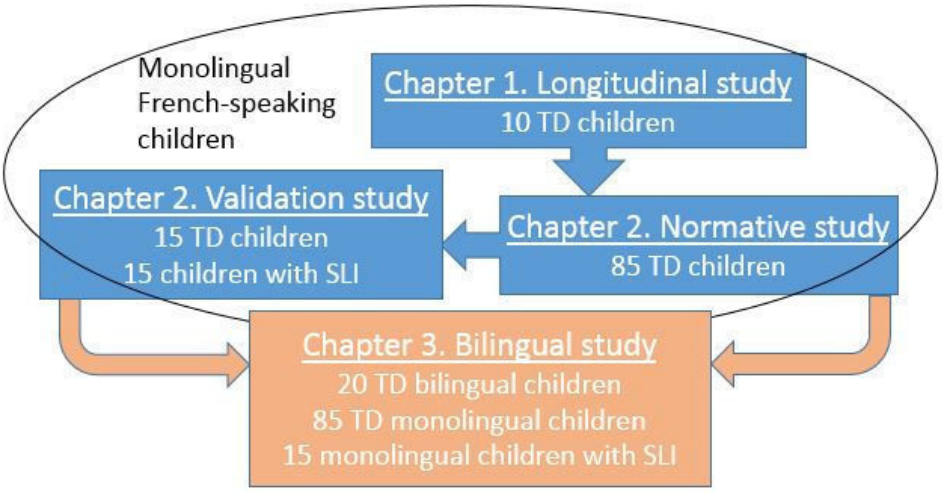


Figure 1: The links between the different studies.

The thesis findings will have direct clinical applications, as they provide validation and preliminary norms for a new screening test for language impairment. They will document the ability of two potential markers, namely the object clitic and subject use, to correctly identify children with SLI, which will contribute to the research body on the nature of language impairment. Finally, the comparison of the profiles of TD bilingual children on the two subscales of the MilBec to the profile of monolingual children with and without SLI will contribute to the research body on the role of input in language acquisition, both typical and impaired. Hopefully, the availability of this knowledge will contribute to the improvement of the early identification of Francophone children with SLI, for both monolingual and bilingual children. The MilBec can also have applications in research, as it is a validated, but still quickly performed, screener of SLI.

Chapter 1: The Development and Preliminary Validation of a French Parent Questionnaire to Assess Language Skills

Marianne J. Paul and Elin T. Thordardottir

McGill University

and

Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS)

Abstract

Purpose: This study presents the adaptation procedures of the Milestones en Français du Québec (MilBec: Paul & Elin Thordardottir, 2010), a French adaptation of the Dutch parent questionnaire by Luinge et al. (2006) aimed at identifying language difficulties in preschool children, and preliminary data on its psychometric properties. **Method**: The article reports: 1) A cross-sectional pilot study with 26 participants aged 1, 3, and 5 years. This study investigates the developmental sensitivity of the preliminary version of the parent questionnaire. 2) A 2-year longitudinal study employing a revised version of the questionnaire with 10 participants between 12 and 66 months at the study onset. This study investigates the developmental sensitivity and aspects of the MilBec's validity and reliability, including comparison with other language measures. **Results**: Study 1 showed that scores on the preliminary version increase with age, but also pointed to some required modifications. Study 2 revealed a statistically significant effect of growth in scores for the MilBec, although a ceiling effect occurred around 42 months. Validity is supported by the concurrent increase of MilBec scores and scores of other measures of receptive language. Test-retest and inter-judge reliability are both above .90. **Conclusions**: The MilBec is a promising clinical tool for assessing overall language skills in preschool children. Additional studies investigating its diagnostic accuracy, validity and reliability are warranted.

Keywords: language development, screener, parent-report, French, preschoolers, validation

Introduction

The term language impairment can be used with children having a wide range of difficulties. The language difficulties of these children are generally classified as a specific language impairment (SLI), when language is the only impairment that the child has, or as a language impairment secondary to another impairment, such as deafness, intellectual disabilities, autism spectrum disorder or other various syndromes. Many children require the help of speech-language pathologists, since the prevalence rate of specific language impairment (SLI) is 7.4% for kindergarten children in the USA (Tomblin et al., 1997). It was also reported to be as high as 9 to 14% in a preliminary prevalence study conducted with 5-year-old francophone children in Quebec (Elin Thordardottir, 2010; Elin Thordardottir et al., 2003-2008).

For children with SLI, providing early intervention is important in order to minimize the consequences of having low language skills. Indeed, it has been found in a systematic review by Nelson, Nygren, Walker, and Panoscha (2006) that for children aged 5 years or younger, intervention significantly improves speech and language outcomes, including articulation, phonology, expressive and receptive language, lexical acquisition and syntax, when compared to control groups. Improvement in other aspects influenced by language abilities, such as socialization skills, self-esteem, and play themes, are also reported in some studies. Finally, providing earlier intervention to children with SLI has greater long-term benefits. For example, a study by Vlassopoulos et al. (2014) compared the outcomes at age 13 of children with SLI who received, or did not receive, therapy in the preschool years. It was found that even though children who had received therapy had, as a group, a more severe impairment than those who did not, they had similar school success at age 13. Furthermore, it is only when a therapy received before the age of 5, and not when the child is older, that the child's behaviour and concentration

at age 13 were found to be improved.

To maximize the long-term gains of therapy, it is thus better to start intervention before the child reaches 5 years of age, which requires an early identification of the problem, although it should be stressed that intervention during the school years is also beneficial. However, many children who will later present with long-term SLI are currently not identified before they enter school. For example, in the prevalence study at age 5 conducted by Tomblin et al. (1997), it is reported that only 29% of the parents whose children were identified has having SLI already knew that their child had language difficulties. Similar results were obtained in Greenslade, Plante, & Vance (2009), in which children between 48 and 68 months were identified as having SLI based on a combination of clinical judgment and an independent classification based on formal testing only. Using that protocol, 31 of the 42 children in the SLI group were identified for the first time as having SLI. In the preliminary prevalence study conducted in Quebec, 30 of the 39 children who failed the assessment protocol had not been previously identified as having SLI (Elin Thordardottir, 2010).

To increase early identification, appropriate tests should be available to assess young children's language skills. However, there is currently a limited number of tests to assess French-speaking children under the age of 5 in Quebec (Canada) (Garcia, Paradis, Sénécal,& Laroche, 2006; Gaul Bouchard, Fitzpatrick, & Olds, 2009; Monetta et al., n.d). Some screening tests developed in Europe are available in French, but proper validation is required to determine if the content and available norms are appropriate for French Quebeckers. Indeed, Frisk, Montgomery, Boychyn, Young, van Ryn, McLachlan, and Neufeld (2009) found that in some cases the use of a screening test from another country, even if the same language is used, required an adjustment of the cut-off score to maintain an appropriate identification level. The authors suggested that the

differences were due to demographic and educational factors between the two countries.

Another complication in the early identification of children with SLI is that most of the available validated tests in French focus on vocabulary. However, recent findings indicate that using low vocabulary knowledge as an early indicator for SLI is not an ideal choice. Indeed, in her extensive review, Rescorla (2011) reported that numerous research studies have concluded that the predictive validity of vocabulary measures is rather low, since the majority of late talkers, those children who are "slow to talk" (p. 142), improve their language skills to the point that they are within the normal range by the age of 5 to 7 years. Furthermore, it has been shown for English that even though early vocabulary development is strongly related to later syntactic development (Bates, Bretherton and Snyder, 1988), vocabulary tests do not adequately identify 4- and 5-year-old children with or without SLI, with many children with SLI scoring within normal limits and many children with typical language development scoring below expectations (e.g. Gray, Plante, Vance, & Henrichsen, 1999; Spaulding, Hosmer, & Schechtman, 2013). It has been proposed that the low ability of receptive vocabulary tests to identify children with SLI is because such tests do not assess the depth of knowledge of the words known by the child (Spaulding et al., 2013). In light of these results, it seemed preferable to have a screening test for identifying children with language difficulties targeting elements from various language domains rather than a test including exclusively vocabulary.

The purpose of this paper is to address the need for a French screening tool for children between 12 and 71 months validated in Quebec. Specifically, the selection and adaptation procedures of a parent questionnaire to screen language impairment will be described in the next section. Then, Study 1 will present a cross-sectional pilot study performed to ensure that parents easily understand the items and the examples that are presented, followed by the description of

the modifications performed to obtain the Milestones en français du Québec (MilBec: Paul & Elin Thordardottir, 2010). Finally, Study 2 presents a longitudinal multiple-cases study investigating different aspects of the MilBec's psychometric properties: its developmental sensitivity, its concurrent validity, and different aspects of its reliability.

Procedures

Development of the MilBec

The first step leading to the creation of a new screening tool in French was to determine whether an adaptation of an already existing screening test was possible or if a new test should be created altogether. To be considered a good candidate for adaptation, a test should have the following characteristics. First, the test's content should target several language domains, not only vocabulary knowledge and early syntax. Second, the administration and scoring should not require specialized training. Third, the administration and scoring time should be very brief, ideally less than 10 minutes, so as to be adequate for large-scale screening. Finally, the age range targeted by the screener should cover at least the period between 24 and 48 months, because it is for this age range that the need for a screening tool is the highest. The last two characteristics were considered important to reduce the cost and difficulty related to managing multiple tests, while facilitating the monitoring of each child's rate of language development. Parent questionnaires were considered particularly good candidates, as they usually entail no specialized training and a reduced administration time. Furthermore, several studies support the validity of such measures (e.g. Boudreault, Cabirol, Trudeau, Poulin-Dubois, & Sutton, 2007; Klee, Carson, Garvin, Hall, Kent, & Reece, 1999).

A review of the available European tests was performed to compare the tests' characteristics to the list of criteria. Several screening tools for young children were found, but

none fulfilled all of the targeted characteristics (for more details on the tests, see Vallée & Dellatolas, 2005). A European French adaptation of the MacArthur Communicative Development Inventories (MCDI: Fenson et al., 1993), which targets vocabulary and early syntax for children under the age of 30 months, does not fulfill characteristics 1 and 4; furthermore, a Quebec equivalent is already available and well documented (Trudeau et al., 1999). The Dépistage et Prévention Langage à 3 ans (DPL-3 : Coquet & Maetz, 1997), the Questionnaire Langage et Comportement 3 ans ½ (QLC_3,5 : Chevrie-Muller, Goujard, & Goujard, 1994) and the Épreuve de Repérage de Troubles du Langage lors du bilan médical de l'enfant de 4 ans (ERTL-4 : Roy & Maeder, 1996) have a very narrow target age range, which varies between 3 and 9 months around the target age, thus they do not fulfill the fourth criteria. Additionally, the ERTL-4 requires specialized training (Maeder & Roy, 2000). The Batterie Rapide d'Évaluation des fonctions cognitives (BREV : Billard et al., 2000) targets children between 4 and 9 years of age, the Bilan de Santé Évaluation du Développement pour la Scolarité (BSEDS : Zorman & Jacquier-Roux, 2002) is intended for children 5 or 6 years of age, and the Protocole d'Évaluation Rapide (PER2000 : Ferrand, 2000) targets children between 3 years 6 months and 5 years 6 months. Thus, these three tests do not meet the fourth criteria, as they target an older age range.

As none of the European French screening tests fulfilled the specified characteristics, the second step consisted of looking for screening tests from other languages that could be found in journal articles (and for which a translation of the items are available in French or English). A fifth characteristic was added to the list for these potential candidates: the items should target general language milestones. Indeed, since the age of acquisition of global language milestones (e.g. age at which babbling starts, period when two-word combinations emerge) is quite stable

cross-linguistically for young children (Slobin, 1969), an adaptation of a test targeting principally such milestones was considered to have a high likelihood of being adequate in the new language.

The Dutch parent questionnaire from Luinge, Wit, Post and Goorhuis-Brouwer (2006), for which an English translation of the items was provided, possessed all the required characteristics and the search for a candidate for an adaptation was ended. This screener is filled out by the parent, it has a short administration and scoring time, and is aimed at children between 12 and 71 months. It originally contained 26 yes/no items, asking if the child says/comprehends/uses certain linguistics elements. It includes items from both the expressive and receptive modalities, by targeting general language milestones related to vocabulary, syntax, narrative skills, and phonological development. The items were selected based on a literature review of language milestones and of various screening tools available in English. Some items targeted skills that can be hypothesized to be relevant for any languages, such as understanding two-word combinations, asking questions, while others might not be relevant to all languages, such as the use of adequate word order and correct use of irregular plurals or past tenses. The items were selected under a "unitary dimension" (p. 924) view of language, under which if a child has a difficulty in one language domain, it is expected that he has or will have a difficulty in other language domains as well.

The authors of this Dutch questionnaire tested its validity using a cross-sectional study of 527 Dutch-speaking children between 12 and 72 months from the Netherlands (Luinge et al., 2006). The authors performed an item analysis to identify the most adequate items for each of the five age groups. Only the items judged to have appropriate scaling properties were kept in the final version, which contained 14 items. To be judged adequate, the items had to fulfill the assumptions of the Mokken model used, which is a type of item response theory. The authors

conclude that the screener could be used to assess the development of language skills. This was further supported by a second study with 98 participants between 33 and 72 months with and without language difficulties, which yield a sensitivity of 83% and a specificity of 94% (Luinge, 2005).

Adaptation Procedures

Once selected, the Dutch parent questionnaire had to be adapted into French. This adaptation procedure involved four steps: 1) A translation and analysis of the original items, to ensure their suitability in French, was performed. 2) A literature review was undertaken to determine the necessity to use additional items targeting characteristics of French not targeted in the original items. 3) When required, a review of literature to select French examples representative of spontaneous language of francophone children was performed. 4) Native speakers of Quebec French not specialized in language development performed a review of the final items. Each of these steps is presented in more detail below.

The first step was the translation into French of the English translations of the 26 items. Indeed, given that the scaling properties of the items might not be the same due to cross-linguistic differences, the totality of the original items were kept for the adaptation. A direct translation was not always favored, because in some instances more casual vocabulary was considered preferable. For example, the direct translation of 'speech' is "parole", but this term is rarely used with this meaning by nonprofessionals. It was thus translated with "ce que dit votre enfant" ('what your child says'). Once translated, the items were analyzed to make sure that underlying concept was equivalent in both languages, which was the case for all 26 items. No items were thus removed or modified based on this analysis.

The second step consisted of a review of the literature on language development in

French to determine whether some important aspects of French, or some elements shown to be particularly difficult for francophone children with SLI, should be added. Three additional items more specific to French were added, two of them related to the basic syntactic structure of French, which has a subject-verb-complement structure. The first one concerned the use of object clitics, although little has been published concerning its acquisition for younger French-speaking children. Indeed, some authors hypothesize that object clitics are particularly difficult for school-aged children with SLI (Jakubowicz et al., 1998; Hamann et al., 2003). The object clitic is a pronoun placed between the subject and the verb, used to refer to the direct object complement. The simpler structure places the complement, preceded with a determiner, directly after the verb. For example, a typical sentence would be "Je veux la pomme" ('I want the apple'), which corresponds to the more complex syntax with an object clitic "Je la veux" ('I it want').

The second item concerned the use of the subject, which is obligatory in most contexts in French. This item was selected because some authors concluded that school-age Francophone children with language delays performed less well than age-matched peers in this aspect (Jakubowicz et al., 1998), although their performance was not different than that of younger children (Hamann et al., 2003). Furthermore, Elin Thordardottir and Namazi (2007) found that children with SLI between 3 years and 4 and a half years of age omitted the first singular pronoun "je" ('I'), more often than age-matched peers. Sentences without subjects are related to an immature sentence construction, as they are occurring when the child uses infinite verbs (Elin Thordardottir, 2005; Hamann et al., 2003). Such sentences are indicative of a root infinitive stage, which precedes the period when children start to use subjects and start to conjugate (Elin Thordardottir, 2005). To determine sentence maturity by analyzing sentence construction, such

as the inclusion or not of a subject, is a clinical tradition in Quebec. Asking parents if their child uses sentence with a subject is a way to estimate the level of maturity of sentence constructions without having to analyze them.

Finally, the third item targeted the child's ability to make gender agreement between a noun and its modifiers. In French, the noun's gender is usually marked on the accompanying determiner and adjective(s), if present. The gender of the noun sometimes concurs with the referent's biological gender, as in "La fille" ('the$_{feminin}$ girl$_{feminin}$'), but most often the referent does not have a biological gender, as in "Un crayon" ('A$_{masculin}$ pencil$_{masculin}$'). Although gender agreement is an element acquired by TD children as early as 30-36 months for the "un/une" ('a') contrast (Rondal, 2001), it was found that children with SLI as old as between 6;11 and 11;3 still made gender mistakes on the determiner or omitted it (Roulet, 2007).

Almost half of the original items of the screener had examples from child speech to help the parents understand the items. The third step consisted in the selection of the examples that would be used in the adapted version, which would have to correspond to real utterances of young native speakers of French. A review of the literature was performed to select examples representative of the typical way a young French-speaking child would talk. Various sources were consulted, notably Bassano (2000), Parisse and Le Normand (2000), Hickmann (1997) and Rondal (2001). The adapted version of the parent questionnaire can be found in Annex 1.

The final step was undertaken in order to ensure that all 29 items were easily understandable to the general population. The items were reviewed by two adult native French-speakers who did not have experience in linguistics or in speech therapy. Following the analyses of their comments, any required modifications (e.g. correction of typing mistakes, reformulation of some sentences, addition of examples) were performed.

Study 1: Preliminary Version of the Parent-Questionnaire

The goals of Study 1 were to determine 1) whether the preliminary version was easily understood by parents, whether the wording of the items was adequate, whether the examples chosen were helpful to parents, and 2) whether the items varied in difficulty, and 3) to obtain preliminary information on the questionnaire's developmental sensitivity. The study used a cross-sectional sample of children in three age groups.

Method

Participants

The parents of 26 monolingual French-speakers participated in the study: nine children between 12 and 23 months (1-year-old group), eight children between 35 and 45 months (3-year-old group), and nine children between 60 and 69 months (5-year-old group). All children had normal development (no diagnosis or parental concerns about the child's development or hearing). Maternal education served as a measure of socio-economic status (SES): the mothers of 24 participants had attended university, the mothers of the other 2 children (one in the 1-year-old group, 1 in the 5-year-old group) had attended CEGEP. In the education system in Quebec, CEGEPs are post-secondary institutions providing a two-year pre-university program, or a 3-year professional program. The first year of CEGEP is equivalent to grade 12 in other Canadian provinces, and the second year is considered to be equivalent to the first year of university (freshman year).

Procedures

The project was approved by the Institutional Review Board of the Faculty of Medicine of McGill University. Parents were then contacted through daycares by a letter of invitation to participate in a study on French development. Parents who signed the consent form were asked

to complete a questionnaire of general information on the child's development, and the adapted French screening tool with an added section for comments.

A score of 1 was assigned for each item on the adapted French screener to which the parent responded 'yes' and a score of 0 for items to which parents responded 'no' or did not answer. The raw score was calculated out of maximum of 29. On two of the items (item 25 and 27) some parents provided written comments on the questionnaire's margin that made their answer both a 'yes' and a 'no' answer. In these cases, a partial credit was given to the child, who received a score of 0.5 for that item; the three children who received a partial credit had a total score of 28.5.

Analyses and Results

Developmental sensitivity. The distribution of raw scores is shown in Figure 1. For the younger groups, the scores increased systematically with age. However, a ceiling effect occurred for the 5-year-old group, with almost all children having a raw score of 28 or 29. Even some children from the 3-year-old group had reached the test's ceiling.

[Insert Figure 1 about here]

Item difficulty. In order to verify if the items varied in difficulty, the percentage of children having a received a score of 1 was calculated for each group, for each item. The items were reordered to have a decreasing percentage of children receiving a score of 1 across groups, as presented in Figure 2. The easiest items had 100% of the children receiving a score of 1 for both the younger and older groups. The hardest items had fewer than 100% of the oldest group receiving a score of 1, an even lower percentage of children in the middle group and none of children in the youngest group.

[Insert Figure 2 about here]

Parents' comments. Using the option provided at the end of the questionnaire, 17 parents indicated that it took less than 5 minutes to fill out the questionnaire; eight parents indicated that it took 5-10 minutes, and one parent took more than 15 minutes. Three parents indicated that the questionnaire was easy to fill out, and/or not long to complete. Other comments were made regarding the formulation of some items, as discussed in the next section.

Revision of the Questionnaire

Changes were made to the items on the questionnaire based on parents' comments. Notably, a few parents requested a clarification on the wording "généralement" ('usually') used on some items; it was thus changed to "plus de 75% du temps" ('more than 75% of the time'). The item "Est-ce que votre enfant parle comme un adulte, en ce qui a trait à la complexité des phrases?" ('Does your child talk like an adult, in term of sentence complexity'), was annotated relatively frequently, with many parents making a comparison to peers rather than to adults. The item was therefore changed to an overall evaluation of the child's language skill, compared to peers: "Est-ce que vous considérez que votre enfant a un langage suffisamment développé, en comparaison aux autres enfants de son âge?" ('Do you consider that your child's language skills are sufficiently developed compared to other children the same age'). Regarding the item on intelligibility, one parent also pointed out that she understood her child's speech, but that she was almost the only one who did. Because this distinction was judged pertinent, an additional item was inserted about the need for a parent to act as interpreter for his or her child. Finally, some annotations pointed to the need of adding examples for some items that did already have them.

The cross-sectional data showed that the questionnaire is developmentally sensitive, with older children systematically obtaining higher total scores than younger children. The items were reordered to have those receiving a score of 1 by children across all age group to appear before

the items receiving a score of 1 only by older children. This reordering of the items highlighted their variety in terms of difficulty. However, a ceiling effect was seen for the 5-year-old group, and some children from the 3-year-old group also scored near or at the maximum score. It was thus judged necessary to select additional items targeting the older age range.

The search for items focused on relatively later acquired skills, including narration and early reading skills. Eight new items were created based on a review article by Sprenger-Charolles and Serniclaes (2003) on reading and writing acquisition in various languages including French, and the article of Justice, Bowles, and Skibbe (2006) on print knowledge in English. The new items targeted narrative skills, metalinguistic skills (including, notably, phonological awareness) and pre-literacy knowledge. For the later source, it was considered highly probable that the relationships between these general areas would also apply to French, since both languages have a similar alphabetic writing system. Another new item targeted an element specific to French, namely contracted articles (e.g. “du” instead of “de le”). Although no studies have investigated their use by children with SLI, the contracted article is reported to be acquired around the age of 4 (Rondal, 2001), it is thus potentially useful to prevent some 3-year-old children from reaching ceiling. Finally, as mentioned earlier, one additional item targeting intelligibility was added, the formulation of some items was changed based on the analysis of the comments provided by the parents, and examples were provided for some items that did not already have them. In total, 10 items were added to the questionnaire and many of the original items were slightly modified.

Finally, it was observed that parents signaled some degree of uncertainty about their answers by writing additional information beside their answer or by expressing it orally to the first author. It was thus decided to replace the dichotomous yes-no answer choices by a four-

choice answer: “oui” (‘yes’) and “il me semble” (‘I believe so’), both scored as 1 point, and “je ne crois pas” (‘I don’t think so’) and “non” (‘no’) scored as 0 points. It was hypothesized that this would reduce the number of parents who indicate their level of (un)certainty on the questionnaire, by providing them a means to indicate it in their answer.

Following these modifications, the final version of the questionnaire contained 39 items: one item asking if the parent thinks their child has a language sufficiently developed for his or her age, 8 items on expressive vocabulary, 2 items on receptive vocabulary, 4 items on expressive vocabulary/syntax, 3 items on receptive vocabulary/syntax, 4 items on expressive syntax, 4 items on narrative abilities, 2 items on language use/communication, 5 items on phonology/articulation, and 3 items on meta-linguistic knowledge. The items were organized in order of increasing difficulty, following the reordering based on difficulty levels (see Figure 2) for the original items and based on the reported age of acquisition in the literature review for the new items.

The title of the final revised French version of the questionnaire, Milestones en français du Québec (MilBec), makes reference to the original paper presenting the Dutch questionnaire that was adapted into French entitled ‘The ordering of milestones in language development for children 1 to 6 years of age’. The MilBec can be found in Annex 2.

Study 2: A Longitudinal Study Using the MilBec

The aim of this study was to examine the MilBec’s developmental sensitivity, concurrent validity, as well as some aspects of reliability of the MilBec using a longitudinal design involving multiple cases.

A longitudinal paradigm was selected to investigate the developmental sensitivity within the same individual. Indeed, in addition to documenting that older children as a group obtain

higher scores than younger children, it is also important to determine that the scores of individual children increase with age. It is possible that at one point (or more) in development some plateau effect occurs, which could be masked in a cross-sectional study due to the variability across subjects. If the questionnaire is used to monitor a child's language development, the presence of such plateaus must be known. Indeed, this could help avoid the incorrect conclusion of the presence of difficulties because of the presence of a plateau, when in fact the child's trajectory mirrors what is seen in TD children.

Another element that can be investigated using repeated testing of the same individual is test-retest reliability and stability of scores (i.e. that once an item receives a score of 1, it still receives it at the next testing session). Whether the parental judgment on the items changed once the parent knows the item or not should be investigated; it is possible that parents become more aware of the child's mistake or success on the targeted skills after being exposed to the items, or that they remember their previous answer and provide the same answer on a subsequent testing. These elements could lead to either an increase or a decrease in test-retest reliability.

It was hypothesized that 1) individual children's raw scores on the MilBec would systematically increase with age; 2) the results in the MilBec would correlate positively with measures of receptive language; and 3) when a child receives a score of 1 at a particular time, a score of 1 for this item should also be received on subsequent testing.

Method

Participants

Participants were 10 typically developing children (6 boys and 4 girls) aged between 12 and 66 months at the beginning of the study and living in the Greater Montreal area. All children had normal hearing, based on a pure tone hearing screening in each ear at 4 frequencies (500Hz,

1000Hz, 2000Hz, 4000Hz), with an intensity level of 10 dB HL (30dB HL at 500Hz because of ambient noise) using a portable screener. All children had normal nonverbal cognition (above 85), based on the Brief-IQ of the Leiter International Performance Scale-Revised (Leiter-R: Roid & Miller, 1997). There was a total of 7 participating parents as two families had more than one participating child (one had two children, and the other had three children). The maternal education of the mother of the two siblings was a high school diploma (which is equivalent to a grade 11), all the other mothers had some university level education. There were four families (four children) recruited from Study 1, and three families (six children) recruited through word of mouth. Based on the information provided by the parents on the demographic questionnaire, all participants were monolingual French speakers, with both parents speaking French at home; seven children were exposed exclusively to French, three were exposed to a language other than French for 5 hours or less a week. All children were judged to have typical language development based on the absence of parental concerns, supported by clinical judgment based on the performance during the conversational language samples during play and the results on an unpublished French translation of the Reynell Developmental Language Scales III (RDLS-III: Edwards, Fletcher, Garman, Hughes, Letts, & Sinka, 1997). For the children who were less than 2-year-old at the intake, the clinical judgment was also based on the evolution of the child's performance as they got older.

Procedures

All parents signed the consent form of the project, which was approved by the Institutional Review Board of the Faculty of Medicine of the McGill University. The demographic questionnaire and the consent form were filled out by the parents during the first testing session. The first author, who is a licensed speech-language pathologist in the Province of

Quebec, evaluated the children 6 times or until they had reached 73 months, with a four-month interval between each testing time. The period between two testing sessions was on average 126 days, with a standard deviation of 17 days (range: 91-162 days). The testing time of each session varied between 30 and 60 minutes, depending on the child's age and cooperation. The children were tested in their home or the home of the tester; the evaluations were video- and audio-recorded using a Canon FS100.

At each testing time between 12 and 71 months, one of the parents filled out the MilBec, and the child's receptive language skills were evaluated using the following measures: 1) An unpublished French translation of the Reynell Developmental Language Scales III (RDLS-III: Edwards et al., 1997). This test was selected because it is widely used in clinical settings and it was developed to assess the language skills of children between 18 months to 7 years. It has also been used in previous published studies on Quebec French children (Elin Thordardottir, Rothenberg, Rivard, & Naves, 2006). However, a plateau effect was observed on the children's score at around age four. 2) The French adaptation of the Test of Auditory Comprehension of Language-Revised (Commonly referred to in French clinical settings as "Carrow": Groupe coopératif en orthophonie, 1999), unpublished Quebec French adaptation of the Test for Auditory Comprehension of Language [TACL-R]; Carrow-Woolfolk, 1985). This test was added to the protocol from the third testing point for children 3-year-old or older, because of the plateau effect seen on the RDLS-III. The Carrow was selected because it had been developed for children between 3- and 9-year-old, and is also widely used in clinical settings, although usually for children in their first school years due to the availability of local norms at that age range. Normative data have been collected for this test for Quebec French children (Groupe cooperatif, 1999; Elin Thordardottir, Keheyia, Lessard, Sutton, & Trudeau, 2010).

In addition to the RDLS-III and the Carrow, the following measures were collected but are not reported here. 3) The Edmonton Narrative Norms Instrument (ENNI: Schneider, Dubé, & Hayward, 2005), was administered for children from the age of 48 months only, as it is not an appropriate task for younger children, and 4) a 10-minute conversational language sample during play, used to inform clinical judgment.

During the last session or in an additional testing session, the hearing screening and the nonverbal intelligence test were administered.

Analyses and Results

Developmental Sensitivity

The first question of interest was whether MilBec scores increased with age for individual children. As shown in Figure 3, the results showed that this was the case to some extent. Indeed, there was a positive linear trend between MilBec scores and age, with a rapid increase in scores up to 44 months, followed by a sudden decrease in the slope around 44-48 months, with some children at that age starting to obtain the maximum score. To confirm the association of increasing scores and increasing age, further analyses were considered. Because there was a non-independence of the data points, since several points were collected from each child, and because of the large variation in age between children, neither Pearson correlations nor MANOVA procedures could be used to confirm the scores increase over time. Instead, the most appropriate procedure for this dataset was judged to be the generalized estimating equations (GEE), an extension of the general linear model. The GEE analysis makes an adjustment to compensate for the dependency between measures, uses all the data points in the calculation of the regression coefficients, and can handle missing data. To perform the adjustment, two elements had to be assumed: the type of development with time (e.g. linear, quadratic) and the

correlation structure of the repeated measures (Twisk, 2013).

[Insert Figure 3 about here]

Preliminary analyses were performed to determine the most appropriate assumptions for the adjustment. As proposed in Twisk (2013), the type of development was determined by performing a MANOVA for repeated measures. The model using a linear relation was the one with the highest F value, and was thus judged to provide the best fit. For the correlation structure, Twisk (2013) proposed that the correlations of the residuals of a linear regression between the variables should be analyzed. For this analysis, 12 age groups with slightly overlapping age-range were created. The series of testing point of each child was assigned to the most age-appropriate consecutive groups, with only one data point per group per child; since the children were not tested 12 times, each child thus had several missing data points, either at the younger or the older age groups. These 12 groups would be used in the GEE analysis, and are thus used to determine the correlation structure. Two of the oldest children had to be excluded from correlation analysis, because they had received the maximum score on the MilBec at each data collection time, and only the correlations from groups with at least four participants were considered. The pattern of correlations that best described the relationship between the nonstandardized residuals of the linear regression between the MilBec score and testing time was the unstructured one, with the correlation varying from strongly negative to strongly positive. Using a type III GEE analysis on all subjects based on the 12 age groups, and assuming a linear relation and an unstructured correlation, a significant effect of time (p = .000) was found. The model fit obtained was 380,736, using a corrected quasi-likelihood under a criterion of independent model.

Concurrent Validity

Visual inspection of Figure 4 showed a positive linear relation between MilBec and RDLS-III scores. The relationship showed a pronounced slope and a high density of scores at the upper right corner. This indicated that when the child was near the ceiling on the RDLS-III (with a score between 55 and 62), the child's score was also near the ceiling on the MilBec (with a score between 35 and 39).

[Insert Figure 4 about here]

The analysis comparing the Carrow and the MilBec included fewer data points, and covered a more restricted range of MilBec scores, because the Carrow was administered only to children older than 3 years from the third testing point. Figure 5 showed a positive linear relation between the MilBec and the Carrow scores, which had a much less pronounced slope than what was observed with the RDLS-III. Also, there was a wide range of Carrow scores corresponding to the maximum score on the MilBec: when the child received a score between 35 and 39 on the MilBec, the concurrent child's score on the Carrow was between 80 and 120. This indicated that the score of children on the MilBec was at the ceiling, while the Carrow captured changes in their language skills.

[Insert Figure 5 about here]

Reliability

Test-retest reliability investigates whether a similar result is obtained at two different points in time. The longer the period between the two evaluations, the more variations between test scores can be expected. In the current study the time elapsed between the two test times was 126 days on average. During a period when language is developing, it is expected that test scores would increase due to maturation. Test-retest reliability was evaluated in this study by comparing

the percentage of items receiving the same score on two consecutive test administrations. First, each pair of test times was classified into one of three categories. The first category encompassed cases were the same parent filled out the questionnaire at both Time X and at Time X+1 (the next testing time), which can be used as an indication of test-retest reliability. Among all the pairs, 66% were from this category. The second category encompassed cases where the parent who filled out the questionnaire at time X and Time X+1 was not the same, which can be used to assess inter-judge reliability. Among all the pairs, only 11% were from this category. Finally, the third category encompassed the cases where it was unknown who filled out the questionnaire either at Time X or Time X+1, making it impossible to determine in which of categories the pairs should be classified. Among all the pairs, 22% were from this unspecified category.

Once the pairs were classified, for each test item the score at Time X was compared to the score at Time X+1. Each pair that received a score of 1 on Time X, but a score of 0 on Time X+1, was identified as a score decrease. If the pair went from 0 at Time X to 1 at Time X+1, it was identified as a score increase. The total number of items with a score decrease and a score increase was calculated for each pair of testing, along with the percentage of items having received the same score on both Time X and Time X+1 (see Table 2). The reliability index was 93.4% for the test-retest category, 91.8% for the inter-judge category, and 90.7% for the unspecified category. When all the 46 pairs of items for all of the 39 items are taken into consideration, the test-retest reliability index of the MilBec was 93.1%.

[Insert Table 1 about here]

As a 4-month period elapsed between Time X and Time X+1, some developmental changes may have occurred. A score increase could be caused either by the child having acquired the skill within the 4-month period between the two test administrations, or by an error factor.

On the other hand, a score decrease could be caused by an error factor, or by other elements related to the acquisition pattern of the items that would make the parent believe that the child no longer possesses an ability he or she previously had. An item with a high proportion of score decreases should be excluded from the screening tool, as the acquisition of the target language skill at one point would not predict its acquisition later. In order to determine if some items were more susceptible to decrease in score, a total number of decreases were calculated for each item on all 46 pairs of testing, regardless of who filled out the questionnaire. This analysis showed that 1.5% of the Time X and Time X+1 paired items showed a decrease of score (28 decreases occurring out of a total of 1794 opportunities across children). These score decreases involved 17 items in total, and the maximal number of decreases for a given item was three. Thus, no items should be considered for removal of the questionnaire based on poor test-retest reliability.

Finally, examination of the questionnaire showed that item #6 asking "is your child producing some words (simplified or not)", was occasionally re-scored by the parent after questioning from the examiner, since some parents answered "no" even if their child was already speaking in sentences. These parents commented that they answered "no" because their child did not use simplified speech. This item should be carefully investigated in a further study to determine if it should be removed from the questionnaire or simply reworded. Also, a few parents requested verbal clarification on one of the items added after the pilot study, namely the use of contracted article; better examples should be provided in a subsequent version of the MilBec to clarify this question.

Discussion

This article presents the MilBec, a French adaptation of the Dutch parent questionnaire by Luinge et al. (2006). The MilBec is a parent questionnaire intended to screen for language

difficulties in children between 12 and 71 months. The results from both the cross-sectional study (Study 1) and the longitudinal study (Study 2) showed that children's scores increased significantly with age. Furthermore, although a ceiling effect occurred using the preliminary version of the screener, the addition of items effectively raised the age at which the ceiling effect occured (see Figure 3). Indeed, the increase in score was quite rapid up to the age of 42 months in the MilBec, and most children were close to or at the ceiling between 60 and 66 months. In contrast, in the preliminary version the ceiling effect started to show from 36 months, with virtually no variability of scores in the 5-year-old group.

The validity of the MilBec was addressed in Study 2 by comparing the children's results on the MilBec with their results on two receptive language tests, namely the RDLS-III and the Carrow. These analyses showed that the MilBec taps into language skills that are acquired at a similar age range as those targeted in the RDLS-III (see Figure 4). Indeed, Brevet and Cambournac (2009) showed that European children also reached ceiling around 4-year-old on the RDLS-III. Furthermore, the groups of 4-year-old and 5-year-old children in their study received similar scores, while the group of 3-year-old children received lower scores. The findings of the study suggest that the MilBec would be sensitive to development up to approximately 36-42 months, with a ceiling effect occurring around 60 months. This does not preclude the use of the MilBec as a language screening tool for children up to 71 months. Indeed, in a screener test the items should capture the minimal skill level under which a language impairment can be suspected, not his or her maximal performance.

Based on these findings, we can conclude that the MilBec has the potential to be a quick and cost-effective screening tool for language difficulties in children. The fact that the MilBec is sensitive to development is most likely related to the fact that the items target different aspects of

language that are emerging at different points in time. Emerging elements are more salient, making them easier for parents to identify and judge as acquired or not by their child. Because the items target different language domains, passing the screener does not mean the same things at different ages. For examples, the earlier items target more the emergence of babbling and early words, whereas for older children the items target more pre-literacy and metalinguistics skills. The results showed that the MilBec is developmentally sensitive for children with TD, particularly before the age of 42 months. Since the MilBec included various aspects of language, it is hypothesized that it will allow the identification of children with preserved vocabulary skills, but who present difficulties in other language domains. Whether children with language impairment will obtain a significantly lower score on the MilBec than children with TD depends on whether the items target elements differentiating these two groups of children. Some authors proposed that the difference between children with and without language impairment stem from a different rate of language development, whereas other proposed that the difference is qualitative, with the presence of deviances (see Leonard, 2014, for a discussion on the different views). For example, Rescorla (2011) concluded in her review of studies on language outcomes that the language skills of children with TD and late-talkers only differ quantitatively, not qualitatively. On the other hand, proponents of the deviant language view advanced that children with language impairment make particular error patterns (e.g. Jakubowicz et al., 1998; Rice & Wexler, 1996; Rice, 2004). Whether the selected items will allow the identification of children at risk of language impairment will have to be investigated in a future study.

The principal limitations of the longitudinal study are the fact that some of the participants were siblings, which decreases the number of participating parents, and the various ages at which the child entered the study. In addition, most children have a mother with a high

degree of education. Since parents from low socio-economic status have been found to answer differently than parents with higher socio-economic status on some parent report measures (Feldman et al., 2000), the present results should thus not be generalized to this population. Before the MilBec can be used as a screener, it has to be normed and its diagnostic accuracy has to be established.

Acknowledgments

We want to thank all the parents and children who participated in the study.

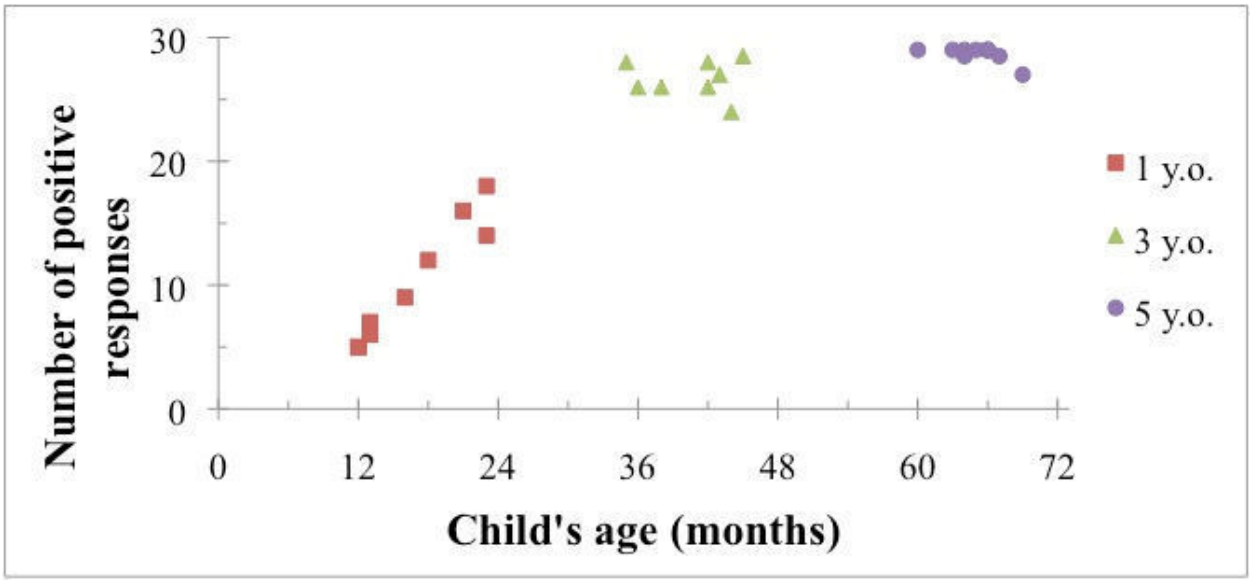


Figure 1. The raw score of each child on the first version of the parent questionnaire as a function of age.

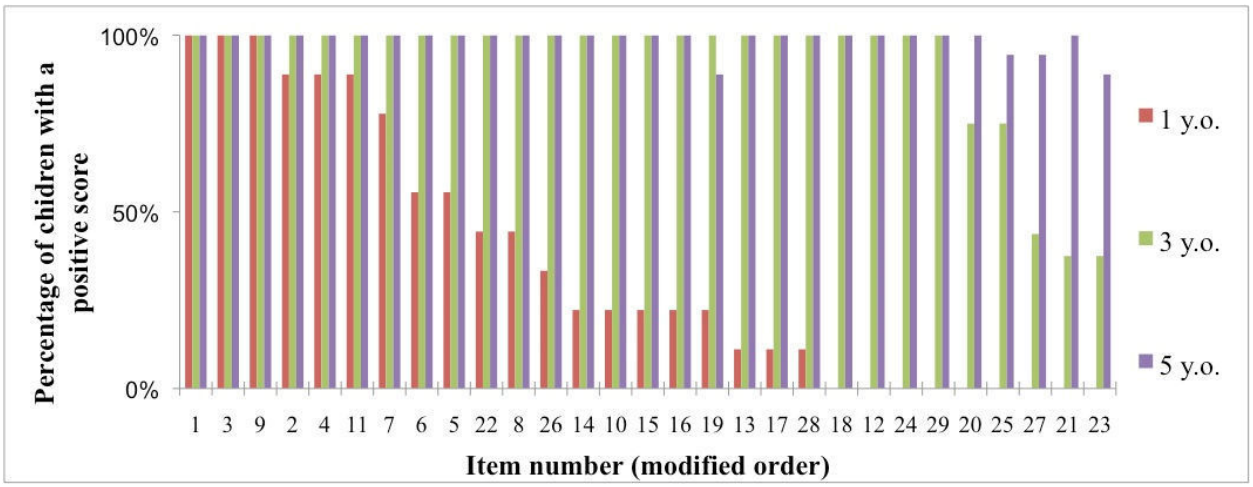


Figure 2. Percentage of children with a positive response for each item on the first version of the parent questionnaire, with items reordered by item difficulty.

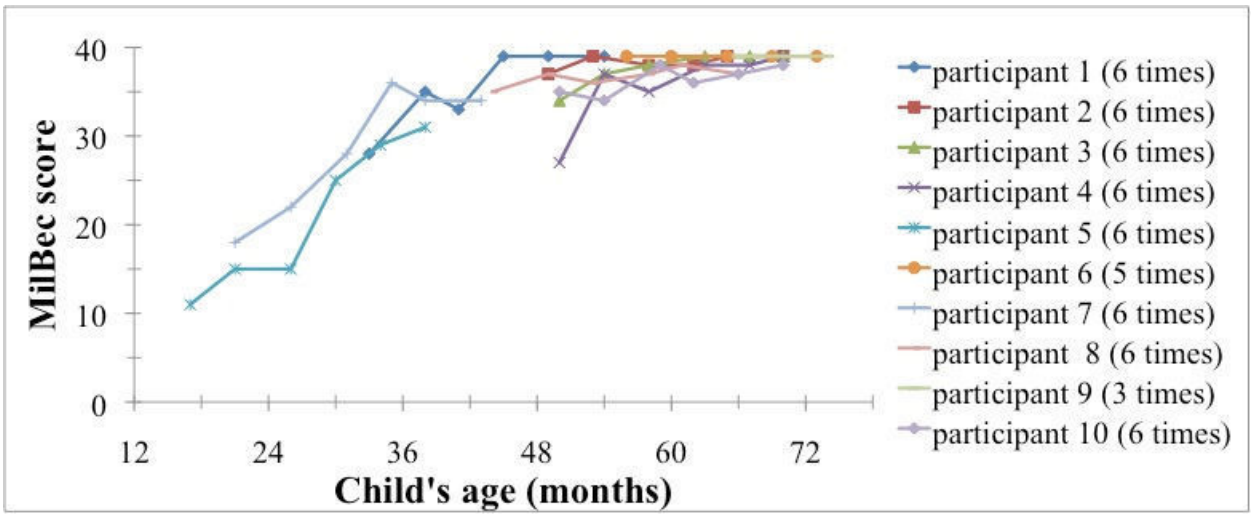


Figure 3. Increase in total scores on the MilBec for each subject at each longitudinal testing point. For each participant, the number of testing sessions performed is indicated.

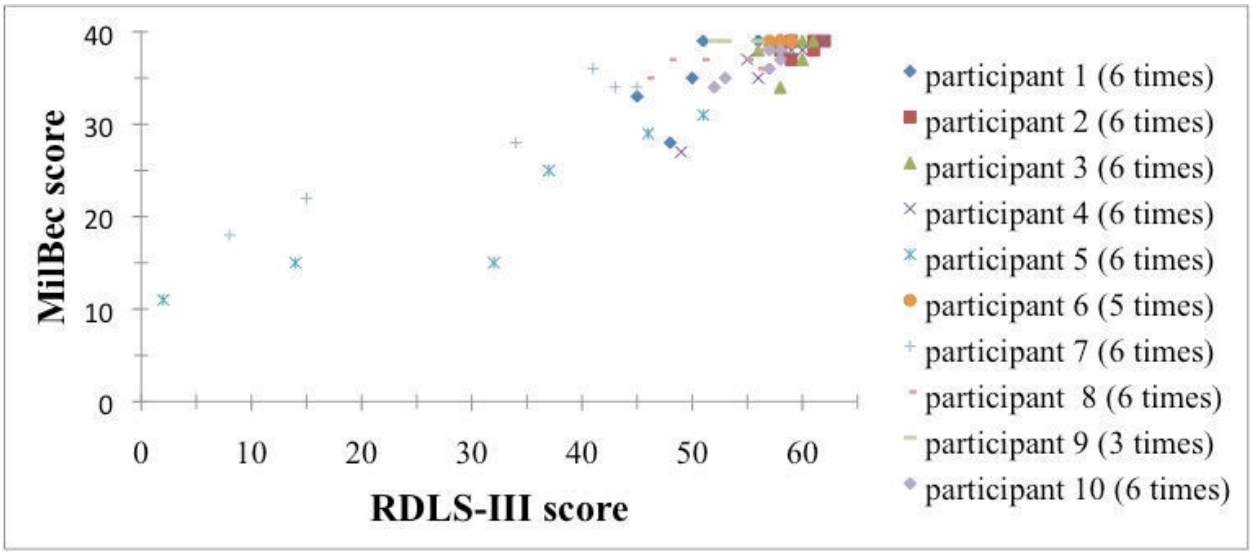


Figure 4. Scatterplot showing longitudinal performance on the MilBec and the RDLS-III for each participant.

Note: for each participant, the number of testing sessions performed is indicated.

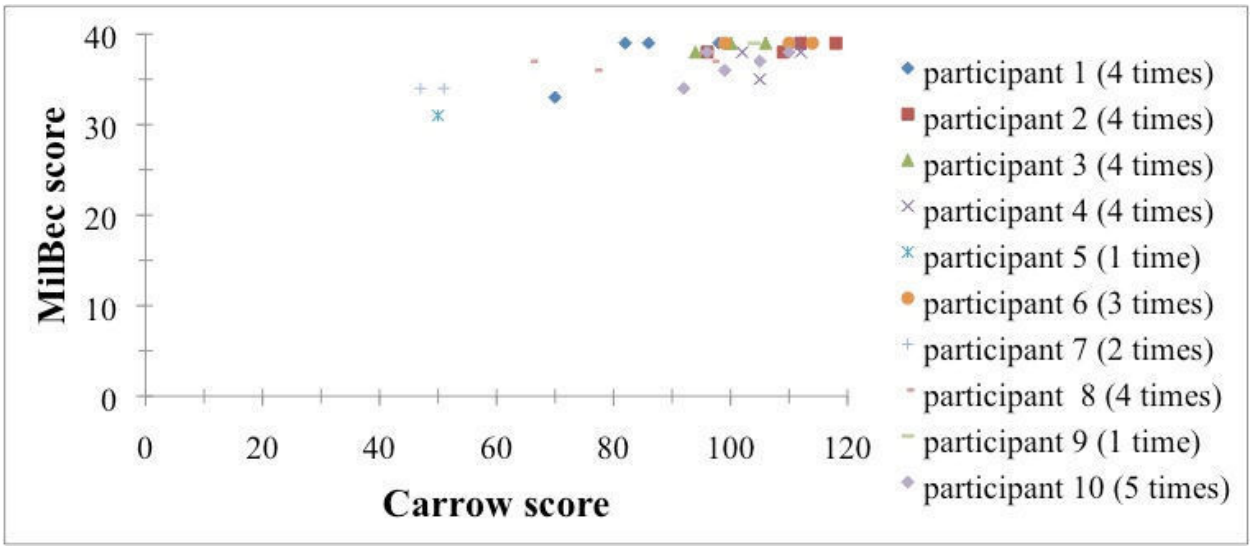


Figure 5. Scatterplot showing longitudinal performance on the MilBec and the Carrow for each participant.

Note: For each participant, the number of testing sessions performed is indicated.

Table 1. Calculation of reliability index

	Number of opportunities	Score decrease at Time X+1	Score increase at Time X+1	Reliability index
test-retest	31 pairs * 39 items = 1209 opportunities	14 decreases (range per pair : 0- 3)	59 increases (range per pair : 0- 10)	1136/1209 = 93.9%
inter-judge	5 pairs * 39 items = 195 opportunities	6 decreases (range per pair : 0- 3)	9 increases (range per pair : 1- 4)	180/195 = 92.3%
unspecified	10 pairs * 39 items = 390 opportunities	8 decreases (range per pair : 0- 2)	28 increases (range per pair : 0- 10)	354/390 = 90.7%
Total	46 pairs * 39 items = 1794 opportunities	28 decreases	96 increases	1670/1794 = 93.1%

Note: Test-retest: the same parent filled out both questionnaires. Inter-judge: different parents filled out the questionnaire. Unspecified: information on the person who filled out the questionnaire is lacking either at Time X or Time X+1.

Annex 1: The adapted version of the Questionnaire

**** Veuillez noter qu`à la fin du questionnaire, nous vous demandons d`indiquer le temps qu`il vous a fallu pour répondre au questionnaire ****

**** Comme le même questionnaire est utilisé pour tous les enfants entre 1 et 5 ans, il est normal que les enfants plus jeunes aient une majorité de réponses négatives ****

Questionnaire ` prière de répondre à toutes les questions	oui	non
1. Est-ce que votre enfant produit une variété de sons qui ressemblent à des consonnes et des voyelles?	☐	☐
2. Est-ce que votre enfant dit « maman » ou « papa »?	☐	☐
3. Est-ce que votre enfant comprend la signification de « non »?	☐	☐
4. Est-ce que votre enfant produit quelques mots (simplifiés ou non)? (ex. *ati* pour « partie »; *la* pour « lait », *non*)?	☐	☐
5. Est-ce que votre enfant peut identifier une ou plusieurs parties du corps, par exemple en répondant à des questions du type « *Où est ton nez?* »?	☐	☐
6. Est-ce que votre enfant dit environs 10 mots différents?	☐	☐
7. Est-ce que votre enfant peut pointer certains objets que vous nommez?	☐	☐
8. Est-ce que votre enfant peut combiner deux mots? (ex. *veux biberon; là bobo; dedans chien)*?	☐	☐
9. Est-ce que votre enfant comprend des demandes simples de deux mots (ex. « viens manger », « assis-toi »)	☐	☐
10. Est-ce que votre enfant fait des suites de trois mots? (ex. : *veut monter Grégoire; pas mettre ça; moi goûter fraises*)	☐	☐
11. Est-ce que votre enfant comprend des phrases de trois-quatre mots (ex. « touche pas à ça », « sur la table », « attend ton tour »)	☐	☐
12. Est-ce que votre enfant fait des phrases complètes de trois ou quatre mots? (ex. : *on dirait une fille; il criait tout le temps; raconte une histoire*)	☐	☐
13. Est-ce que votre enfant nomme correctement certaines couleurs?	☐	☐
14. Est-ce que votre enfant pose des questions?	☐	☐
15. Est-ce que votre enfant utilise le bon ordre des mots dans ses phrases?	☐	☐
16. Est-ce que votre enfant mentionne le sujet dans ses phrase, c'est-à-dire est-ce qu'il indique qui fait l'action? (ex. *Martin* dans « Martin va à la piscine », *tu* dans « Tu viens? »)	☐	☐
17. Est-ce que votre enfant utilise des mots qui qualifient / décrivent d'autres mots? (ex. : *grande* et *rouge* dans « grande maison rouge »)	☐	☐

Questionnaire ` prière de répondre à toutes les questions	oui	non
18. Est-ce que votre enfant peut répéter une histoire en se basant sur des images?	☐	☐
19. Est-ce que votre enfant raconte spontanément des événements de sa journée? (ex. : quelque chose qui est arrivée à la garderie)	☐	☐
20. Est-ce que votre enfant utilise le masculin et le féminin correctement la majorité du temps?	☐	☐
21. Est-ce que votre enfant utilise généralement le pluriel correctement? (ex. : yeux/œil, chevaux/cheval)	☐	☐
22. Comprenez-vous environs la moitié (50%) de tout ce que votre enfant dit?	☐	☐
23. Est-ce que votre enfant utilise correctement les passés composés irréguliers? (ex. : couru, mis, pris)	☐	☐
24. Est-ce que votre enfant fait de longues phrases avec plusieurs verbes? (ex. : *Quand le soleil se couche, il fait noir*; *Maman dit tu dois venir*)	☐	☐
25. Est-ce que votre enfant remplace parfois le mot qui désigne un objet par un pronom? (ex. : *la* dans « Je la veux », au lieu de dire « Je veux la pomme »)	☐	☐
26. Comprenez-vous environs les trois-quarts (75%) de tout ce que votre enfant dit?	☐	☐
27. Est-ce que votre enfant parle comme un adulte, en ce qui a trait à la complexité des phrases?	☐	☐
28. Est-ce que votre enfant comprend des consignes à deux étapes ou plus? (ex. : « Tu dois ranger tes jouets avant d'aller jouer dehors »)	☐	☐
29. Comprenez-vous la quasi-totalité (près de 100%) de tout ce que votre enfant dit?	☐	☐

- Combien de temps vous a-t-il fallu pour répondre à ce questionnaire ?
  ☐ moins de 5 minutes ☐ 5-10 minutes ☐ 10-15 minutes ☐ plus de 15 minutes

Merci beaucoup de votre participation, elle est très appréciée!

Si vous le désirez, vous pouvez nous faire parvenir vos commentaires concernant le projet de recherche, le questionnaire (les questions, le choix d'exemple, etc.) en utilisant la feuille ci-jointe.

Les commentaires seront pris en considération lors de la révision du questionnaire, qui suivra l'analyse des résultats obtenus. Alors, n'hésitez pas à nous indiquer ce que vous avez aimé ou pas, et de nous indiquer les questions que vous avez trouvées plus difficiles à répondre!!!

Commentaires :

_____

_____

Annex 2: The MilBec

MilBec - Test de dépistage du langage pour enfants de 12 à 71 mois

Nom de l'enfant :	Genre : ☐ masc. ☐ fém.
Date de naissance (jj-mm-aaaa) :	Âge (mois):
Complété le (jj-mm-aaaa) :	par : ☐mère ☐père ☐ autre :

Consignes : Indiquez « oui » si la réponse est vraie présentement, ou l'était lorsque votre enfant était plus jeune.

Questionnaire – prière de répondre à toutes les questions *Comme le même questionnaire est utilisé pour tous les enfants, il est normal que ceux plus jeunes aient une majorité de réponses négatives.*	oui	il me semble	je ne crois pas	non
1. Est-ce que vous considérez que votre enfant a un langage suffisamment développé, en comparaison aux autres enfants de son âge?	☐	☐	☐	☐
2. Est-ce que votre enfant produit, ou produisait quand il était petit, une variété de sons qui ressemblent à des consonnes et des voyelles ?	☐	☐	☐	☐
3. Est-ce que votre enfant comprend la signification de « non »?	☐	☐	☐	☐
4. Est-ce que votre enfant comprend des consignes simples de deux mots ? (ex. « viens manger », « assis-toi »)	☐	☐	☐	☐
5. Est-ce que votre enfant dit « maman » ou « papa »?	☐	☐	☐	☐
6. Est-ce que votre enfant produit quelques mots (simplifiés ou non)? (ex. « ati » pour *partie*; « la » pour *lait*, *non*)?	☐	☐	☐	☐
7. Est-ce que votre enfant comprend des phrases de trois-quatre mots ? (ex. « touche pas à ça », « sur la table », « attend ton tour »)	☐	☐	☐	☐
8. Est-ce que votre enfant vous montre du doigt les objets qui l'intéressent?	☐	☐	☐	☐
9. Est-ce que votre enfant peut identifier une ou plusieurs parties du corps, par exemple en répondant à des questions du type « Où est ton nez? »?	☐	☐	☐	☐
10. Est-ce que votre enfant dit environ 10 mots différents?	☐	☐	☐	☐
11. Est-ce que votre enfant peut combiner deux mots? (ex. *veux biberon; là bobo; dedans chien; papa parti)*?	☐	☐	☐	☐
12. Comprenez-vous environ la moitié (50%) de tout ce que votre enfant dit?	☐	☐	☐	☐
13. Comprenez-vous environ les trois-quarts (75%) de tout ce que votre enfant dit?	☐	☐	☐	☐
14. Est-ce qu'il vous est inutile de « traduire » ce qu'a dit votre enfant pour qu'une personne non familière le comprenne, plus des trois-quarts (75%) du temps?	☐	☐	☐	☐
15. Est-ce que votre enfant fait des suites de trois mots? (ex. : *veut monter Grégoire; pas mettre ça; moi goûter fraises*)	☐	☐	☐	☐
16. Est-ce que votre enfant pose des questions (avec des phrases complètes ou non)? (ex. : *Papa parti?* **;** *est où Maman?* ; *pourquoi?*)	☐	☐	☐	☐
17. Est-ce que votre enfant fait toujours ses phrases avec les mots dans le bon ordre?	☐	☐	☐	☐
18. Est-ce que votre enfant mentionne le sujet dans ses phrases, c'est-à-dire est-ce qu'il indique qui fait l'action? (ex. *Martin* dans « Martin va à la piscine », *tu* dans « Tu viens? »)	☐	☐	☐	☐
19. Est-ce que votre enfant raconte spontanément des événements de sa journée? (ex. : quelque chose qui est arrivée à la garderie)	☐	☐	☐	☐
20. Est-ce que votre enfant nomme correctement certaines couleurs?	☐	☐	☐	☐

MilBec - Test de dépistage du langage pour enfants de 12 à 71 mois

Questionnaire – prière de répondre à toutes les questions *Comme le même questionnaire est utilisé pour tous les enfants, il est normal que ceux plus jeunes aient une majorité de réponses négatives.*	oui	il me semble	je ne crois pas	non
21. Est-ce que votre enfant possède dans son vocabulaire trois mots ou plus qui qualifient ou décrivent d'autres mots? (ex. : *grande* et *rouge* dans « grande maison rouge »; *très* dans « très vite »)	☐	☐	☐	☐
22. Est-ce que votre enfant comprend des consignes à deux étapes ou plus? (ex. : « Tu dois ranger tes jouets avant d'aller jouer dehors »)	☐	☐	☐	☐
23. Est-ce que votre enfant fait des phrases complètes de trois ou quatre mots? (ex. : *on dirait une fille; il criait tout le temps; raconte une histoire*)	☐	☐	☐	☐
24. Est-ce que votre enfant peut répéter une histoire en se basant sur des images?	☐	☐	☐	☐
25. Est-ce que votre enfant fait de longues phrases avec plusieurs verbes? (ex. : *Quand le soleil se couche, il fait noir*; *Maman dit tu dois venir*)	☐	☐	☐	☐
26. Comprenez-vous la quasi-totalité (près de 100%) de ce que votre enfant dit?	☐	☐	☐	☐
27. Est-ce que votre enfant utilise le masculin et le féminin correctement la majorité du temps? (ex. *la pomme*, *la gentille fille*, *un tapis*, *le beau chien*)	☐	☐	☐	☐
28. Est-ce que votre enfant remplace parfois le mot qui désigne un objet par un pronom? (ex. : *la* dans « Je la veux », au lieu de dire « Je veux la pomme »)	☐	☐	☐	☐
29. Est-ce que votre enfant utilise le pluriel correctement plus de 75% du temps? (ex. : yeux/œil, corail/coraux)	☐	☐	☐	☐
30. Est-ce que votre enfant utilise correctement le passé composé des verbes irréguliers fréquemment utilisés? (ex. : couru, mis, pris)	☐	☐	☐	☐
31. Est-ce que votre enfant utilise les articles contractés correctement plus de 75% du temps? (ex. *du* pour *de le*; *au* pour *à le*)	☐	☐	☐	☐
32. Est-ce que votre enfant commente parfois la similitude entre des mots liés par le sens? (ex. : *la robe fleurie a des fleurs?; la feuille est lignée parce qu'elle a des lignes*)	☐	☐	☐	☐
33. Est-ce que votre enfant est capable de trouver des mots qui riment? (ex. : *moufette* va avec *toilette*; *chat* va avec *rat*; *tapis* va avec *souris*)	☐	☐	☐	☐
34. Est-ce que votre enfant est capable de trouver des mots commençant avec le même son? (ex. : *part* va avec *petit*; *lapin* va avec *loupe; manteau* va avec *melon*)	☐	☐	☐	☐
35. Est-ce que votre enfant informe plus de 75% du temps du lieu et des personnes impliquées de manière suffisante, lorsqu'il raconte un événement de sa journée?	☐	☐	☐	☐
36. Est-ce que votre enfant indique clairement plus de 75% du temps dans quel ordre les événements ce sont déroulés, lorsqu'il raconte une histoire?	☐	☐	☐	☐
37. Est-ce que votre enfant peut réciter l'alphabet sans erreur plus de 75% du temps?	☐	☐	☐	☐
38. Est-ce que votre enfant est capable de reconnaître plus de 3 mots écrits? (ex. : son nom, papa, maman, marque de commerce)	☐	☐	☐	☐
39. Est-ce que votre enfant regarde un livre en le tenant à l'endroit, en commençant au début et en tournant les pages une à la fois, plus de 75% du temps?	☐	☐	☐	☐
Sous-totaux				
Total				

Preface to Chapter 2

In the previous chapter, the procedure used to adapt the Dutch parent questionnaire by Luinge et al. (2006) into French was presented. In Study 1, conducted as part of the adaptation procedure, a cross-sectional study pointed to some required modifications. The final version of the questionnaire, named the MilBec, was then tested in Study 2 with a longitudinal study aimed at documenting the MilBec's validity, reliability, and developmental sensitivity. The results suggested that the MilBec is a promising tool to screen for language impairment. However, the test was not equally sensitive to developmental changes: there was a nonlinear relationship between age and MilBec score and some children reached ceiling from around 36-42 months. Whether the MilBec can be an adequate screening tool for language impairment for older children should thus be determined.

In this chapter, the MilBec's usefulness in identifying children with language impairment is investigated using a three-step process. In Study 1, cross-sectional data from typically developing (TD) monolingual children between 12 and 71 months will allow the documentation of the psychometric properties of the MilBec and provide preliminary norms. Study 2 is a diagnostic accuracy study conducted with TD children and children with specific language impairment (SLI) between 36 and 71 months, with the purpose of further investigating the MilBec's concurrent validity and reliability. This study will also explore the MilBec's sensitivity and specificity, two elements used to determine a test's diagnostic accuracy. Finally, the usefulness of two potential clinical markers of SLI, namely use of subjects and object clitics, will be investigated in Study 3 by analyzing the performance of the children on two specific items of the MilBec.

It was decided to exclude from Study 1 children who are at higher risk of language delay, such as when the parents have important concerns regarding their child's language development.

Indeed, Peña, Spaulding, and Plante (2006) showed in a simulation study that the diagnostic accuracy of tests decreased when the normative data included the score of children with a disorder, as this inclusion decreases the group's average score and increases its standard deviation. The inclusion of a higher proportion of children with a disorder in the norming sample has also been proposed as one of the possible factors leading the PPVT-4 to have a lower discriminative value than the PPVT-3 (Spaulding, Hosmer, & Schechtman, 2013).

Together, the results of these two studies will provide a first evaluation of the MilBec's adequacy as a screening tool for the identification of children with SLI. Once this is established, further analyses can be conducted to investigate the language profiles of monolingual children with and without impairment, and how bilingual children's profiles compared to them.

Chapter 2: Preliminary Normative Data, Concurrent Validity and Diagnostic Accuracy of the MilBec, a Screener for Language Impairment in French-Speaking Children between 12 and 71 Months from Quebec.

Marianne J. Paul and Elin T. Thordardottir

School of Communication Sciences and Disorders at McGill University, Montréal, Canada

Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS)

To be submitted in a journal for family doctors or pediatricians.

Abstract

Purpose: This study investigated the psychometric properties of the MilBec, a French adaptation of the Dutch parent questionnaire by Luinge et al. (2006), aimed at identifying language difficulties in children between 12 and 71 months. The diagnostic accuracy of two items targeting potential clinical markers, the use of the subject and object clitic, was also investigated. **Method**: Three studies were conducted: 1) A cross-sectional normative study of 85 children between 12 and 71 months (17 subjects per 12-month age groups); 2) a validation study with 15 children with typical development (TD) and 15 children with moderate to severe specific language impairment (SLI) between 36 and 71 months; 3) an analysis of the performance of children with and without SLI on the items targeting the two potential clinical markers. **Results**: *Study 1*: The correlation of the MilBec score with age is extremely high (r = .921; p = .000) for 12- to 39-month-old children (n = 42) and is high (r = .598; p = .000) for 40- to 71-month-old children (n = 43). *Study 2*: MilBec scores were significantly different for children with and without SLI, with a very large effect size (Cohen's d = 3.4). MilBec scores were significantly correlated with other language tests. The patterns of correlations varied between groups when the groups were analyzed separately. The MilBec's sensitivity was 100% and specificity was 93%. *Study 3*: The potential clinical markers had a low sensitivity (20% for subject use after 27 months and 53% for object clitic use after 31 months), but a high specificity (100% and 97%, respectively). **Conclusions**: The MilBec is a promising screening tool for identifying children with language impairment. Failing the items about subject use for children over 27 months or object clitic after 31 months are strong indicators of SLI, but do not capture most cases of SLI.

Keywords: Specific language impairment (SLI), French, screening test, parent questionnaire, validation, norms, diagnostic accuracy, preschoolers

Introduction

The role of family physicians in screening for language impairment is important and has been highly recommended since the 1980s (Tervo & Balaton, 1980). The importance of language screening is highlighted in the Rourke Baby Record (Rourke, Leduc, & Rourke, 2014) and the ABCdaire du suivi périodique de l'enfant de 0 à 5 ans ("ABC of periodical follow up of children between 0 and 5 years", free translation) (Brunet, Cossette, Cousineau, & Lemieux, 2015). In these two interview guides used during well-child visits, language is one of the many aspects of the child's development included. Furthermore, parents have high regard for their family physician's opinion regarding their child's language development. For example, Carscadden et al. (2010) mention the following anecdote: several parents whose child had failed a language screening declined the offer of a free evaluation by a speech-language pathologist, because their physician told them not to worry about their child's language. However, recent findings (e.g. Rescorla, 2011) showed that the proposed items in well-care child checklists are not sufficient to identify many of the children with low language skills for whom a more complete language evaluation is warranted.

It has been proposed that family physicians should use validated parent questionnaires in order to more accurately identify children with delays and use less stringent thresholds at which referrals should be made (King & Glascoe, 2003). Indeed, elicitation of parental concerns regarding speech and language development yielded a specificity of 83% and sensitivity of 72% (Glascoe, 1991). There are three main advantages of using parental reports instead of direct testing. First, it reduces the time the professional has to spend on the assessment. Second, it may provide a more representative estimation of the child's skills, since the parent can report on skills that are infrequently used and their observations are based on the usual performance of the child, whereas the professional's assessment may be affected by the child's shyness or unwillingness to

interact with a stranger. Finally, using parental report avoids possible negative effects of direct assessment on the child's self-esteem and self-consciousness.

Studies have documented the psychometric properties of different parent questionnaires and the adequacy of this method to assess language skills in young children. For example, the Language Development Survey (Rescorla, 1989) can correctly identify the majority of children with language impairment (sensitivity of 91%) and with typical development (specificity of 87%) (Klee, Carson, Garvin, Hall, Kent, & Reece, 1998). One of the most widely used parent report measures is the MacArthur-Bates Communicative Development Inventory (CDI: Fenson et al., 1993) which assesses vocabulary knowledge and first word combinations in children between 8 and 30 months old. This measure had been used by many researchers, adapted to many languages, and repeatedly shown to have good psychometric properties (e.g. Boudreault, Cabirol, Trudeau, Poulin-Dubois, & Sutton, 2007; Elin Thordardottir & Ellis Weismer, 1996; Heilmann, Weismer, Evans, and Hollar, 2005; Marchman & Martinez-Sussmann, 2002; Childers, Vaughan, & Burquest, 2007). One issue related to parent questionnaires is that some parents may fail to answer some of the items if, for example, they do not know the answer or understand the question. Issues related to how these missing data are to be interpreted, and when it is considered that there is too much missing information for the test to be reliable, have to be decided beforehand.

Buschmann et al. (2008) had also stressed the importance of earlier referral, because early language delay would be an important indicator of various developmental problems involving language difficulties as a primary or secondary problem. Rescorla (2011) reached a similar conclusion regarding the importance of language screening in children between 18 and 35 months. She reported that this would help to identify children who have autism, intellectual disabilities, hearing impairment, receptive language delay, or who are at greater risk of later

language delay due to low socio-economic status. When language difficulties are the primary concern, the child is said to have a specific language impairment (SLI), which is characterized by the presence of language difficulties in the absence of hearing impairment, autism, intellectual disabilities, or other health-related problems impacting language development. For very young children, at around age 2 years, the term "late-talker" is often used to identify those children with low expressive language skills, with or without low receptive skills.

Many of these late-talking children who are delayed in their expressive language skills will only catch up with their peers in the following years (Law, Boyle, Harris, Harkness, & Nye, 1998; Rescorla, 2011). However, increasing evidence showed that this recovery is partly deceptive, since their language skills are significantly lower than children without a history of being a late-talker (Rescorla, 2011). The decision to recommend intervention or not is thus not an easy task. A speech-language pathologist should consider both the long-term and the short-term benefits of the intervention, which includes the child's emotional and behavioral development (Olswang, Rodriguez, Timler, 1998). In any case, providing language intervention to young children, even if they would have naturally reached the normal range, would not be detrimental to them. Furthermore, early intervention has been shown to be effective at improving speech and language difficulties (for a review, see Wallace et al., 2015).

For children still exhibiting mild to moderate language impairment at the age of 5 years, numerous related long-term negative effects have been documented, such as psychosocial problems and psychiatric disorders, or lower school success and educational attainment (e.g. Beitchman & Brownlie, 2005-2010; for a review, see Law et al., 1998). Recent research reached similar conclusions. Notably, Yew and O'Kearney (2013) found in their meta-analysis that adolescents diagnosed as having SLI between 3 years and 8 years 8 months were twice as likely as TD adolescents to have emotional problems and more than twice as likely to have behavioural

problems. Similarly, Conti-Ramsden, Mok, Pickles and Durkin (2013) concluded that adolescents with SLI are twelve times more likely then TD adolescents to self-report important difficulties with peer relationships, as well as more behavioural difficulties and emotional symptoms.

Providing early intervention to a child with low language skills is crucial to decrease the likelihood of long-term negative effects and to maximize the child's potential; it is as important as providing intervention for older children with persistent difficulties. To increase early referral for speech and language services, appropriate screening tools should be available to guide physicians in their recommendations. However, there currently are few validated tests to assess language skills in Canadian French (Gaul Bouchard, Fitzpatrick, & Olds, 2009), a fact already pointed out in the 1990s (Gregoire, 1993). This situation could lead to under-identification of children with language impairment (Webster, Majnemer, Platt, & Shevell, 2004).

Any tests should demonstrate very good reliability (the test's results would be the same even if the testing situation is different) and validity (the test actually evaluates what it supposed to). Furthermore, a diagnostic test should possess excellent sensitivity (the ability to identify correctly children with SLI as having SLI) and specificity (the ability to correctly identify TD children as not having SLI). Plante and Vance (1994) suggested values of at least over 80% (preferably over 90%) of correct identification. Taken together, sensitivity and specificity inform about the test's diagnostic accuracy, or discriminative value. A systematic review of language screening tools filled out by parents found a median sensitivity of 81%, but could be as high as 94%, whereas the median value for specificity was 87%, with a maximum value of 96% (Wallace et al., 2015). Before using any tests, these sensitivity and specificity values should be available for the population with which the test will be used. Indeed, it has been shown that four American screening tests were not equally good at identifying 54-month-old Canadian children with a receptive or an expressive language delay. In some cases, the diagnostic accuracy of the test for

the Canadian children could be improved by selecting new cut-off scores (Frisk et al., 2009). The authors hypothesized that this difference stem from differences between the two countries in terms of demographic characteristics and educational systems.

In addition to these elements, the predictive validity of a test should be considered, as it is informative of the relation between the current performance level and the performance the child should eventually reach when he or she is older. Many researchers investigated the predictive value of the children's vocabulary and early grammar skills at age two years, two criteria traditionally used to identify young children as being late-talkers. In an extensive literature review on late-talkers, Rescorla (2011) concluded that there was accumulating evidence showing that many children identified with language impairment at age 5 years were not late-talkers at age 2 years. A similar conclusion was reached by Poll and Miller (2013), who used a retrospective study design to investigate the predictive value of early delay in expressive vocabulary and two-word combination. They found that 36% of children with low language skills and 18% of TD children at age 8 years had delayed expressive vocabulary at age 2 years. In contrast, 23% of children with low language and 8% of TD children at age 8 years did not produce two-word combination at age 2 years. Using either early measure would thus not have identified the majority of children still exhibit language difficulties in the early school years. This indicates that a test should include additional facets of language development in order to have a good predictive value for SLI. A vocabulary test alone has limited value as a diagnostic test for SLI; in fact, vocabulary size does not correctly identify many preschool children with SLI. Indeed, Spaulding, Hosmer, and Schechtman (2013) reported that children with SLI scored on average only -0.4 standard deviations below their TD peers, well within the normal limit range, on the latest versions of the Peabody Picture Vocabulary Tests (PPVT-III: Dunn & Dunn, 1997; PPVT-IV: Dunn & Dunn, 2007). These tests are the most frequently used for receptive vocabulary, but their

sensitivity to SLI is .80 for both versions, with a specificity of only .75 for the PPVT-III and .70 for the PPVT-IV. Even though the assessment of vocabulary skills should be part of a child's language evaluation, judging if the child has SLI or not should not rely heavily on the child's performance on that test alone.

The sensitivity and specificity values are used to determine the confidence we can have that all children with (and without) the impairment have been correctly classified as having (or not) the impairment. To determine the confidence that can be placed on the result of a given child, the positive (and negative) predictive values of a test can be used. These measure the percentage of children with (or without) the impairment who failed (or passed) the test, from all the children who failed (or passed) it. Another possibility is to use the positive (or negative) likelihood ratios, which provides the likelihood that a child who has failed (or passed) the test really has the impairment; these ratios are calculated using the sensitivity and specificity values. Likelihood ratios are superior to the predictive values, because they are less affected by the impairment's prevalence, although their interpretation is less intuitive (Attia, 2003).

In order to address the need for a screening test for at risk of SLI in French-speaking preschoolers, the Milestones en français du Québec (MilBec: Paul & Elin Thordardottir, 2010), was developed. The MilBec is an adaptation of the Dutch parent questionnaire for children between 12 and 71 months by Luinge, Wit, Post and Goorhuis-Brouwer (2006). The parents are asked to indicate whether their child shows evidence of the specific language ability described in each item by answering 'oui' ('yes') or 'il me semble' ('I believe so'), which are scored as 1 point, or answering 'je ne crois pas' ('I don't think so'), or 'non' ('no'), scored as 0 points. The MilBec contains 39 items targeting both the receptive and expressive modalities in various language domains (vocabulary, syntax, narrative, language use, phonology, metalinguistic skills). By including different language domains, the test includes skills that emerge at different

developmental levels; for example, word combinations are present at least sometimes in 90% of 22-month-old Francophone children (Boudreault, Cabirol, Trudeau, Poulin-Dubois, & Sutton, 2007), whereas metalinguistic skills start to develop at 4 years of age (e.g. for French, see Lefebvre, Girard, Desrosiers, Trudeau, & Sutton, 2008). The detailed adaptation procedures as well as a preliminary validation from a longitudinal study were presented in Chapter 1. The results showed that the MilBec is developmentally sensitive, with a rapid increase in score up to age 36-42 months, with a ceiling effect occurring around 60 months. High levels of test-retest and inter-judge reliability and good concurrent validity were also reported.

In addition to the total score on the MilBec as a way to identify children with SLI, the usefulness of two specific items of the questionnaire is of interest. These items target potential clinical markers of SLI in French, namely the use of object clitics and use of subjects (Jakubowicz, Nash, Rigaut, & Gérard, 1998; Hamann et al., 2003). The first item targets the object clitics, which are pronouns replacing the direct object complement, usually a noun and its determiners. The object clitics are placed between the subject and the verb. For example, in "Je la veux" ('I it want'), "la" is the object clitic, replacing "la pomme" (the apple) in the typical sentence structure "Je veux la pomme" ('I want the apple'). The second item targets the use of subject, which is obligatory in most contexts in French and is an area of difficulty for children with SLI. Indeed, some researchers found that children with SLI performed significantly less well than age-matched children with TD, but similarly to TD younger children (Hamann et al., 2003; Jakubowicz et al., 1998; Thordardottir & Namazi, 2007).

The purposes of the present studies were to continue the examination of the psychometric characteristics of the MilBec and to examine its ability to screen children for SLI. In Study 1, preliminary norms were collected using a cross-sectional design with children between 12 and 71 months; the MilBec's reliability was also investigated. In Study 2, the concurrent validity and

diagnostic accuracy of the MilBec were investigated with children between 38 and 71 months, building on the normative data collected in Study 1. In Study 3, the diagnostic accuracy of the two items targeting potential clinical markers was investigated.

Study 1: Preliminary normative data

This first study was performed to collect preliminary normative data for monolingual French-speaking children between 12 and 71 months on the MilBec, in order to 1) further document the developmental sensitivity of the questionnaire, 2) to further investigate the MilBec's reliability, notably its internal consistency.

It was hypothesized that the raw scores on the MilBec would increase with age, and that the mean scores would be significantly higher for some of the older groups compared to the younger groups. In addition, based on the results presented in Chapter 1, it was hypothesized that the relationship between age and MilBec scores would be positive and linear with a decreased slope and lower variability of scores from the age of 36-42 months.

Methods

Participants

The parents of 85 monolingual French-speakers (44 boys, 41 girls) between 12 and 71 months participated in the study, with 17 participants per 12-month age group. Children should have been exposed to another language for less than 5 hours per week, should have normal development (no previous diagnosis or serious parental concerns about the child's development) and no parental concerns regarding hearing. One child born prematurely was excluded from the study, because of the specific risks to language development associated with prematurity that continues until the school years (Van Noort-van der Spek, Franken, & Weisglas-Kuperus, 2012). Two additional children were excluded because of serious parental concerns regarding language development mentioned in the demographic questionnaire. Socio-economic status (SES) was

assessed by self-reported maternal education. All subjects were living in the Province of Quebec, mainly in the Greater-Montreal area (n = 69). The data of 10 participants were obtained from the longitudinal study presented in Chapter 1. When more than one MilBec score was available for a child, only one was included in the cross-sectional normative data: the first one was used for eight participants from the longitudinal study; for the other two participants, the first one placing the child in the 2-year-old group, because this group had the lowest number of participants.

Background variables. The demographic questionnaire used with the participants from the longitudinal study presented in Chapter 1 was not the same as the one used in the current study. In the longitudinal study, the question on maternal education level asked the participants to select one of four possible levels of education: no high school diploma, high school diploma, CEGEP and university. As the demographic questionnaire used in the current study asked for maternal education in years, the maternal education of the participants from the longitudinal study were converted the following way: high school diploma was considered as 11 years of education, CEGEP as 13 years of education, and university as 16 years of education. For each of the 5 age groups, the average age and maternal education was calculated, as well as the standard deviation and range (see Table 1). One-factor ANOVA showed that the 5 age groups of children did not differ significantly on maternal education ($F(4.79) = .297$, $p = .879$). One-factor ANOVA confirmed that the 5 age groups of children differed significantly on age in months ($F(4.80) = 447{,}215$, $p = .000$), with post-hoc Tukey showing a statistically significant group difference between all groups (all $p = .000$).

[Insert Table 1 about here]

Procedures

The parents were invited to participate in a study on the validation of a parent questionnaire on language development via daycares, sports centers, school and public billboards

using e-mail, posters, or billboards posting. All parents signed the consent form of the project, which was approved by the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS). Once they confirmed that they wished to participate, parents were given the option to fill out the MilBec and the demographic questionnaire using an on-line survey created using LimeSurvey (n = 39) or using a paper-pencil version (n = 46). The demographic questionnaire included items regarding parental education level, medical and developmental history of the child, and other items regarding language development and language use at home. All statistical analyses were performed using the French version of IBM Statistic SPSS version 23.

Analyses and Results

Gender Effect

The language development of boys and girls may differ, with a slight advantage for girls in early language development (Wallentin, 2009). In a study of early development in French conducted in Quebec, it has also been shown that girls had better expressive skills than boys, particularly between 17 months and 28 months, although the effect size associated with this difference was small (Bouchard, Trudeau, Sutton, Boudreault, & Deneault 2009). The appropriateness of analyzing the data across genders thus had to be verified. To determine if there was a gender effect on MilBec scores, the five 12-months age groups had to be combined to have a minimum of 10 participants in each subgroup. The children were thus divided in three age groups (n = 27 for the 12-31 months, n = 32 for the 32-51 months, and n = 26 for the 52-71 months) and further divided by gender. The number of participants per subgroup and descriptive statistics are presented in Table 2. The distribution of scores as a function of age per gender is presented in Figure 1. Visual inspection indicated that there was no obvious effect of gender on the distribution of MilBec scores. A two-way factorial ANOVA was performed, using age and

gender as factors. There was a statistically significant effect of age (F(2.79) = 146,066, p = .000), with post-hoc Tukey showing a statistically significant group difference between all age groups (p = .000 or p = .003). The effect of gender was nonsignificant (p = .157) and there was no significant interaction effect (p = .375). These results support the adequacy of combining the results across genders in the normative data. The gender variable was not be further considered in the subsequent analyses.

[Insert Table 2 about here]

Developmental Sensitivity

The distribution of MilBec scores is presented in Figure 2 using a boxplot. The median score increased with age, with a greater difference between the youngest groups than between the oldest groups. The largest within-group variability occurred at age 2 and decreased for the older groups. Similar results were found based on the mean standard deviation and range of score (see Table 3). To determine if the effect of age was significant, an ANOVA was performed. A significant Levene test (p = .000) indicated that the data violated the assumptions of homogeneity of variance. The Brown-Forsythe adjusted F-test was used, which showed a significant group difference (F(4,50,024) = 97,983, p = .000). A post hoc Tukey test indicated that Age 1 and Age 2 groups were significantly different from all the other groups (all p = .000). Age 3 group was not significantly different from Age 4 group (p = .884), but was significantly different from Age 5 group (p = .050). Age 4 was not significantly different from Age 5 group (p = .347).

[Insert Table 3 about here]

[Insert Figure 1 about here]

To determine the type of relationship that existed between age and MilBec score, a scatterplot of MilBec scores as a function of age in months was produced (see Figure 2). The relationship between age and MilBec score was not linear. To fit a smooth curve without

specifying an a priori relationship between the variables, a local regression (loess) adjustment curve was performed on all the data (Jacoby, 2000), with the default Epanechnikov adjustment using 50% of the data points (see Figure 2). Visual inspection of the pattern of distribution indicated that the relationship between MilBec score and age followed two distinct linear slopes, one for the younger children and one for the older children, with a relatively short period of transition around 40 months. The strength of the relationship was explored using a Pearson correlation for each age subgroup separately. For the 12- to 39-month group (n = 42; mean age = 26.5 (standard deviation (SD) = 8.5); mean MilBec score = 22.6 (SD = 10.0)), the correlation was r = .921 (p = .000). For the 40- to 71-month group (n = 43; mean age = 56.0 (SD = 9.0); mean MilBec score = 36.4 (SD = 2.9)), the correlation was r = .598 (p = .000).

[Insert Figure 2 about here]

The children's variability in performance was illustrated using a scatterplot, where the two regression lines were indicated (one for the 12- to 39-month group and one for the 40- to 71-month group). To determine the extent to which the children were performing close to the expected level, additional lines adjusted to take into account the standard deviation were calculated for each group (see Figure 3). To determine the standard deviation to use for each of the two groups, the children were separated into 10 groups of 6-month intervals. The standard deviation was calculated for each of these ten groups, and then the average standard deviation was calculated (see Table 3). Older children were excluded from this calculation, because of the much smaller standard deviations for these subgroups. For the younger children, the standard deviation of the five groups between 12 and 41 months were used, leading to an average of 4.25. For the older children, the average standard deviation was based on the two groups of children between 42 and 53 months, for an average of 3.5. Visual inspection of Figure 3 shows that the majority of the children in both age groups are distributed around the regression line representing

the expected level, with some children under the -1 SD line, and very few under the -2 SD line. This pattern of distribution indicates that these lines could be useful in the selection of cut-off values for the identification of children with SLI.

[Insert Table 3 about here]

[Insert Figure 3 about here]

Reliability

To investigate the internal consistency of the MilBec, Cronbach's alpha (α) was calculated based on all 39 items, with a resulting α = .967; if any of these 39 items were to be removed, the new α varied between .968 and .965. The removal of any of the 39 items would thus not improve the internal consistency of the questionnaire. However, items 3 and 5 showed no variability, since all participants received a score of 1. Because it is proposed that the MilBec be used differently for younger and older children, and that the relationship between age and MilBec score is different for younger and for older children, Cronbach's alpha was also calculated for the younger and the older children separately. For children in the 12- to 39-month group (n = 42), α = .961 f, and for the 40- to 71-month group (n = 43) α = .763. When separated this way, the older group showed an additional 13 items with no variability.

Discussion for Study 1

An important goal of this first study was to confirm the MilBec's developmental sensitivity for children between 12 and 71 months. The results showed that most of the age groups were significantly different from each other, except for two comparisons: the group of 3-year-old was not different from the group of 4-year-old and the group of 4-year-old was not different from the group of 5-year-old. This finding supports the developmental sensitivity of the test, with older children receiving higher scores. Furthermore, there was a very strong correlation between age and MiBec score, particularly for the younger children. However, the rapid increase

in MilBec score for the younger children led to a high variation in mean scores between the groups with a 12-month range. Such a large variation between group means can limit the usefulness of the preliminary norms. Indeed, when a child's age is at the junction between two groups, a very different impression of the child's language skills will be arrived at depending on the comparison group used. For example, a score of 12 obtained by a child tested the day before his second birthday would be very close to the mean score of the 1-year-old group (z-score = -0.16), but if the same child is tested two days later, the same score would be very far from the mean of the 2-year-old group (z-score = -2.4). Such a large difference does not mirror the expected relative stability of language skills. Although there will always be children on the cut-off of an age group, if the difference between mean scores of two consecutive groups is smaller, the impact of this is reduced. For example, if the means and standard deviations of the 6-month intervals are used instead (see Table 3), the difference between the two groups for a score of 12 becomes smaller ($z = -0.96$ at 23 months and $z = -1.58$ at 24 months). To use more restricted age range, additional normative data must be collected. However, because the 12-month grouping of children over 40 months demonstrated a lower rate of score increase on the MilBec, it is believed that starting at that age, the 12-months age bands are appropriate.

The internal consistency of the MilBec, a measure of construct validity, was found to be adequate based on Cronbach's alpha ($\alpha = .967$). This analysis also pointed to two items that might be considered for removal in a revised version of the test, as all participants received a score of 1. However, it would first have to be established whether or not these items are indicative of language difficulties at some point between 12 and 71 months.

One element that should be considered in a follow-up study is the possible effect of schooling on test items. Indeed, some target metalinguistic skills that are taught during kindergarten. The fact that the average maternal education level was high limits the

generalizability of the results. Lastly, for the MilBec to be considered a valid screening tool of language, it is not sufficient for it to be developmentally sensitive. It is also important to determine that it is valid (i.e. that it is assessing language skills), and that it correctly identifies children with and without language impairment (i.e. that it has good discriminative values). These elements are investigated in the next study.

Study 2: Diagnostic Accuracy

The goal of this second study was to determine if the MilBec is a good screener for language impairment for children between 36 and 71 months by 1) providing evidence of concurrent validity, and 2) documenting the MilBec's sensitivity and specificity for the detection of language impairment. The measures used in this study can be separated into three types. The first one included measures targeting a specific language skill, more specifically a measure of expressive vocabulary, receptive vocabulary, comprehension of morphosyntax, spatial and pre-arithmetic concepts comprehension. The second type included a sentence imitation task, selected because it was shown to have one of the best sensitivity and specificity values for identification of 5-year-old children with SLI (Elin Thordardottir et al, 2011). The last type was another French parent questionnaire, which was longer and validated in Europe. Because the MilBec targets various language domains, it was hypothesized that MilBec scores would have at least a moderately strong positive correlation with language measures targeting specific language domains, and a strong positive correlation with the other parent questionnaire.

Methods

Participants

Two groups of 15 French-speaking monolingual children between 36 and 71 months (inclusive) matched group-wise on age and gender participated in the study. The first group included children with specific language impairment (SLI). To be included in this group the child

had to be currently receiving or awaiting a language therapy in a rehabilitation center for a moderate to severe language delay or language impairment. Potential participants for this group were identified based on the intake information at the Institut Raymond-Dewar (Montreal, Quebec, Canada). If a parent approached for Study 1 indicated that his or her child had language impairment, the parent was invited to participate in Study 2. If the child fulfilled the inclusion and exclusion criteria, he or she was included in the SLI group. All children with SLI recruited for the study had thus been previously assessed by a certified speech-language pathologist, as this was a requirement to be included in the waiting list of a rehabilitation center. Children with pervasive developmental disorder or with a semantic-pragmatic profile of impairment were excluded. Children with a diagnosis or suspicion of childhood apraxia of speech, which is primarily a motor impairment, were included only if they also showed evidence of language impairment. Three additional children with SLI were referred, but were excluded: one because of severe stuttering, one because language had normalized, and one because the child also had a diagnosis of autistic spectrum disorder.

Participants in the second group were children with typical development (TD) recruited from Study 1. Most of them were invited to participate to Study 2 at the same time as Study 1; some were recruited later, if they had agreed to be contacted for participating in another study on the MilBec. The parents were invited to participate in Study 2 if their child had the age or the gender required to obtain a group-wise match with the SLI group. The judgment of normal language development of all children in this group was based on the lack of parental concerns, and later confirmation by test results and informal evaluation. Three children with TD were excluded from the study: two because they were adopted from a non-francophone country, one because of a diagnosis of anxiety disorder.

All children had normal hearing and normal nonverbal intelligence. For the 4- and 5-year-

old participants who understood and accepted to do the task, a pure tone hearing screening conducted at 4 frequencies (500Hz, 1000Hz, 2000Hz, 4000Hz) was performed in each ear with a portable audiometer, at an intensity level of 10 dB HL (30dB HL at 500Hz because of ambient noise). For the untested children, normal hearing was assumed based on parent report; a few parents even reported that their child had passed a hearing screening as an infant. A few children in the SLI group had also been evaluated by an audiologist in the past, with the conclusion of a normal hearing. Nonverbal cognition was assessed using the Brief-IQ of the Leiter International Performance Scale-Revised (Leiter-R: Roid & Miller, 1997) to rule out intellectual disabilities. To be included in the study, a standard score of 70 or higher was required. The Socio-Economic status (SES) of the children, as assessed by the maternal education, was determined based on parental report. All participants were living in the Greater-Montreal area.

Background variables. The participants' background characteristics are presented in Table 3. The difference between the groups in terms of the number of boys was not significant ($\chi 2(1) = .556$, $p = .456$). For the background variables (age in months, maternal education and nonverbal IQ), the Levene test for homogeneity of variances was not significant (respectively $p = .295$, $p = .729$, $p = .161$). Independent sample t-test for groups with equal variance was selected, and no significant group difference was found for age ($p = .469$), but was significant for maternal education ($t(27) = -4{,}293$, $p = .000$). The average maternal education for TD children was higher, at 16.9(3.1), compared to the 12.2(2.7) for children with SLI. It was also significant for the Brief IQ scores ($t(26) = -4{,}759$, $p = .000$), with a higher average IQ for children TD, at 114.1 (11.2) compared to 97.0 (7.5) for children with SLI. Despite this difference, the average score for the group of children with SLI was not statistically different from the expected average score of 100 ($t(13) = -1{,}494$, $p = .159$). This group difference between children with and without SLI was congruent with Gallinat and Spaulding's (2014) meta-analysis, where it was found that a lower

nonverbal IQ score should be expected for children with SLI.

[Insert Table 3 about here]

Procedures

For children in the SLI group, the project was presented to parents of children fulfilling the inclusion and exclusion criteria during a therapy session at the Institut Raymond-Dewar. Once parental consent was obtained, an appointment for testing was given. All parents signed the consent form of the project, which was approved by the ethics committee of the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS).

The children were individually tested by the first author, a certified speech-language pathologist from the Ordre des Orthophonistes et Audiologistes du Québec (OOAQ) and native speaker of Quebec French. The children were tested in their home, at a home chosen by the parent (e.g. babysitter home), or in a therapy room at the Institut Raymond-Dewar. Most testing was completed in a single visit, with breaks whenever the child needed one, but some were divided into two visits by parents' request or when required due to lack of cooperation from the child. The evaluations were video-recorded using a Canon FS100 for scoring and reliability purposes. All statistical analyses were performed using the French version of IBM Statistic SPSS version 23.

In addition to the MilBec and the same demographic questionnaire used in Study 1, the following measures were administered to each child. The Échelle de Vocabulaire en Image Peabody (EVIP: Dunn, Theriault-Whalen, & Dunn, 1993) evaluates receptive vocabulary; it is a Canadian French adaptation of the Peabody Picture Vocabulary Test-Revised (PPVT-R: Dunn & Dunn, 1981). For the between-group comparisons, the standard score obtained from the test's manual was used (EVIPss) to take into consideration the expected performance at the child's age. However, it was decided to use the raw scores (EVIPraw) for the correlational analysis, since raw

scores were used with all the other measures.

Four subtests of the Nouvelles Épreuves pour l'Examen du Langage (NEEL : Chevrie-Muller & Plaza, 2001) were administered : Subtest 7 (morphosyntax), with the reported measure being the total score of section a and section b. This task evaluates the comprehension of morphemes (e.g. gender of the subject pronoun) and syntax (e.g. reversible sentences). Subtest 8 (positional concepts) evaluates the comprehension of spatial terms (e.g. "sur" (under), "à côté" (beside)). Subtest 8 (arithmetic) evaluates pre-arithmetic concepts (e.g. things counting, ordering terms). Subtest 11 (expressive vocabulary) had two parts. The first part provides a long and a short version, the short version being suggested for younger children or those with severe difficulties. The long version was tried for all participants, but for some children only the short version was completed. The part 2 of the test was administered to all children. The reported results are the total raw score of the short version of part 1 added to the raw score of part 2.

A sentence repetition task, adapted from the Sentence Imitation in Context subtest of the Clinical Evaluation of Language Fundamentals – Preschool (Wiig, Secord, & Semel, 1992) by Royle and Elin Thordardottir (2003), was administered. The scoring procedures followed the original English guidelines, with a discontinuation rule after five consecutive scores of 0, and each sentence received a score of 0-3 depending on the number of errors. The procedure of calculating the percentage of correct word used in Elin Thordardottir et al. (2010; 2011) with 5-year-old children, which required the administration of all test items, was not selected because some children with SLI refused to complete the task, but they had reached the discontinuation rule allowing the original procedures to be followed.

A spontaneous language sample was collected during a 10-15 minute play session using a standard set of toys including a Little People small castle, figurines and accessories. This data was collected to support clinical judgment of language status, and was not analyzed for the

present study.

A parent questionnaire to screen for language impairment, namely the Inventaire du Développemente de l'Enfant – forme langage (IDE: Duyme, Capron, & Zorman, 2010) was completed during the evaluation or less than three weeks before the testing session. The IDE is a parent questionnaire of 100 items targeting mainly vocabulary and syntax, equally divided into an expression subscale (IDE-lex) and a comprehension subscale (IDE-lco). The items are divided into 5 levels, each level presenting items from the IDE-lco subscale and the IDE-lex subscale. Scoring takes 2-4 minutes, including the calculation of a total score for each subscale and the calculation of the limits of the category "at risk" and the category "at high risk", which are based on the chronological age. The order of presentation of the MilBec and IDE was counterbalanced across subjects. For both the MilBec and IDE, if an item was not answered it was scored 0. In one case, a whole section of the IDE was not completed by the parents. For this participant, the IDE was considered missing data.

Analyses and Results

Language Tests and Screening tools

For each of the nine language measures, the group means, standard deviation and range of scores are presented in Table 6. Most tests showed a ceiling effect for TD children, except three tests, namely the EVIP, the sentence repetition, and the NEEL11 (expressive vocabulary). Three tests showed a floor effect for children with SLI, namely sentence repetition, NEEL7 (morphosyntax), and NEEL8 (arithmetic). These last two subtests thus had both a floor effect for children with SLI and a ceiling effect for TD children. Except for the EVIP and the measures with a floor effect in the SLI group, there was greater score variability within the group of children with SLI, with a standard deviation that could be twice as large as the standard deviation found for the group of TD children.

[Insert Table 6 about here]

Analysis of group differences was performed using a series of nine independent-sample t-tests. A Bonferroni correction was applied to reduce the Type 1 error related to multiple comparisons, leading to a critical alpha level set at .0055. The mean and standard deviation (SD) obtained by each group on all the tests, as well as the *p*-value and effect size based on Cohen's *d* are presented in Table 6. There was a significant result of Levene's test of equality of variance for the IDElex, sentence repetition, NEEL8 (positional concepts) and NEEL11 (expressive vocabulary). For these variables, the t-test that did not assume equal variance was reported. Significant group differences using the stringent *p*-value of .0055 were obtained for all nine language measures tested (with *p*-values of .000 or .001), except for the expressive vocabulary which was close to significant ($p = .006$), and would have been if the Bonferroni correction had not been applied. For this reason, the effect size was also indicated for this test, despite the difference between groups was not strictly speaking significant. All effect sizes were very large, being much greater than .8 (Cohen, 1992), with the largest effect size obtained by the IDE-lex with $d = 5.1$, followed by the NEEL8 (positional concepts) with $d = 4.9$, and the MilBec with $d = 3.4$; all the other measures had a large effect size, with Cohen's *d* between 1.2 and 2.1.

The MilBec's Concurrent Validity

To examine the concurrent validity of the MilBec, the strength of the relationship between MilBec scores and the raw score of some of the language tests and screening tools was calculated using Pearson's correlations. To reduce the risk of Type 1 error, a Bonferroni correction must be applied, which required to limit the number of comparisons to conserve statistical power. Seven measures, for a total of 21 correlations, were used, leading to a stringent *p*-value of .0024 for the tests to be considered significant; a *p*-value of .002 was considered significant. The two tests showing both a floor effect for children with SLI and a ceiling effect for children with TD were

excluded from the analyses. These tests, namely the NEEL7 (morphosyntax) and the NEEL8 (arithmetic), would not have been informative for either the SLI group nor the TD group. The remaining seven measures were included in the correlational analysis (see Table 7), which was first performed for all children and then performed separately for each group.

When all children were considered simultaneously, MilBec scores were significantly correlated with IDE-lex (r = .921), IDE-lco (r = .783), EVIP raw score (r = .545), sentence repetition (r = .797) and NEEL11 (expressive vocabulary) (r = .591). The correlation with NEEL8 (positional concepts) (r = .492) would have been significant if the Bonferroni correction had not been applied (p = .008). For children with SLI, the MilBec scores were significantly correlated only with IDE-lex (Pearson's r = .756), but would have been significant with IDE-lco if the Bonferroni correction had not been applied (p = .031). For TD children, the MilBec scores were significantly correlated with IDE-lex (r = .741), IDE-lco (r = .884), and EVIP raw score (r = .731). The correlation with sentence repetition (r = .715) would have been significant if the Bonferroni correction had not been applied (p = .009). The lack of correlation with NEEL8 (positional concepts) was expected for this group, given the ceiling effect. As predicted, the correlations between the MilBec and the two subscales of the IDE, the other parent-questionnaire, were stronger than the correlation with the other language tests, although the sentence repetition task also showed a very high correlation.

[Insert Table 7 about here]

The MilBec's Diagnostic Accuracy

MilBec scores as a function of age for children with and without SLI are presented in Figure 4, where it can be observed that the raw scores of children with SLI were lower than those of children with TD. To determine the MilBec's diagnostic accuracy, which informs on how well the individual children were correctly identified as having SLI or TD, cut-off scores had to be

determined. Two analyses were performed to determine the best cut-off values. The first one was based on the regression lines and average standard deviation for each of the two age groups presented in Study 1. The cut-off was set at -1 SD, as the scores of majority of the TD children from the normative study were above that line. As presented in Figure 4, the 15 children with SLI were under the cut-off lines, correctly identifying them as having language impairment, which corresponds to a sensitivity of 100%. For the TD children, 14 out of 15 scores were above the cut-off lines, thus correctly identifying them as TD, which corresponds to a specificity of 93%. These sensitivity and specificity values can be used to calculate the likelihood ratios. A negative likelihood ratio indicates the probability that a child receiving a "pass" score does indeed have typical development, and is calculated using (1-sensitivity)/specificity; a negative likelihood inferior to 0.1 is a strong indicator that the child does not have the disability (Jaeschke et al. 1994). For the MilBec, the negative likelihood was 0, which indicated that if a child had a score above the cut-off, it was a very strong evidence to rule out a moderate-to-severe SLI. The positive likelihood ratio indicates the probability of a child who receives a "fail" score was indeed having an impairment. It is calculated using sensitivity/(1-specificity), and a result superior to 10 is a strong indicator that the child really has the disability (Jaeschke, Guyatt, & Sackett, 1994). For the MilBec, the positive likelihood was 14.28, which indicated that a score under the cut-off value on the MilBec was a very strong indicator that the child indeed had a moderate-to-severe SLI.

[Insert Figure 3 about here]

The second manner to determine the best cut-off value consisted of using a Receiver Operating Characteristic (ROC) curve analysis, taking into account all children regardless of age. This analysis allows to determine the degree of overlap between the distribution of two populations, in this case children with SLI and those without SLI, calculated at different cut-off

scores (Hajian-Tilaki, 2013). The cut-off leading to the highest level of sensitivity, while preserving the best possible specificity, was a score of 31.5. The corresponding sensitivity was 100% and specificity 93%; the area under the curve estimated, which is a combined measure of sensitivity and specificity at every possible cut-off scores (Hajian-Tilaki, 2013), was of .973. Both procedures thus led to a similar outcome.

Discussion for Study 2

This Study 2 documented the MilBec's concurrent validity and diagnostic accuracy. The MilBec strongly correlated with other language measures, particularly when the measure was another parent-questionnaire, which supports the concurrent validity of the MilBec. The higher correlation between the MilBec and the IDE might rise from two sources: 1) they are both parent-questionnaire, whereas the other tests are formal tests administered to the child by an unfamiliar adult, and 2) the IDE targets two language domains, whereas formal tests targets only one (except the sentence repetition task, as will be further discussed below), making it more similar to the MilBec which targets various language domains. A possible explanation for the lack of statistical significance when using the Bonferroni correction between the MilBec and the NEEL8-positional concepts task might be due to the very specific kind of knowledge assessed with this task, namely the comprehension of spatial terms. When all participants were taken together, there was a larger range of test scores, which can explain why stronger correlations were found between the MilBec and other language tests.

When the two groups were considered separately, the correlation patterns were not the same for children with and without SLI. Notably, the correlation with the EVIP, a receptive vocabulary test, went from r = .545 (p = .002) for all children regardless of language groups to r = .731 (p = .002) for TD children only. In typically developing children, vocabulary is often used to assess overall receptive language skills (e.g. study 2 in Petersen et al., 2013), and even as a quick

indicator of verbal intellectual quotient (Campbell, Brown, Cavanagh, Vess, & Segall, 2008); for the MilBec to be highly correlated with receptive vocabulary thus supports its use as a screener for language difficulties. For children with SLI, however, language development is not necessarily synchronized the same way as what is found in TD children, with some linguistic elements being more delayed than others (Leonard, 2014). Furthermore, it is not always the same element that is particularly delayed in all children, with findings repeatedly showing that vocabulary skills are not delayed in every child with SLI (e.g. Gray, Plante, Vance, & Henrichsen, 1999; Spaulding, Hosmer, & Schechtman, 2013). Consequently, vocabulary skills should not be used to estimate overall language abilities in that population; the lack of significant correlation between vocabulary skills and MilBec score for children with SLI is thus not a sign that the MilBec is not an adequate screener. Finally, the correlations between the MilBec and sentence repetition was of the same magnitude as the correlations with the other parent questionnaire, and contrasted with the less strong correlations with other language tests. This higher correlation might be caused by the fact that sentence repetition can be viewed as a measure of overall language skills, since it requires adequate speech perception and production, vocabulary and grammatical knowledge (Klem et al., 2015), whereas the other language tests target a specific language skill.

A different pattern of correlations for TD children and those with language difficulties were observed. Notably, whereas the correlation between the MilBec and receptive vocabulary was high for TD children, it was no longer significant for children with SLI. This finding is coherent with the earlier review of vocabulary skills of children with SLI, where it was pointed out that their vocabulary was not always below normal limits. Differences between the abilities of children with and without language difficulties have been found in other studies as well. For example, it was found in a longitudinal study by Moyle, Ellis Weismer, Evans, and Lindstrom

(2007) that TD children between 2;6 (years;months) and 5;6, demonstrated both lexical and syntactic bootstrapping, whereas late-talking children showed mainly evidence of lexical bootstrapping. Lexical bootstrapping occurs when the lexical development contributes to the syntactic development, with a certain number of items in the lexicon being required for extracting rules and regularities necessary for syntactic productivity. In contrast, syntactic bootstrapping occurs when syntactic knowledge is used to help learn new words. These differences in correlation patterns between the MilBec and other measures of language might also be partially explained by the great variability in language profiles, not only vocabulary, of children with SLI, as well as by differences in severity levels.

The capacity of the different tests to discriminate between TD children or children with SLI can be assessed using sensitivity and specificity measures. Using the extreme groups of TD children and children with a moderate to severe SLI, the MilBec was found to have a sensitivity of 100% and a specificity of 93%. This makes the MilBec a screening tool with a good discriminant validity based on Plante and Vance's (1994) criteria. Using the effect size based on Cohens' *d* (see Table 4), a measure that indicates to which extent the two groups' means are far from one another, the MilBec had a very large effect size with $d = 3.4$, the third-highest value out of the nine measures used in the study. For children between 36 and 71 months, the MilBec was found to be a better screener than many direct tests, including tests of vocabulary skills. These results support the use of the MilBec to screen for SLI in preschool children. To have a quick parent questionnaire whose results can be interpreted with confidence to screen for language impairment would help to save time and resources.

Limitations of the study include the use of intervention status as the gold standard, without considering the performance on a standardized test battery. Indeed, the lack of validated tests in Canadian French makes the identification of children with language impairment more

reliant on clinical judgment. As more tests will be developed and validated, it will become feasible to use a gold standard that includes both clinical judgment and a rigorous test protocol. The low number of subjects is also a limitation of this study.

Study 3: Diagnostic Accuracy of Two Potential Clinical Markers

Two items included in the MilBec have been identified in the literature as potential clinical markers of SLI in French: subject use and object clitic use. The age at which these items are acquired by TD monolingual children, as reported by parents in the MilBec, was determined. Then, their diagnostic accuracy in identifying monolingual children with SLI was investigated. This analysis determined the value of these elements in identifying specific language impairment.

Methods

Participants and Procedures

There were two groups in this study. The participants from the first group were the same 85 typically developing (TD) monolingual French-speakers between 12 and 71 months detailed in Study 1. The second group included the 15 participants with specific language impairment (SLI) between 36 and 71 months detailed in Study 2.

The scores of each participant for the subject use (item 18: does the child use subject) and object clitic use (item 28: does the child replace a word to identify an object by a pronoun) were retrieved from the MilBec. The parent had to select one of four answer choices: “oui” (‘yes’) or “il me semble” (‘I believe so’), which were both scored as 1 point, or "je ne crois pas” (‘I don’t think so’) or “non” (‘no’), both scored as 0 points. To determine at what age these items were reported to be used by typically developing children, a scatterplot of the children’s score as a function of age in months was produced (see Figure 5). Visual inspection of Figure 5 shows three distinct periods of variability in the reported usage across children: the first period occurred when there was no participant reported as using the element, the second period when some children

were reported using it, and the third period when all children were reported as using it. The age at which there was variability in the parent's report concerning the use of the item by their TD children was determined for each item in a two-step process. First, the boundary between the first and second period was set, by identifying the age at which the first child was reported as using the item. Then, the boundary between the second and the third period was set at the age at which all TD children, excluding outliers, were reported as using the item. Using this method, the periods' boundaries for the subject use occurred at 19 and 27 months. The periods' boundaries for the object clitic use occurred at 26 and 31 months.

[Insert Figure 5 about here]

To determine the usefulness of the specific items in identifying SLI in children based on the information from the MilBec, the pattern of acquisition should be compared between the groups of children with and without SLI. For each item and for each group, the number of children using the subject and use the object clitic were determined (see Table 8), for each of the three periods based on Figure 5 previously detailed. The numbers of children in the third period, which corresponds to the period during which it is expected that parents will report their child as using the item, would then be used to calculate the sensitivity and specificity value of each item. In turn, the sensitivity and specificity values would be used to calculate the likelihood ratios.

[Insert Table 8 about here]

Analyses and Results

Subject use. None of the 11 participants with typical language development under 19 months were reported as using the subject by their parents (see Table 7). In the second period, between the age of 19 and 27 months, there were 8 out of 13 participants reported as using them. Finally, in the third period occurring between 27 and 71 months, all of the 61 participants were reported as using subjects. In contrast, from the 15 children between 36 and 71 months in the SLI

group, 12 were also reported using subject. The corresponding the sensitivity was only 20%, with only 3 of 15 children with SLI receiving a score of 0 for this item; the specificity was 100%, since all 61 TD children received a score of 1 for this item. In turn, these values can be used to calculate the likelihood ratio indicating the strength of the evidence to rule in or rule out SLI for children over 27 months. The negative likelihood was 0.8, which indicates that when the child used the element, it was only a weak evidence toward ruling out a moderate-to-severe SLI. On the other hand, the positive likelihood was excellent (.20/0), which indicated that receiving a score of "0" for the item was a strong indicator of language impairment.

Object clitic use. Concerning object clitic use, the first period included 23 TD children under 26 months, all of them being reported by their parents as not using object clitics (see Table 7). For the four children between 26 and 31 months, which corresponds to the second period, two of them are reported using them. In the third period occurring between 31 months and 71 months, there were 56 of 58 children reported using them. During this period, 7 out of 15 children between 36 and 71 months with SLI were reported using object clitics. The sensitivity was 53% for this item, with 8 of 15 children with SLI receiving a score of 0; the specificity was 97%, with 56 of 58 TD children having received a score of 1. Using these values, the likelihood ratios can be calculated for children over 31 months: the negative likelihood was 0.5, which indicates that if the child used the element it was not a strong evidence toward ruling out a moderate-to-severe SLI, whereas the positive likelihood, with a value of 13.25, was a strong indicator of the presence of SLI.

Discussion for Study 3

Subject use. The results from Figure 3 showed that the period of variability in the reported use of subjects is between 19 and 27 months. This is consistent with the findings of Parisse and Le Normand (2001), which showed that Francophone children used verbs

accompanied by a subject a little more than 60% of the time between 24 and 27 months of age. It also concurred with findings from Elin Thordardottir (2005), where it is reported that in the spontaneous speech of 24-month-old Francophone children, the percentage of correct use of personal pronouns was very high with 74% of correct use for “je” (‘I’), and 90% for “il/elle” (“he/she”), even though they are one of the three main types of omitted words.

However, the results highlighted the limited value of subject use as a clinical marker of SLI for children over 27 months. Indeed, although it provided an excellent specificity value of 100%, its sensitivity was extremely low at 20%, which led to a poor negative likelihood ratio of 0.5. These results concurred with previous findings questioning the status of subject use as a clinical marker of SLI in French (Elin Thordardottir & Namazi, 2007), since the use of the subject did not allow the identification of the majority of children with SLI of the current study. Indeed, a clinical marker should identify all children currently having the condition up to the individual whose difficulties have resolved (Bishop, North, & Donlan, 1996). Nonetheless, a parental judgment of non-usage of the subject would be a strong indicator of SLI, as shown with the very high positive likelihood ratio. The lack of subject would be a strong indicator of immature sentence construction.

It is to be noted that in the MilBec, the parents were asked if their child was using the subject at all, in contrast to most research investigating clinical markers where the percentage of use was analyzed. It is still unknown what is the level of use required for parents to judge that their child uses a particular element; it is possible that the parent correctly reported that the child was using subjects, but a more detailed analysis would uncover a low percentage of use.

Object clitic use. The period in which some parents reported that their child used object clitics was between 26 and 31 months. This finding indicates a slightly earlier age of use than the reported age of emergence, which was reported to be between 30 and 36 months (e.g. Hamann &

Belletti, 2006). However, this difference may stem from two elements that may have affected the reported age of acquisition across studies: the type of elicitation procedures and the definition of emergence used. Nonetheless, in the current study some children over 48 months were not reported by their parents as using object clitics, suggesting that they may take longer to acquire for some TD children. It is possible that these children have skills at the lower end of normal development, making them more similar to children with SLI than the average child. Alternatively, it could be that this item was harder for parents to judge. The greater variability in the development of TD children may explain the lower specificity value for this item, which is 97% for children over 31 months, compared to a specificity of 100% for subject use.

It is to be pointed out that the age at which use of object clitics is first identified occurs when the children are all reported to be using subject. This indicates that, as one might expect, an element indicative of a more complex syntax appeared in a child's language once a more basic element of an utterance was also included. Indeed, the subject is an obligatory element in an utterance, whereas object clitics are a more complex syntactic way to include the direct object. This difference in complexity might be the reason why the item on the use of object clitics was less often reported as mastered for children with SLI over 31 months than was subject use, which had a sensitivity value of 53%. Although this value is better than the sensitivity value of 20% found for subject use, it is still much below the proposed 80% value to be considered as a diagnostic criterion with a "fair" value (Plante & Vance, 1994). Many children with SLI may have reached a language development stage where they are including subjects, but some had not yet reached the more complex stage of using object clitics. The fact that some TD children over 31 months were not reported using object clitics also supports its greater complexity. As was the case with subject use, the results on object clitic use concurred with the literature refuting it as a clinical marker of SLI (Elin Thordardottir & Namazi, 2007), as it did not allow the correct

identification of many children previously diagnosed with a moderate to severe SLI. Nonetheless, object clitic use was a strong indicator of the presence of SLI when a child receives a score of 0 on this item. Indeed, this item's negative likelihood was 13.25, which is above the criteria of a minimum value of 10 to be considered a strong evidence to confirm the presence of an impairment (Jaeschke et al. 1994).

The principal limitations of this study is using cut-off scores based on preliminary norms. However, before undertaking a long and costly process of collecting appropriate normative data, it was judged essential to first establish the test's potential using preliminary results. Also, using dichotomous data rather than more graded accuracy data may have provided a coarser view of the acquisition patterns. However, the coarser dichotomous data was preferred, because the MilBec's primary goal is to screen for language impairment and not provide a graded view of the items' acquisition profile. Another limitation is that some children with SLI were receiving language intervention while others were not, and such information was not taken into account. Intervention may have some impact on the children's results on their standardized testing, and the greater awareness of their child's language status and difficulties may have influenced how parents filled out the MilBec and IDE. A future study would have to be performed to determine if parents' answers on the MilBec change when their child starts intervention.

Discussion and Conclusion

The MilBec is a short parent questionnaire aimed at identifying language impairment. The results from the first two studies supported its use as a screening tool to identify children between 36 and 71 months who are at risk of having a moderate to severe language disorder. In the first study, the MilBec was shown to be developmentally sensitive and to have adequate construct validity. In Study 2, it was shown that the MilBec had adequate concurrent validity and diagnostic accuracy. Study 3 added to the body of research investigating potential clinical

markers of SLI, namely subject and object clitic use. An important clinical application of the reported findings on clinical markers of SLI is that the absence of subjects at 27 months and object clitics at 31 months use should be considered a sign of SLI, although their use does not indicate typical development.

Many measures assessing French-speaking children at age 5, when they are used with a cut-off score of -1 standard deviation, were found to have a sensitivity value below .55 and a specificity value above .80 (Elin Thordardottir et al., 2011). From the 12 measures reported in that study, six showed this pattern: mean length of utterance in words and morphemes, story grammar and first mentions from the Edmonton Narrative Norms Instruments (ENNI: Schneider, Dubé, & Hayward, 2005), rapid automatized naming of animals (RAN; Catts, 1993) and forward digit span (Celf-4 subtest; Semel, Wiig, & Secord, 2006). These results highlighted the fact that children with SLI may have very different underlying difficulties that lead to impaired language skills. It might be because the MilBec targets language skills from various language domains that it had a high diagnostic accuracy. Indeed, having low skills in any language domain included in the screener would lead to a lowered score, even if skills from the other language domains were preserved. This explanation is coherent with the results in Elin Thordardottir et al. (2011), where it was found that nonword repetition and sentence repetition were associated with good diagnostic accuracy, most likely because these tests require skills of many different types. For these tests, if the child's ability level on any one of the required skills was low, his or her test results would be affected by it.

Although the two potential clinical markers did not have a sufficient level of sensitivity, despite a good level of specificity, the sensitivity and specificity of the MilBec total score were adequate. This highlighted the adequacy of investigating elements from different language domains, and the usefulness of including elements from different developmental periods. Indeed,

the finding that a skill acquired earlier, namely subject use, was less sensitive to SLI than the later acquired object clitic use, pointed to the fact that language becomes more complex with age, even in children with SLI. Some children who were able to acquire an earlier developing item were still struggling with the later acquired element; it is possible that some children were not delayed in their acquisition of subject, but nonetheless experienced difficulties at acquiring object clitics. This would be coherent with the findings that early sign of delay in syntactic complexity, namely the emergence of two-word combinations, allowed the identification of only a minority of children who would demonstrate signs of language impairment at age 8 (Poll & Miller, 2013). To confirm the hypothesis that as the children's language becomes more complex, a gradually higher number of children would demonstrate difficulties acquiring language, a detailed analysis of response patterns on the MilBec in a longitudinal study would have to be performed. Finally, further investigation of the MilBec's diagnostic accuracy in a population sample is warranted, to collect norms with groups having a smaller age range and with participants that are representative of the general population.

Acknowledgments

We want to thank all the parents and children who participated in the study. The cooperation of the therapists who accepted that the project be presented to their patients' parents during their therapy session was also highly appreciated. A special thanks to Eve Julie Rioux and Martine Gendron for their assistance, notably for the recruitment and making arrangements for testing materials or testing rooms.

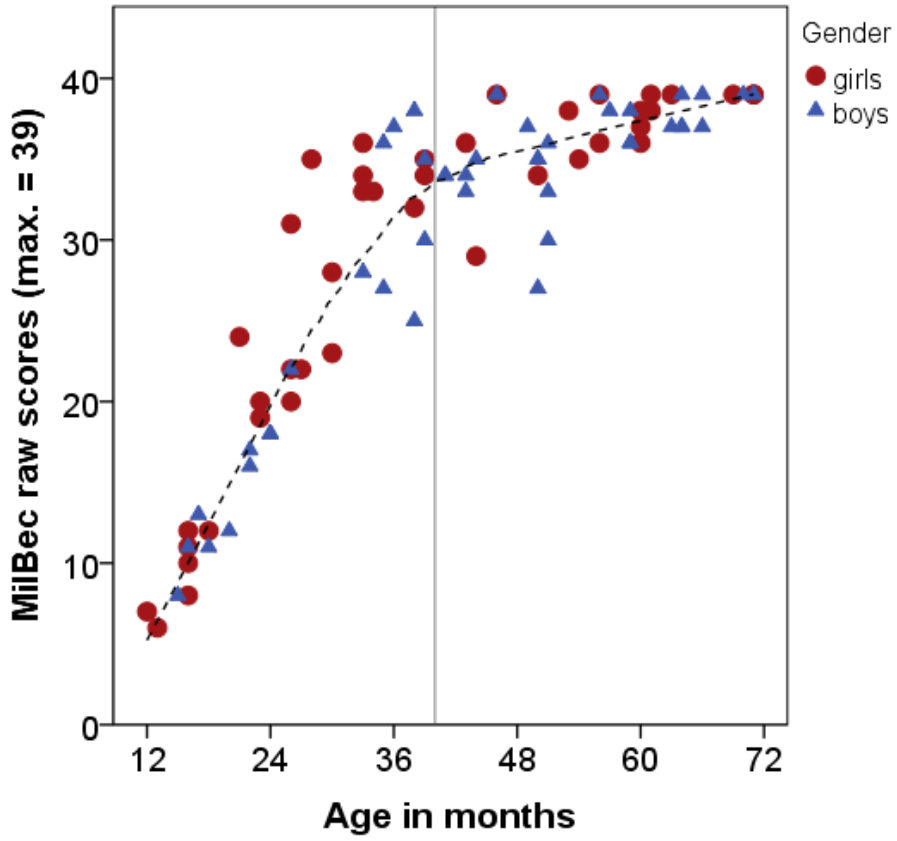


Figure 1. Scatterplot of MilBec score as a function of age, with boys and girls as subgroups. Note: The dotted line indicates the Loess adjustment curve. The vertical gray line at age 40 months indicates the age at which it is judged that the change in slope occurs.

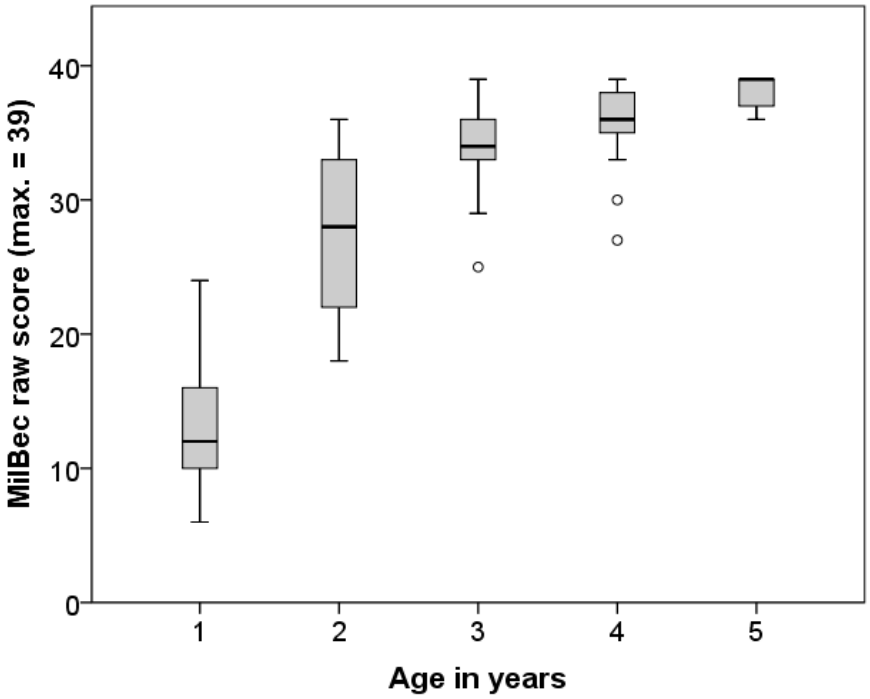


Figure 2. Boxplot of MilBec score per age group (in years).

The median is shown as the line in the middle of the box, which itself indicates the range encompassing 50% of the scores; the whiskers show the range of the top and bottom 25% of the scores; the dots represent outliers.

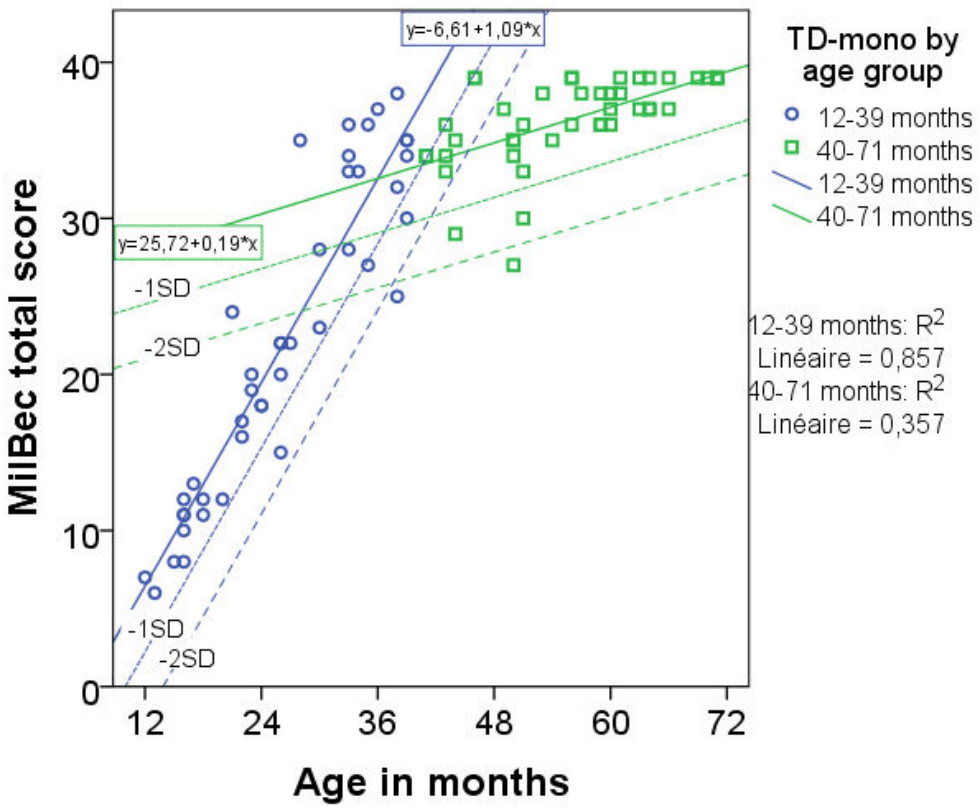


Figure 3. Scatterplot of the total score of the MilBec as a function of age (in months).

Note: The TD monolingual children are separated into two groups: between 12 and 39 months and between 40 and 71 months. The linear regression lines and its formula is shown for each age group in full lines, along with the -1 SD and -2 SD reference lines based on the average standard deviation of each group in dotted lines.

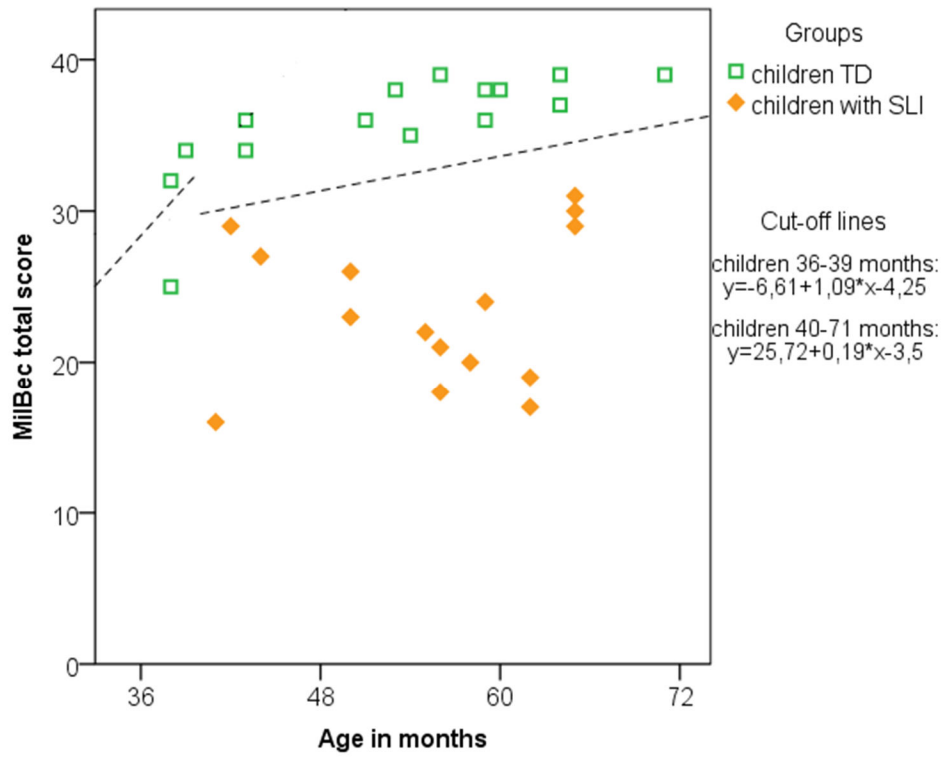
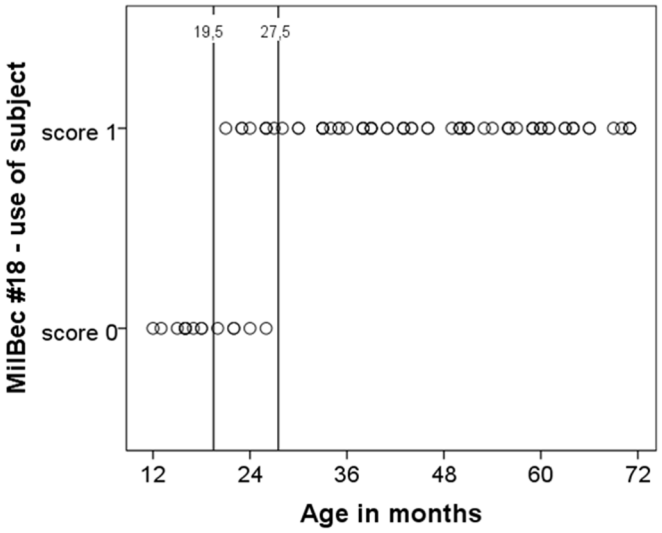


Figure 4. Scatterplot of MilBec scores as a function of age, per language group.

Note: there is a cut-off line for children between 36 and 39 months and one for children between 40 and 71 months.

a)



b)

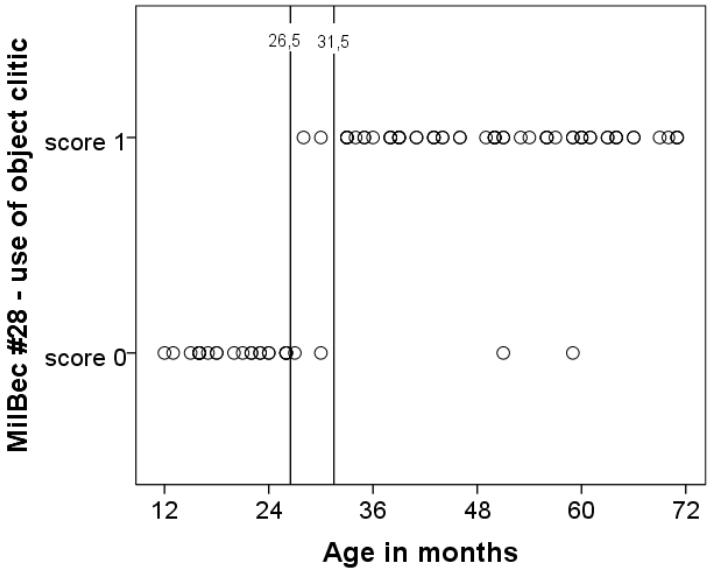


Figure 5. Scatterplot of the distribution of scores for the TD-mono group as a function of age in months for a) subject use, and b) object clitic use.

Note: A score of 1 is obtained when the parent answer "Yes" or "I think so"; a score of 0 when the parent answer "No" or "I don't believe so". The two vertical lines show the period of emergence of the item, with the specific value used indicated in the label at the top.

Table 1: Age, MilBec score and maternal education of the participants per age group.

For each variable, the mean and standard deviation (SD), and the range of score (minimal – maximal values) are indicated in each age group.

Group	Number of participants (Nb of boys)	Age in months : mean (SD), range	Maternal education : mean (SD), range
1 y.o.	17 (7 boys; 10 girls)	17.9 (3.4) range : 12-23	16.2 (2.5) range : 12-21
2 y.o.	17 (6 boys; 11 girls)	29.6 (3.9) range : 24-35	16.6 (3.0) range : 11-20
3 y.o.	17 (11 boys; 6 girls)	41.0 (3.0) range : 36-46	16.3 (2.1) range : 13-19
4 y.o.	17 (12 boys; 5 girls)	53.6 (3.6) range : 49-59	16.4 (3.2) range : 11-24
5 y.o.	17 (8 boys; 9 girls)	64.9 (4.1) range : 60-71	15.8 (1.8) range : 11-18

Table 2: MilBec scores for each subgroup separated by gender, presenting the number of subjects (n), the mean, the standard deviation (SD), and range.

			MilBec score		
	Group	N	Mean	SD	Range
12-31 months	girls	17	18.2	8.7	6-35
	boys	10	14.6	4.3	8-22
32-51 months	girls	11	34.1	2.5	29-39
	boys	21	33.2	3.9	25-39
52-71 months	girls	13	37.9	1.4	35-39
	boys	13	37.8	1.2	36-39

Table 3: MilBec score for each 1-year group, presenting the number of subjects (n), the mean, the standard deviation (SD), and range.

Group	n	mean	SD	Range
1 y.o.	17	12.8	4.9	6-24
2 y.o.	17	27.4	6.4	15-36
3 y.o.	17	34.1	3.6	25-39
4 y.o.	17	35.4	3.1	27-39
5 y.o.	17	38.1	1.1	36-39

Table 4: MilBec score for the 10 subgroups of 6-month intervals used to calculate the average standard deviation to use with the regression line.

The number of participants (n), the minimum, maximum, mean and standard deviation (SD) are indicated for each subgroup.

Age subgroups	Age range	n	Minimum	Maximum	Mean	SD
For children between 12 and 39 months						
1	12-17 months	9	6	13	9.6	2.4
2	18-23 months	8	11	24	16.4	4.6
3	24-28 months	8	15	35	21.5	6.0
4	30-35 months	9	23	36	30.9	4.5
5	36-41 months	10	25	38	33.4	3.7
					Average SD	4.25
For children between 40 and 71 months						
6	43-46 months	7	29	39	35.0	3.5
7	49-53 months	9	27	38	33.9	3.5
					Average SD	3.5
Exclude from the calculation for children between 40 and 71 months						
8	54-59 months	8	35	39	37.1	1.6
9	60-64 months	10	36	39	37.7	1.1
10	66-71 months	7	37	39	38.7	0.8

Table 5: Background characteristics of children with and without SLI.

The number of participants, the mean and standard deviation (SD), the range of score (minimal – maximal values) for age, maternal education and nonverbal IQ of the participants are indicated.

	SLI group	TD group	*p*-value
Number of participants (Number of boys)	15 (10 boys)	15 (8 boys)	.456
Age in months : mean (SD), range	55.3 (8.3) range : 41-65 n = 15	52.8 (10.5) range : 38-71 n = 15	.469
Maternal education : mean (SD), range	12.2 (2.7) range : 8-18 n = 15	16.9 (3.1) range : 13-24 n = 14	.000**
Nonverbal IQ : mean (SD), range	97.0 (7.5) range : 80-111 n = 14	114.1 (11.2) range : 93-137 n = 14	.000**

Note. The p-value for the comparison between groups (khi-square for the number of boys, independent t-test for age, maternal education and nonverbal IQ). The number of participants with available data is also indicated for each measure. NOTES: *: Comparison significant at p = .05; **: Comparison significant at p = .01.

Table 6: Performance of the children with and without SLI on the various language tests.

For each language test and screener, the number of participants (n), the mean, standard deviation (SD), the range of score obtained by each group, the maximum score that can be obtained on the test (test max.), as well as the *p*-value for the independent t-test comparing the two groups and the effect size based on Cohen's *d*.

	Group	n	Mean	SD	Range	Test max.	*p*-value	Cohen's *d*
MilBec	SLI	15	23.47	4,984	16-31	39	.000*	3.4
	TD	15	35.73	3,634	25-39			
IDElex	SLI	15	29.60	6,367	18-37	49	.000°*	5.1
	TD	14	46.36	3,272	37-49			
IDElco	SLI	15	36.20	8,291	16-45	51	.001*	2.1
	TD	14	45.86	4,688	32-51			
EVIP	SLI	15	82.73	17,576	66-127	>160	.000*	1.7
(stand. score)	TD	15	115.80	19,369	84-144			
Sentence	SLI	12	11.17	7,975	0-22	50	.000°*	1.5
Repetition	TD	12	30.08	13,007	11-47			
NEEL7 –	SLI	15	1.47	1,302	0-4	6	.000*	1.3
morphosyntax	TD	13	3.69	1,653	1-6			
NEEL8-	SLI	15	5.93	3,411	1-12	12	.000°*	4.9
positional	TD	13	11.23	1,092	8-12			
concepts								

NEEL8- arithmetic	SLI TD	15 13	2.27 5.85	2,120 2,996	0-6 0-10	10	.001*	1.2
NEEL11 expres. voc.	SLI TD	15 13	40.00 52.62	14,147 6,862	16-60 41-64	70	.006°	(1.8)

Note. Cohen's *d* is indicated in parentheses when there is not a significant group difference using the alpha level based on the Bonferroni correction, but would have been using a *p*-value of .05.

°: t-test that does not assume equality of variance

*: Comparison significant using a Bonferroni correction, with an alpha level of .0055.

Table 7: Pearson's correlation, between the MilBec score and the other language tests and screeners, when all children are considered together and per language group.

The p-value and the number of participants for each correlation are also indicated.

Group			IDE lex	IDE lco	EVIP Raw	Sentence Repetition	NEEL8 positional concepts	NEEL11 expressive vocabulary
All	MilBec	Pearson r	.921	.783	.545	.797	.492	.591
		p-value	.000*	.000*	.002*	.000*	.008	.001*
		n	29	29	30	24	28	28
SLI	MilBec	Pearson r	.756	.556	-.265	.486	-.393	.439
		p-value	.001*	.031	.340	.109	.147	.102
		n	15	15	15	12	15	15
TD	MilBec	Pearson r	.741	.884	.731	.715	.472	.174
		p-value	.002*	.000*	.002*	.009	.103	.569
		n	14	14	15	12	13	13

Note. The number of participants (n) used in the analyses are also indicated.

*: Comparison significant using a Bonferroni correction, with an alpha level of .0024. The correlation with a p-value of .002 are considered significant.

Table 8: The percentage of children with and without SLI who received a score of 1 on the two potential clinical markers.

For each item and per each period found in Figure 5, the percentage of children with or without SLI being reported by their parent as using subject and using object clitic, with the exact number of participants indicated in parentheses.

	TD	SLI
Subject use (item 18)		
under 19 months	0% (0/11)	-
between 19 and 27 months	62% (8/13)	-
over 27 months	100% (61/61)	80% (12/15)
Object clitic use (item 28)		
under 26 months	0% (0/23)	-
between 26 and 31 months	50% (2/4)	-
over 31 months	97% (56/58)	47% (7/15)

Note. The age subgroups are based on the periods determined in Figure 5.

Preface to Chapter 3

In the previous chapter, the preliminary normalization and validation of the MilBec was presented. In Study 1, the results indicated that the MilBec is developmentally sensitive for children between 12 and 71 months, with a particularly high rate of score increase between 12 and 39 months, with a ceiling present for the older children. In Study 2, the pattern of correlations between the MilBec and other language tests supported its concurrent validity for children between 36 and 71 months. Furthermore, high sensitivity and specificity values supported its usefulness at identifying monolingual children with SLI. The MilBec is a promising screening tool to use in large-scale screening of SLI in children between 36 and 71 months.

However, in the Province of Quebec, there are many bilingual children in the population, with 12.6% of Quebeckers speaking French and another language at home (Statistique Canada, 2011a). Furthermore, 26% of children under the age of 14 years are able to speak a language other than French (Statistique Canada, 2011b). In order to be maximally useful in large-scale screening, the ability of the MilBec to correctly rule out SLI in bilingual children should also be investigated. Indeed, the language development of bilingual children differs from that of monolinguals, because the child's exposure to language is divided between each of his or her languages. If his or her language skills in one language were compared to monolinguals' skills during the acquisition period, the difference would be greater as the child's exposure to the language would be lowered (e.g. Elin Thordardottir, 2011, 2015a). Bilingualism thus affects language acquisition, although for a different reason than language impairment does. The pattern of skills affected by bilingualism, when each language is viewed independently and compared to monolinguals' skills, is, however, different from the pattern of skills affected when the child has SLI. On the one hand, children with SLI show language skills that are lowered in several language domains when compared to TD age-matched peers (Elin Thordardottir, 2011). On the

other hand, bilingual children have particularly lowered skills only on some language domains, namely vocabulary and syntax, and usually only when the language assessed is the one the child is the least exposed to (e.g. Elin Thordardottir, 2014, 2015a). Because the MilBec includes items targeting various language domains, this difference between TD bilingual children and monolingual children with SLI could be exploited to help correctly identify bilingual children as having, or not, SLI.

Each item of the MilBec, based on the expected impact of a reduced exposition related to bilingualism on the language domain it targets, was included in one of two subscales: the "language general subscale", which includes the items expected to be the least affected by bilingualism, and the "language specific subscale", which include the items expected to be the most affected by bilingualism. In Chapter 3, the rationale behind the separation of the 39 items of the MilBec into one of the two subscales will be presented. An exploratory study was also performed to test whether these subscales are indeed affected differently by different levels of exposure to French. This study will compare the performance of monolingual children with and without impairment, using the same participants from Chapter 2, with the performance of 20 bilingual children between 12 and 71 months.

Chapter 3: Language Profile on a Parent Questionnaire of Typically Developing Bilingual Children and Monolingual Children with and without Language Impairment.

Marianne J. Paul and Elin T. Thordardottir

School of Communication Sciences and Disorders at McGill University, Montréal, Canada

Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS)

Abstract

Purpose: The typically developing (TD) bilingual child's language acquisition in each language can be slowed when compared to monolinguals', because two languages are being learned side by side. To determine if a bilingual child's low language skills are normal given the child's exposure history or are a manifestation of specific language impairment (SLI) can be particularly difficult. This study investigates the ability of the MilBec, a French adaptation of the parent questionnaire by Luinge et al. (2006), to correctly rule out SLI. The MilBec contains 18 items targeting skills less prone to a slowed acquisition in a bilingual environment, notably narrative skills (excluding microstructure skills, which are related to syntax), articulation, and phonological awareness. These items are included in the language general subscale. The 21 other items target elements more affected by the child's reduced exposure in each language, namely vocabulary and syntax, and form the language specific subscale. **Method**: The performance of 85 TD monolinguals (TD-mono), 15 monolinguals with SLI (SLI-mono), 12 TD bilinguals with high exposure to French (BIL-high), and 8 TD bilinguals with low exposure to French (BIL-low), all between 12 and 71 months, was compared. **Results**: Compared to the TD-mono group: 1) the distribution of scores was generally similar for both bilingual groups and the TD-mono group; however, the proportion of children failing was higher for bilingual children from the BIL-low group and only on the language specific subscale, 2) the SLI-mono group received lower scores on both subscales, 3) for all children, there was an unexpectedly large variability of scores on the language-general subscale. **Conclusions**: Most TD bilingual children have an overall performance similar to the performance of TD monolingual children on the MilBec, making this screener test a promising tool to correctly rule out SLI in bilingual children.

Keywords: language impairment (SLI), French, bilingual, screener test, parent-questionnaire, validation, preschool

Introduction

There is a vast literature on the impact of bilingual upbringing, notably investigating the mental advantages that bilingualism can provide, the effects it has on language acquisition and the attained level of proficiency (e.g. Bialystok, 2015; 2016; Elin Thordardottir, 2011, 2015a; Elin Thordardottir & Brandeker, 2013; Genesee, 2015). In a review on childhood bilingualism, Genesee (2015) reported that the type of bilingual upbringing experienced by the child is an important aspect to consider when assessing a child's language skills. Amongst the elements to consider, there is whether the child is equally exposed to both languages, when the second language was introduced, whether the two languages spoken to the child have the same status in the community (majority or minority languages), and in which contexts each language is being used and by whom. As pointed out in Luk and Bialystok (2013), because of the complexity of these elements, it is not possible to provide a single number to quantify proficiency that would fully represent the knowledge of a bilingual individual. Nonetheless, the relative exposure to each of the child's languages can be used to classify bilingual children into more homogeneous groups. Indeed, it was found that as the relative exposure in a language increased, the child's language skills increased as well (e.g. Elin Thordardottir, 2011, 2015a).

The fact that bilingual children receive reduced exposure to each of his or her languages will have an impact on some, but not all, language competencies (Elin Thordardottir, 2011, 2015a; Genesee, 2015). When a task requires knowledge that must be gained in each language separately, lower performance in one language is expected when the child has been less exposed to that language. For example, it was shown in Elin Thordardottir and Brandeker (2013) that receptive vocabulary in French was highly influenced by the level of previous French exposure; however, for the two measures of processing skills, namely sentence repetition and particularly nonword repetition, the amount of past exposure to French had less influence. An overview of the

language domains included in the MilBec, and to what extent each of them may be affected by the bilingual child's reduced exposure, will be presented in the next section.

Bilingual children produce similar errors as typically developing (TD) monolingual children (Elin Thordardottir, 2015a; Elin Thordardottir & Ellis Weismer, 2002; Genesee, 2015; Paradis, Tremblay, & Graco, 2014). However, children with specific language impairment (SLI) also exhibit errors similar to those made by TD children (for a review of SLI in French, see Elin Thordardottir, 2016). Because of that, it may be difficult to determine if bilingual children have typical language development or not (Bedore & Peña, 2008). For example, a widely recognized clinical marker of SLI in English is the extended presence of the optional infinitive stage (EOI: Rice & Wexler, 1996; Rice, Wexler, & Cleave, 1995). However, Elin Thordardottir (2015a) found that it was also present at age 5 years in TD French-English bilingual children whose dominant language exposure was in French. The presence of these difficulties in bilingual children with lower exposure to English cannot be attributed to an impaired ability to acquire language: the same children had well developed verbal morphology in French, the language to which they were most exposed. Furthermore, bilingual children who had balanced exposure or an English-dominant exposure did not exhibit a similar extended period of use of optional infinitive. These findings were interpreted by the author as an indicator of the necessity of having been exposed to a large amount of English input to exit the stage where finiteness is optional in English.

This finding highlighted the necessity of adapting the process by which language impairment is identified in bilingual children. This had been known for some time, although without many definite answers on how to proceed. A worldwide network of researchers was brought together to form the European Cooperation in Science and Technology Action ISO804, with the goal of improving the assessment of children in minority language groups and bilingual

children. As a member of this group, Elin Thordardottir (2015b) proposed a procedure to identify children with SLI that can be used for monolinguals across languages and for simultaneous bilingual children exposed to a variety of languages. Many recommendations were made, including the use of different cut-off criteria to identify SLI based on the level of past exposure in the language being assessed; these criteria have to be empirically determined for each test. For example, for non-balanced simultaneous bilingual children, a cut-off score of -1.5 to -1.75 could be used in the language with the higher level of exposure, whereas in the language in which a lower exposure is received, a cut-off of -2.25 to -2.5 could be used instead. Based on this recommendation, the performance of bilingual children on the Milestones en français du Québec (MilBec: Paul & Elin Thordardottir, 2010) should be investigated to determine the impact of different levels of exposure to French on the scores obtained by bilingual children on the MilBec subscales. This investigation is the first step toward identifying the cut-off score to be used to identify bilingual children at risk of SLI.

The MilBec is a language screening tool adapted from the Dutch parent-questionnaire by Luinge, Post, Wit, and Goorhuis-Brouwer (2006) for children between 12 and 71 months (for the details of the adaptation procedures, see Chapter 1). The MilBec contains 39 items asking the parent to judge if their child is able to understand or use certain linguistic elements (e.g. combine words, make rhymes, speak intelligibly). The items of the MilBec are mostly general language milestones targeting expressive and receptive skills, from various language domains (vocabulary, syntax, morphosyntax, articulation/phonology, metalinguistic skills, pre-literacy, and narration). The parent is asked to indicate if the child is using the element mentioned in the item, by answering “Oui” (yes) or “Il me semble” (I think so), both scored as 1 point, or “Je ne crois pas” (‘I don’t believe so’) or “Non” (‘no’), both scored as 0 points. The MilBec has been shown to be developmentally sensitive for monolingual French-speaking children, to have good concurrent

validity and good reliability (see Chapters 1 and 2), and is considered promising as a quick screening test for language impairment with good diagnostic accuracy (see Chapter 2).

In order for the MilBec to be useful with bilingual children, it was proposed to separate its items into two subscales, depending on the extent to which the domains they target are more or less affected by bilingualism. The rationale for the inclusion of each domain as pertaining to one or the other subscale is provided in the next section. The 39 items were then separated into a language-specific scale containing the items hypothesized to be more affected by bilingualism, and a language-general subscale containing the items hypothesized to be less affected by bilingualism (see Table 1). There are respectively 21 items and 18 items on each subscale. In order to determine if language impairment equally affects the child's language performance on domains from both subscales, a brief review of the effect of language impairment will also be presented.

[Insert Table 1 about here]

Language Development of TD Bilingual Children

The literature on bilingual children suggests that some language domains are more vulnerable to the effects of reduced exposure than are others. In this section, a review of the domains included in the MilBec is presented. First, the language domains whose acquisition relies on the exposure received in each language, the language-specific domains, will be presented. Then, the language domains that depends less on the exposure in a particular language, the language-general domains, will follow.

Language-specific domains. From the language domains included in the MilBec, two are particularly affected by a lowered language input, namely vocabulary and syntax. Indeed, it has been repeatedly shown that bilingual children and adults have a significantly lower vocabulary

than monolinguals when each language was considered separately (e.g. Bialystok & Feng, 2009; Bialystok, Luk, Peets, & Yang, 2010; Uccelli & Páez, 2007). Elin Thordardottir, Rothenberg, Rivard, and Naves (2006) investigated the vocabulary and syntactic development of French-English balanced bilingual children at age 2 years (i.e. receiving the same amount of exposure in both languages). They compared the performance of these children to the performance of monolingual French- and monolingual English-speaking children of the same age. In Montreal (Quebec, Canada) where the study was conducted, these two languages are both considered majority languages. The results showed that the bilingual children's total conceptual vocabulary, which is a combined vocabulary score that takes into consideration the non-overlapping known words in both languages (Pearson, Fernández, & Oller, 1993), was significantly lower only when it was compared to monolingual English children. However, the total vocabulary, which adds the vocabulary in each language known by the child, was not significantly different from the monolinguals' vocabulary. Finally, when the bilingual children were treated as monolinguals in each language, considering their knowledge in that language only, they received lower scores on vocabulary and syntax, for both the expressive and receptive modalities. However, the extent of the differences varied depending on the language of comparison, even though the children had no a priori language dominance, having had similar levels of exposure in both languages. The authors concluded that these findings highlighted the complexity of the assessment of bilingual children with a balanced exposure in each language to identify children with language impairment. Indeed, developmental paths are not the same across languages, and these differences can be observed in comparison of the children's performance on tests that were similar in the two languages. Notably, differences were found between monolingual children speaking different languages for vocabulary size (Elin Thordardottir, 2005; Elin Thordardottir et al., 2006), vocabulary composition (Bornstein et al., 2004), as well as mean length of utterances

(Elin Thordardottir, 2005).

In a continuation of this study, Elin Thordardottir (2014) conducted a series of studies to better understand the effect of language input, defined as the percentage of exposure received in a particular language. The participants were 3- and 5-year-old children from Montreal (Canada), classified as monolingual French, bilingual with more French (with a greater exposure to French), French-English bilingual (with equal exposure in both languages), bilingual with more English (with a greater exposure to English) and monolingual English. The classification of the children was based on the relative exposure in each language calculated from a detailed parental report. One of the main findings of these studies was that the development of vocabulary in a language was highly dependent on the level of exposure in that language. For the 5-year-old bilingual, it was found that to reach a level similar to the monolinguals in receptive vocabulary, an exposure of 40% to 60% in the language was required, whereas for expressive vocabulary the required exposure was closer to 70%. The findings for the 3-year-old were congruent with the thresholds found for the 5-year-old. These results indicated that although vocabulary was affected by bilingualism, when a child received a high level of exposure in either language, his or her skills would be within normal limits when compared to monolingual children. It is when the exposure to the language was lower than 50% that a reduced language performance could be expected for both expressive and receptive vocabulary skills.

For the grammatical development of these children, Elin Thordardottir (2015a) reported that mean length of utterance was not affected strongly by bilingualism, but the diversity and accuracy of grammatical morphology in each language were dependent on the level of exposure in each language. On the one hand, for bilingual children who did not receive equal exposure in both languages, the grammatical skills in the language in which they received a low exposure were closer to the skills of children with SLI. On the other hand, it was found that bilinguals have

much better grammatical skills in the language in which they receive a higher level of exposure, with skills similar to those of monolingual speakers of the same language. For bilingual children with an equal exposure to both languages, a level comparable to that of the monolinguals was found for both of their languages. Such a similarity between monolinguals and bilinguals receiving a high level of exposure in the language was also found in a series of studies reported in Paradis et al. (2014). In these studies on 4- or 6-year-old French-English bilingual children living in Edmonton (Canada), where the language of the majority is English, the morphosyntactic abilities of the children were investigated. As was the case for vocabulary skills, when the child received more input in one language, his or her skills in that language were found to be similar to the skills of monolingual speakers of the same age.

These results thus supported the high dependency of vocabulary and syntax/morphosyntax on language input in the specific language being measured, making them language-specific skills. However, a reduced performance was expected only when the child receives a limited input in that language. The required level of exposure to score similarly as monolinguals was between 40% and 70%, depending on the particular language domain or modality. It is thus important to control for the level of exposure in each language when investigating language performance in bilingual children.

Language-general domains. Many language domains included in the MilBec are much less influenced by bilingualism than vocabulary and syntax, notably narrative skills, which are usually assessed for children over 4 years of age (Paul, R. & Norbury, 2012), articulation/phonology, phonological awareness, pragmatics, and pre-literacy skills. Findings supporting the low impact of bilingualism on these language domains will be presented in this section.

Narrative abilities of children are targeted in three items of the MilBec. These items are

classified as belonging in the language-general domain, notably because of the results of Hipfner-Boucher, Milburn, Weitzman, Greenberg, Pelletier, and Girolametto (2014). They investigated the ability of 4- and 5-year-old bilingual children to perform a story retell task, and they compared the results of children attending an English school, but who were predominantly using English or a minority language at home. Their conclusions were that the macrostructure abilities (i.e. related to story grammar) of both groups of bilingual children were similar to those of the monolingual children. Also, Uccelli, and Páez (2007) found that the narrative skills of English-Spanish bilingual children were moderately and positively correlated across the child's two languages, with children receiving high scores in both languages, or low scores in both languages. Narrative skills were thus included in the language-general subscale.

Other skills that are considered to be rather unaffected by a reduction in language exposure due to bilingualism, or for which there is a closer relationship between the skills of a child in both of his or her languages, are articulation/phonology, and meta-linguistics. For example, Schmitt, Simoës and Laloi (2015) investigated a wide range of language abilities of bilingual children between 5- and 7-year-old with an exposure to French between 7%-81%. The results showed that bilingual children's skills for phonology and meta-phonology were similar to those of monolinguals. The language-general status of articulation is also supported by Rvachew, Marquis, Brosseau-Lapré, Paul, Royle, and Gonnerman (2013), who reported that in French, the articulation abilities of bilingual children in kindergarten and first grade, as assessed using a test eliciting single words, were similar to monolingual children's skills. Chiang and Rvachew (2007) showed that for English children learning French, the English phonological awareness skill was the most important predictor of French phonological awareness skills. This indicated that when a child's phonological awareness skills were high in one language, they can be expected to also be high in the other language. The closer relationship between the two languages in the domain of

phonological awareness explained its classification in the language-general subscale. Finally, it is even proposed that some metalinguistic skills can be enhanced in bilingual children, when the tasks require the child to ignore misleading information, and in no case bilingualism would lead to a disadvantage (for a review, see Bialystok, 2012).

Finally, the MilBec includes items related to print knowledge, a pre-literacy precursor of reading achievement. Print knowledge encompasses different elements: print interest, print form, which includes knowledge about words and letters, organization and directionality of print, functions of print, and part-to-whole relationships, which includes alphabet knowledge (Justice & Ezrel, 2004). Dickinson, McCabe, Clark–Chiarelli, and Wolf (2004) found that bilingual children performed in a highly similar way in both of their languages. In their study, it was even decided to cease the administration of the test assessing many of these print knowledge elements in one of the languages, but accepting as correct an answer given in either language.

These studies suggested that, in contrast to vocabulary and syntax, many language skills are relatively unaffected by reduced exposure associated with bilingualism. The items from these relatively unaffected domains were included in the language-general subscale.

Language Development of Monolingual Children with SLI

In this section, the characteristics of the performance of children with SLI on the various language domains included in the MilBec is presented. Language acquisition of children with SLI is generally found to be delayed on a variety of language abilities required for efficient language communication. In French, Elin Thordardottir and Namazi (2007) compared the spontaneous language samples of children between 3- and 4 and a half years-old with and without language impairment, and to TD younger MLU-matched control children. The results showed that the abilities of children with language impairment were generally similar to the abilities of younger children with TD in the areas of morphosyntax (mean length of utterance in words and in

morphemes, and diversity of grammatical morphology used) and vocabulary (receptive vocabulary, number of different words, proportion of closed-class words and social words). Furthermore, this study showed that in French, in contrast to what was found in English, children with SLI made very few morphological errors. The authors concluded that language impairment had an impact across language domains. Similar findings are reported in Elin Thordardottir et al. (2011) for children with and without SLI between 54 and 71 months. In this study, children with SLI showed reduced abilities in a variety of linguistic domains and processing abilities, notably on receptive vocabulary, receptive morphosyntax and complex syntax, nonword repetition, sentence imitation, following directions and forward digit span. Children with SLI can thus be predicted to have reduced skills on the domains included in the language-specific subscale of the MilBec.

Concerning the abilities related to the narrative skills included in the language-general subscale of the MilBec, a greater range of abilities were reported for children with SLI. Hilaire-Debove, and Kern (2013) described a multiple case study on three monolingual French-speaking children, one aged 4 years 8 months, one aged 6 years and one 9 years. The results showed that there was a high variability in the performance of children with SLI on narrative macrostructures using the Frog story. Additionally, Elin Thordardottir et al. (2011) showed that on a test of narration using a more structured task, children with SLI scored lower on story grammar and first mention (a measure of use of cohesive ties), although the comparison with age-matched peers revealed no statistically significant differences. It is to be kept in mind, when talking about narrative skills, that there is a great variability in narrative tasks and types of analyses, and that these variations do not lead to a similar impact on the children's performance (e.g. Duinmeijer, de Jong, & Scheper, 2012; Heilmann, Miller, & Nockerts, 2010; Kadevarek, & Sulzby, 2000).

Many language domains previously identified as pertaining to the language-general

subscale have been found to be impaired in monolingual children with SLI, namely articulatory, phonological and metaphonological skills. For example, Maillart and Parisse (2006) investigated the phonological abilities of French-speaking children with SLI, comparing them to younger children matched on mean length of utterance and phonemic inventory. The results showed that children with SLI were more likely to omit or add syllables when producing a word, and they had a lower percentage of phonemes correct than TD (younger) children matched on mean length of utterance and phonemic inventory size. Furthermore, Piérart (2014) investigated the language skills of French-speaking children with SLI using a multiple-case study and found important delays in articulation, phonology, and lexical skills, with an important variability of profiles amongst children. Finally, concerning phonological awareness of 5- and 6-year-old French-speaking children with SLI, it was found on a task of phoneme deletion that all children performed more than 1 standard deviation below the normative sample's mean (Zourou, Ecalle, Magnan, & Sanchez, 2010).

Other domains targeted in the MilBec by a couple of items are language use and use of gesture, also classified as pertaining to the language-general subscale. A study on English-speaking preschoolers with SLI concluded that these children had some degree of impairment in pragmatic skills such as turn taking, rate of communication, joint attention (for a review, see Weiss & Paul, R., 2010) and use of gesture (Schwartz, 2009). These difficulties were reported to vary across children, with some having more prominent difficulties in some areas, and other children showing different profiles of impairment. Difficulties with the use of gesture to communicate had also been investigated in a longitudinal study by Thal and Tobias (1992) on a population of late-talkers. It was found that based on the classification using the one-year follow up status (normalized or not), the late-talkers who had normalized their language skills were using more gestures the year before than both the age-matched and language-matched controls.

On the other hand, the late-talkers who were still behind in their language skills were found to perform significantly lower on the gesture tasks the year before. Furthermore, children with SLI between 4- and 8-year-old were found to perform significantly lower than TD children between 4- and 7-year-old in tasks of gesture production and gesture comprehension (Wray, Norbury, & Alcock, 2016).

Finally, the last aspect of language included in the MilBec is early literacy skills. The ability of English-speaking children with SLI in this domain had been investigated by Boudreau and Hedberg (1999). They found that overall, children with SLI performed significantly lower than age-matched peers in print concept abilities. However, an analysis by item categories showed that this difference arose from difficulties with metalinguistics- and reading-related vocabulary. In Cabell, Lomax, Justice, Breit-Smith, Skibbe, and McGinty (2010), a great variability in scores of children with SLI was found for emergent literacy skills. According to McGinty and Justice (2009), the variability in the children's scores would not be primarily due to varying language skills. Indeed, they found that the most important predictor of pre-literacy skills was the quality of home literacy mediated by attentional abilities of the children that explained the variability in print knowledge of children with SLI.

This review confirmed the language difficulties of children with SLI in many language domains, and highlighted the great variability in performance of these children. This high variability of language skills was observed particularly on pre-literacy and narrative skills, which are both language-general skills. Nonetheless, the domains in which a child with SLI may have difficulty can be from either from the language-specific subscale or the language-general subscale, as defined in the previous section.

Hypotheses

This study is an initial exploration to determine whether the MilBec, because it targets

items from various language domains, can provide additional insights into the impact of the reduced exposure associated with bilingualism on language acquisition for TD children. Indeed, since bilingualism does not have an impact on all language domain to the same extent, it was proposed to separate the items of the MilBec into two subscales (see Table 1). The language-specific subscale contained 21 items targeting language domains the most likely to be affected by bilingualism, namely vocabulary and syntax. The language-general subscale contained 18 items targeting domains relatively unaffected by bilingualism, namely narrative skills (excluding narrative microstructure), use of gesture, language use, phonology/articulation, meta-linguistic skills and print-knowledge. Furthermore, because the domains targeted in the language-specific scale are not expected to be affected equally for all bilingual children, depending on each child's level of exposure to French, the participants' level of French exposure should be considered. Indeed, it is generally only in the language in which the child received the lowest level of exposure that lower language performances would be expected. It was considered in this study that a child had a low level of French exposure when less than 55% exposure was received in French, and a high level of exposure when it was 55% or more.

The goal of this study was to compare bilingual children's scores to the scores obtained by the larger sample of monolingual French-speaking children that was previously tested using the MilBec. Performing group comparisons would be possible only if a homogeneous group with a narrow age range was recruited and then compared to a matched subsample of the TD children. However, the children who were recruited in the study were of various ages; the analyses were thus performed by comparing the scores of each child with the cut-off scores set for his or her age. The advantage of the selected method, compared to group comparison, was that it would be more informative of the diagnostic value of the subscales. Indeed, even if statistically significant group differences were found, the difference may not be sufficiently large to be meaningful or

that monolingual normative data should not be used. Furthermore, it was at the level of individual children that the effect bilingualism on the children's performance had to be documented, as it was at this level that a clinician would have to rule in or rule out SLI. The specific hypotheses that would be tested were thus the following:

Compared to the language skills of monolinguals with TD, as assessed by the MilBec:

1. The language skills of bilingual children with high French exposure would be within the normal range when compared to TD monolingual children for both subscales.
2. Bilingual children with low French exposure would obtain lower scores on the language-specific subscale only, and a higher proportion of children would fail this subscale than the language general subscale.
3. The language skills of monolingual children with SLI would be much lower than those of children with TD for both subscales, and a high proportion of children with SLI would fail the two subscales.

Method

Participants

Four groups of children were included in this study, all between 12 and 71 months. The first group, composed of TD monolingual children, was the reference group to which the other groups were compared. The other three groups of children were two bilingual groups having a different level of exposure to French, and one group of monolingual children with SLI. Given the exploratory nature of this study, these groups were not matched on any background variables.

The reference group included 85 (44 boys, 41 girls) TD monolingual French-speakers (TD-mono) between 12 and 71 months, with 17 participants per 1-year age group. These data were collected for the normative study presented in Chapter 2. The children had no parental

concerns regarding language, cognition, or audition. From these participants, 10 had a negligible exposure to another language.

The TD bilingual group included 20 children aged 12 to 71 months recruited for this study. These children were from various language backgrounds in terms of the specific languages and the number of languages to which they were exposed. For each child, a parent questionnaire on language exposure detailed elsewhere (Elin Thordardottir et al., 2006) was used to determine the relative exposure to each of the child's languages. Children were then separated into one of two groups, depending on the child's percentage of exposure they had had to French: twelve children (6 boys) had an exposure to French between 55% and 95%; these children formed the BIL-high group. The remaining 8 children (5 boys) had an exposure to French between 5% and 54.9%; they formed the BIL-low group.

The SLI-mon group included 15 (10 boys, 5 girls) monolingual children with SLI between 36 and 71 months. These data were collected for the validation study presented in Chapter 2. The children were recruited at a rehabilitation center, mostly from the Institut Raymond-Dewar (Montreal, Quebec, Canada), all had a moderate to severe language impairment or a severe language delay.

Background variables. The background variables (age in months, maternal education, and percentage of French exposure since birth) are described for the four groups of participants in Table 2. An ANOVA showed that the groups differed on maternal education (p = .000). Because the groups had unequal sample size and that they had unequal variance, Games-Howell post-hoc comparisons were performed, showing that the SLI-mono group was significantly different from the TD-mono group (p = .000) and from the BIL-high group (p = .007). Group comparisons for the other variables were not performed, as the groups were not matched on any variables.

[Insert Table 2 about here]

Procedures

All parents signed the consent form of the project, which was approved by the Ethics committee of the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR-CIUSSS). The recruitment of the bilingual children was performed at the same time and following the same procedure as the recruitment for the normative study presented in Chapter 2. The parents were invited to participate in a study on the validation of a parent questionnaire on language development via daycares, sports centers, school and public billboards using e-mail, posters, or billboard postings. Parents were given the option to fill out the MilBec and a demographic questionnaire using an on-line survey created using LimeSurvey or using a paper-pencil version. For the monolingual children with SLI and a subgroup of TD-mono children, various tests were administered (for more details, see Chapter 2) although only results from the demographic questionnaire and the MilBec are used in the current study.

The demographic questionnaire was used to determine the bilingual children's exposure to French, which was then used to assign the child into the low- or high-exposure group. Then, for each child, the MilBec total score and scores on the two subscales were calculated. All statistical analyses were performed using the French version of IBM Statistic SPSS version 23.

Analyses and Results

To test the hypotheses, the proportion of children having failed each subscale had to be compared across groups. In particular, the hypotheses predicted the performance of the bilingual and SLI groups in relation to the performance of the TD children. The reference values used for this comparison were calculated separately for each subscale and involved the proportion of TD monolingual children who received a score below the set cut-off value, as detailed in the next section. For the other groups, the proportions of children who failed each subscale were then calculated and compared to the reference proportions.

Identifying Normal Limits

For of each subscale, the regression line based on the mean performance of TD monolingual children was calculated; then, two additional regression lines adjusted to take into consideration the standard deviation (SD), one at -1 SD and one at -2 SD, were calculated. The same procedures presented in Study 1 of Chapter 2 were followed: because of the nonlinearity of the data, the regression lines were produced for two age subgroups: one for children under 40 months and one for children 40 months old or older. The SD used for the younger children was the average SD of the five subgroups of 6-month intervals between 12 and 41 months. For the older children, it was based on the average SD of the two groups of 6-month intervals of children between 42 and 53 months; children 54 months and older were not included in this calculation, because the presence of a ceiling effect in these groups greatly reduced the value of the standard deviation (see Table 3). It is worth noting that for the language specific subscale, the average SD of the older group was still half the value of the average standard deviation of the younger group.

[Insert Table 3 about here]

[Insert Figure 1 about here]

The optimal cut-off value allowing the majority of TD children to be classified as being "not at risk" (i.e. above the line) was determined based on visual inspection of Figure 1, which presents a scatterplot of the scores of the TD monolingual children as a function of age against the three regression lines. This visual inspection showed that there was much greater variability in the relationship between scores and age in the language general scale than in the language specific subscale, an unexpected finding that will be addressed in more detail in the discussion. Also, Figure 1 showed that for the two subscales, for both age groups, the majority of the TD-mono children fell above the -1 standard deviation reference lines, and very few children fell

under the -2 standard deviation reference lines. However, in the language specific subscale for the older group there was a somewhat higher number of children falling between the -1 standard deviations and the -2 standard deviation reference lines. Based on these observations, it was decided to use a cut-off score of -1 standard deviation, except for the language-specific subscale for older children for which a -1.5 standard deviation was used to obtain a similar proportion of children failing the subscale.

Performance on the Subscales

The raw scores of all the participants on both subscales as a function of age in months compared to the cut-offs are presented in Figure 2; the formula used for each reference line is indicated beside its corresponding line. For both subscales independently and for each child, it was determined using the formula whether he or she had a score under the cut-off value, thus failing the screening (see Table 4, presenting the younger and older children groups separately). A comparison of the number of TD monolingual children having failed the screening depending on the age group was performed using a khi-square test; this comparison was not performed for the other three groups because of their small sample sizes. The analysis showed no statistically significant difference between the two age-groups, neither for the language-specific subscale nor the language-general subscale (respectively $\chi2(1) = .668$ at $p = .414$, and $\chi2(1) = .605$ at $p = .437$ respectively). This indicated that the proportion of children failing the two subscales did not differ between the younger and older children for the TD-mono group. It was thus judged appropriate to combine them to determine the proportion of children receiving a score under the cut-off values for the whole age-range, for each subscale (see Table 5). For the TD-mono group, which was the group of comparison, 6 of the 85 participants (7.06%) received a score under the cut-off on the language-specific subscale; for the language-general subscale, 8 out of 85 (9.41%) participants received a score below the cut-off values. These values were used as the reference

proportions in the comparisons with the other groups.

[Insert Figure 2 about here]

[Insert Table 4 about here]

[Insert Table 5 about here]

In order to determine whether the proportions of failed screenings were the same for the other groups, the proportion of children failing the screening was calculated across the whole age range for the SLI-mono, the BIL-high and the BIL-low groups using the same procedure presented above (see Table 5). A nonparametric khi-square test was then performed for each of the three groups for both subscales, comparing their proportions to the reference proportions. It is to be noted that the khi-square test could not be performed for the language-specific subscale for the SLI-mono group, because 100% of the children failed it; it can be said, however, that the proportion of SLI-mono group failing that subscale of the questionnaire was statistically different from the reference proportion, which was 7.06% for that subscale. For the SLI-mono group, the difference in proportion was also significant for the language-general subscale (p = .000). Concerning the two bilingual groups, the only statistically significant difference was found for the BIL-low group on the language-specific subscale (p = .048), with more subjects failing the test than expected.

Finally, the number of participants having failed both the language-general and the language-specific subscales was calculated (see Table 6). For the TD-mono group, 5 of the 85 (5.88%) children failed both subscales, which was lower than the number of children having failed the language-specific subscale (n = 6) or the language-general subscale (n = 8). This indicated that in the TD-mono group, 5 children failed both subscales, 1 failed only the language-specific subscale, and 3 failed only the language-general subscale. For the other groups, the

children who failed the language-general subscale also failed the language-specific subscale; one child in the SLI-mono and one child in the BIL-high group failed only the language-specific subscale. The proportion of TD monolingual children having failed both subscales was compared to the proportion of children in the other groups that had failed both subscales. A statistically significant difference was found for the monolingual children with SLI and the bilingual children with a low exposure to French (p = .000 and p = .022, respectively), indicating that more children in these groups failed both subscales.

[Insert Table 6 about here]

Visual inspection of Figure 2a showed that the distribution of scores of both groups of bilingual children was close to the distribution of the TD monolingual children on the language-specific subscale. However, of the children scoring under the reference line, the lowest scores belonged to bilingual children, who had a performance similar to the performance of monolingual children with SLI. These children with SLI received overall much lower scores, with some children scoring closer to the reference line. Concerning the language-general subscale, visual inspection of Figure 2b showed a variability of scores for bilingual children as wide as the TD monolinguals'. Concerning monolingual children with SLI, there seemed to be two subgroups of children, one group receiving scores close to the reference line and one receiving much lower scores.

Discussion

The first hypothesis stated that children in the BIL-high group would be within the normal range when compared to the TD-mono group for both subscales, whereas the second hypothesis stated that the BIL-low group would obtain lower scores on the language-specific subscale only, with a higher proportion of children failing this subscale than the language general subscale. The

results for the bilingual children from both groups on the language-general subscale are congruent with these hypotheses. However, the results on the language-specific subscale, which included items targeting vocabulary and syntax, are only partly congruent with these hypotheses. On the one hand, the congruent findings are that the difference in proportion of children having failed the language-specific subscale in the BIL-high group was not statistically different from the performance of the TD-mono group. Furthermore, the proportion of children in the BIL-low group failing this subscale was statistically significantly different from that of the TD-mono group ($p = 0.048$). On the other hand, the non congruent results are that the performance of the children in the BIL-low group did not show a generalized lower performance when compared to the TD-mono group. Indeed, visual inspection of Figure 2 showed that even the bilingual children with a lower exposure to French often scored as high as TD monolinguals. These results suggest that bilingual children with a wide range of language exposure to French and language experiences perform similarly to TD monolingual children in the language-specific subscale of the MilBec. Concerning the third hypothesis, which stated that the language skills of monolingual children with SLI would be much lower than those of children with TD for both subscales, and a high proportion of children with SLI would fail the two subscales, the results are congruent with all children failing both subscales. A discussion of these results for each subscale is presented next.

Language-specific subscale

A possible explanation for the overall high performance of bilingual children, particularly for those with a low exposure to French, is that some parents may have answered the items while considering the language skills of the children across languages and not only based on their skills in French. For example, a parent might have answered that their child uses subject even if the language in which they used it in not French. Indeed, it was not specifically mentioned in the

instructions to consider exclusively the child's skills in French, even though all the examples and the items were written in French. However, even if it was the case, it would not reduce the usefulness of the MilBec as a screener. Indeed, if the child had strong language skills in his or her other language then he or she would be unlikely to have SLI, because SLI manifests in both languages (for a review, see Kohnert, 2010). These results could also be partially explained by the fact that the amount of exposure required for bilingual to perform in a similar way as monolinguals in the MilBec might be less than the chosen level of 55% of French exposure.

Alternatively, Paradis (2010) proposed that bilingualism could have an impact on the organization of the child's linguistic system, with some abstract aspects of the system being used for the two languages, thus leading to a facilitated learning that could at least partially compensate the reduced exposure received in each language. Reoper (2011) made a similar proposition in his explanation of the Multiple Grammars, a view of language acquisition under the Universal Grammar view. Transfer would be possible only when the two grammars are compatible, and would be better explained as a merging of the two grammars in a single underlying representation, requiring the addition of some elements to a pre-existing representation. In any cases, it has to be remembered that assessing a single language of a bilingual child is assessing only part of his or her abilities, which are distributed amongst both languages. More data from children with different types of French exposure have to be collected and compared to their performance on direct language assessment to better understand their performance. The way parents answer the items (considering only French or any languages) should also be investigated to better understand the pattern of performance of bilingual children.

A closer visual inspection of Figure 2 revealed another difference between monolingual and bilingual children. Indeed, for the few subjects falling below the cut-off reference line on the language-specific subscale, the TD monolingual children are found to be close to the reference

line. In contrast, regardless of their level of exposure to French, the TD bilingual children who failed the subscale received similar scores to monolingual children with SLI, which were generally much below the cut-off reference lines. This finding could be solely due to sampling error, given the small sample size of the bilingual groups and the heterogeneity of the participants within the groups in terms of bilingual experiences (the specific language combinations, the status and number of the children's languages, and the quality of language exposure received by the child). Also, the possibility that some of these children have true language difficulties cannot be ruled out, since their language skills were not formally assessed. In a larger study, detailed analyses at the subject level could be performed to determine if the children's characteristics, such as the number of languages spoken and precisely which languages are spoken, could account for this finding. For these bilingual children, it should be investigated whether a less stringent cut-off score, as proposed in Elin Thordardottir (2015b), would be better to correctly separate TD bilingual children from bilingual children with SLI.

However, if this finding is a manifestation of a true difference in bilingual children's performance, an additional explanation taking into consideration the child's processing skills could be put forward. A generalized low performance of monolingual children with SLI on both language tests and processing measures was reported in Elin Thordardottir et al. (2011). Also, Elin Thordardottir and Brandeker (2013) found that different language measures can be independently affected by language impairment or by a reduced exposure in a bilingual context. These double sources of score reduction for bilingual children with SLI were found for the vocabulary measure, but not for a task less dependent on language knowledge, namely the nonword repetition task. Similarly, Orgassa and Weerman (2008) found that bilingual children with SLI had even lower language scores than monolingual children with SLI, which would stem from the combined effect of reduced processing skills combined with a reduced exposure due to

bilingualism. As pointed out in Elin Thordardottir (2015a), low language scores can be obtained for two different reasons: low processing skills for children with SLI, but a reduced exposure to the language for TD bilingual children. Indeed, the scores of TD bilingual children on many language tests was found to be highly dependent on the lifelong exposure level received in the tested language (Elin Thordardottir, 2011, 2015a).

In the same line of thinking, the absence of bilingual children slightly below the reference line could be related to an additive effect of reduced exposure due to bilingualism and processing abilities in the lower range of the normal variation. This explanation has to be considered along with Gathercole's (2002) hypothesis stating that a certain amount of input is required before a grammatical skill can emerge, that is when a threshold based on a 'critical mass' of exposure is crossed. Once the new grammatical skill is learned, additional exposure would not lead to a greater mastery of that skill. In the case of bilingual children, the findings by Marchman, Martínez-Sussmann, and Dale (2004) also have to be considered. The authors showed that for English-Spanish bilingual toddlers with different levels of exposure in each language, the notion of critical mass was highly language-dependent. Indeed, there was a strong within-language relationship between the vocabulary level requires for early syntax to emerge. In contrast, there was only a weak relation between grammatical skills in one language and vocabulary knowledge in the other language. The authors concluded that each language of a bilingual child developed independently. A similar conclusion was reached by Elin Thordardottir (2015a) for 3- and 5-year-old French-English bilingual children.

It is possible that children with normal, but weak, processing skills will have different language skills depending on whether they have a reduced exposure due to bilingualism or not. The fact that a child does not have strong processing skills would not be apparent if the child is monolingual, because the amount of input received would be sufficient to reach the critical mass.

However, because the bilingual child receives less input in each of his or her language, it is possible that this reduced exposure would not sufficient to reach it, which lead to an apparent slowed language abilities in that language, with the child performing in a similar way to monolingual children with SLI. On the other hand, a bilingual child with strong processing abilities, even if he or she receives a reduced language exposure due to bilingualism, his or her higher processing skills may allow to gain more from the received input, allowing a performance similar to the monolingual children's. It is to be highlighted that an assumption is made in this explanation: the child's processing abilities are having an impact on his or her efficacy at extracting the pertinent elements from the received input. When a child has low processing skills, more exposure would be required before the child can reach the critical mass, whereas children with good processing skills would need a lower amount of exposure to reach the same level of performance. More research investigating these hypotheses, in opposition to more innate grammar views of language learning, should be undertaken.

Language-general Subscale

The results of the bilingual children and monolingual children with SLI on the language-general subscale are congruent with the hypotheses. Indeed, the proportion of children having failed the screening did not differ significantly between the TD monolingual and both groups of bilingual children, and all but one monolingual child with SLI failed the language-general subscale. Visual inspection of Figure 2 showed that the distribution of scores was similar between the monolingual children and both groups of bilingual children. However, there was a subgroup of children with SLI obtaining scores comparable to the scores of TD children in the lower range; the rest of the children with SLI obtained much lower scores than what was found in the other groups. This pattern of performance was thus quite different from what was observed for the language-specific subscale.

For the group of monolingual children with SLI, these results could simply be explained by varying level and type of impairment across children: some children may have no or only a mild impairment on the aspects targeted on the language-general subscale and a more severe impairment on vocabulary and syntax; other children can have a more severe profile of impairment for all language domains. This explanation is supported by the fact that pre-literacy skills and narrative skills, which were targeted with 3 items each in the MilBec, are not always reported as areas of difficulty for children with SLI. For example, it was found in Elin Thordardottir et al. (2011) that a test of narrative skills was not an adequate test to identify children with SLI, since a high proportion of children with SLI received a score within normal limits. Furthermore, a great variability in pre-literacy skills was found in children with SLI in Cabell et al. (2010), where the quality of home environment and attentional skills of the child were found to be better predictors of pre-literacy skills than language skills. A closer examination of the pattern of response on these items, in relation to attention and home environment, would have to be performed to confirm if this explanation is adequate.

As hypothesized, both groups of TD bilingual children received scores closer to the scores of the TD monolingual children for this subscale, which targeted elements that were less affected by a reduced exposure due to bilingualism. This finding is congruent with the proposals by Paradis (2010) and Roeper (2011) presented in the previous section, which stipulated that some elements learned in one language can be used to facilitate learning of the other language. In addition to those proposals, Cummins's (2000) interdependence hypotheses regarding the transferability of skills from one language to another in sequential bilinguals, notably in literacy-related skills, can also be mentioned. Indeed, it is possible that children can rely on knowledge gained from one language to advance his skills in the other language for the domains targeted in this language-general subscale. This would reduce the consequences of receiving a lower input in

each language. However, it has to be mentioned that bilingual children were exposed to a variety of languages and lived different types of bilingual contexts. This is an important factor contributing to the variability in the children's scores, as transfers between languages have been shown to vary depending on which language combinations a child is learning (for a review, see Bialystok, 2012).

Finally, these results could also be interpreted in light of the very speculative discussion presented earlier regarding processing skills, critical mass, and exposure levels. Indeed, it is possible that the critical mass required for a language skill to be acquired is more easily reached for the language domains included in the language-general subscale. If the elements targeted by the items from this subscale required a low exposure before it can be considered acquired, children with SLI or bilingual children without strong processing skills would more rapidly obtained a score within normal limits. Another possible consequence of a low critical mass is that the elements would more rapidly show improvement following a language intervention. Indeed, the children with SLI in this study had been assessed and their parents received advice from a Speech-Language Pathologist in the past, as to be accepted in the caseload of a rehabilitation center, the child had to have been referred after a language evaluation. If the parents had put into practice some of the recommendations, it is expected that a skill targeting an element with a less stringent requirement will be acquired earlier.

Another difference between the two scales was that the skills targeted in each one seem to follow a different developmental pattern. Visual inspection of Figure 2 showed that for the language-specific subscale, the rate of acquisition was very rapid with a very low between-child variability in the early years. The rate of acquisition was much lower around 39 months, when many children reached ceiling, even if some variability was found up to 60 months. In contrast, in the language-general subscale, a wider variability was found across the whole age range. The

rate of acquisition for children under 39 months was less rapid, and the difference with the rate of acquisition of the older children was much less pronounced, leading to a ceiling effect in the language-general subscale occurring only around 54 months. The most likely explanation is that the skills targeted in each subscale are not acquired at the same age. The language-specific subscale targets vocabulary and syntax, more specifically elements that are emerging early in development such as first words and 2- or 3- word combinations. In contrast, many items in the language-general subscale target later developing skills, such as narrative skills and metalinguistic abilities. Because the items included in the MilBec were not originally selected to create these two subscales, no attempt was made to match the subscales in terms of age of acquisition of the items.

Other possible factors can be advance to account for the unexpectedly high variability of scores on the language general subscale. Notably, the parental education level, the inclusion in the home of educational activities and the familial income (Davis-Kean, 2005; Melhuish et al., 2008; Mercy & Steelman 1982) were found to have an impact on the child's language skills. Similarly, parenting behaviours were also found to have an impact on the child's language skills: these include the manner in which the parent interacts with the child and how the child is spoken to (Landry, Smith, Miller-Loncar, & Swank, 1997; Rindermann & Baumeister, 2015), is influenced by the number of persons living in the same house, known as household crowding (Evans, Maxwell, & Hart, 1999). To determine whether these elements are really influencing the children's score on the MilBec will have to be investigated in a follow-up study. Why these more environmental elements would have a greater influence on the language-general subscale, and not the language subscale, is also a question left to be answered. It might be that some of the skills included in the language-general subscale are learned differently from the skills included in the language-specific subscale. For example, whenever a child hears someone speak it requires the

use of vocabulary and syntax; incidental learning would be sufficient to gain an appropriate level of skill for TD children for these language domains. However, other skills, such as the macrostructure of narration and preliteracy skills, might be learned only if the child is experiencing a particular type of language experience, such as book reading. If the child is not sufficiently exposed to such experiences, those skills would develop at a lowered pace.

To conclude, these results concurred with the large body of literature documenting the widespread difficulties of children with SLI (for a review, see Weiss & Paul, R., 2010), with almost the totality of the children with SLI having failed both language subscales. In contrast, children from both groups of TD bilingual children were not more likely to fail the language-general subscale then TD monolingual children, and only a slightly higher proportion of bilingual children with a low level of exposure failed the language-specific subscale. This difference in the pattern of performance could help to correctly rule out SLI in bilingual children: only when a child failed the two subscales that it would be a strong evidence of the presence of SLI. These results indicated that the MilBec is a promising tool to be used with the bilingual population, particularly when the child's performance across the two subscales is considered.

Limitations of the Study and Future Research

Given the exploratory nature of this study, the principal limitations are the low number of participants in the groups other than the TD monolingual children and the large age range. Also, the fact that the bilingual children are from very different background in terms of number and type of languages spoken may limit the generalizability of the findings. The fact that most children were not directly tested to ascertain that they have a typical language development is also an element limiting the interpretation of the results. Finally, as previously mentioned, the MilBec's items were not selected to test these hypotheses, and were thus not matched in terms of expected age of acquisition.

In a follow-up study, a detailed analysis of the items should also be performed to better understand the reasons behind the unexpectedly large variability of scores found for the language general subscale. This high variability is indicative that element(s) other than age is influencing the children's performance. Possible explanations include that some items might be hard to evaluate by parents, that some items are more affected by the quality of language input received by the child, or other elements related to socio-economic status. Some items might also be sensitive to the children's characteristics other than their language skills, such as attention skills. These elements, which were not controlled for in the current study, should be taken into consideration in future researches in order to shed light on their possible role in language acquisition and how parent perceives their child's language skills.

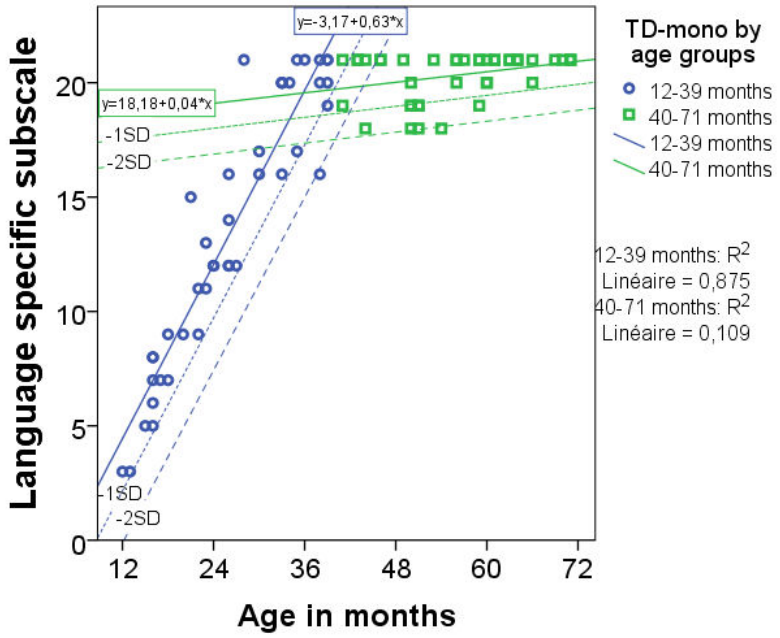
Conclusion

The MilBec is a new parent questionnaire to screen for language impairment in children between 12 and 71 months. The preliminary data collected on bilingual children indicates that it is a promising test to use with this population, as they obtained scores similar to those of monolingual children, even though a higher proportion of children with lower exposure to French failed the language-specific subscale. In contrast, all of the monolingual children with SLI failed the language-specific subscale, and all but one failed the language-general subscale. The results are thus indicative that the MilBec is a promising tool to correctly rule out SLI in TD bilingual children having received different levels of exposure to French. More research needs to be undertaken to better understand the variables influencing the child's performance on the MilBec, for both monolingual and bilingual children. The sensitivity of the MilBec to detect SLI in a bilingual population should be investigated in another study.

Acknowledgments

We want to thank all the parents and children who participated in the study.

a)



b)

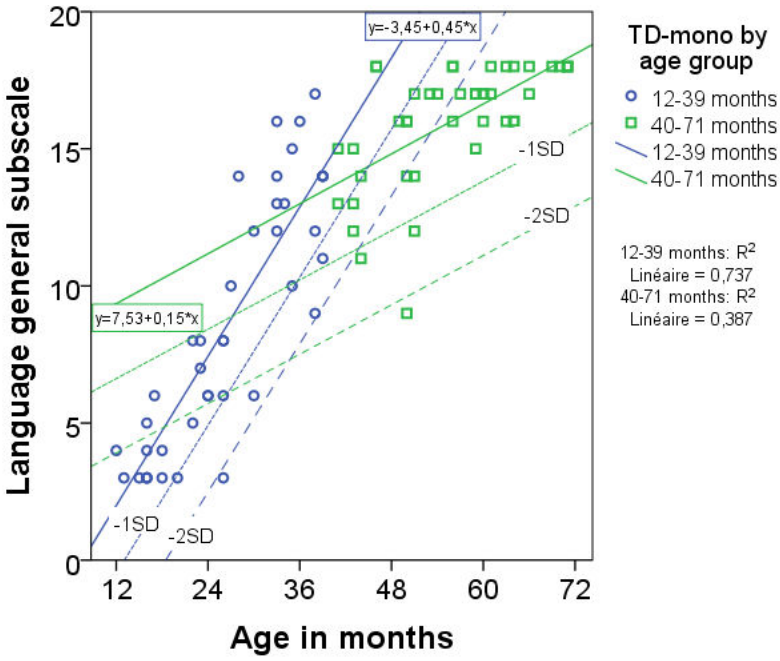
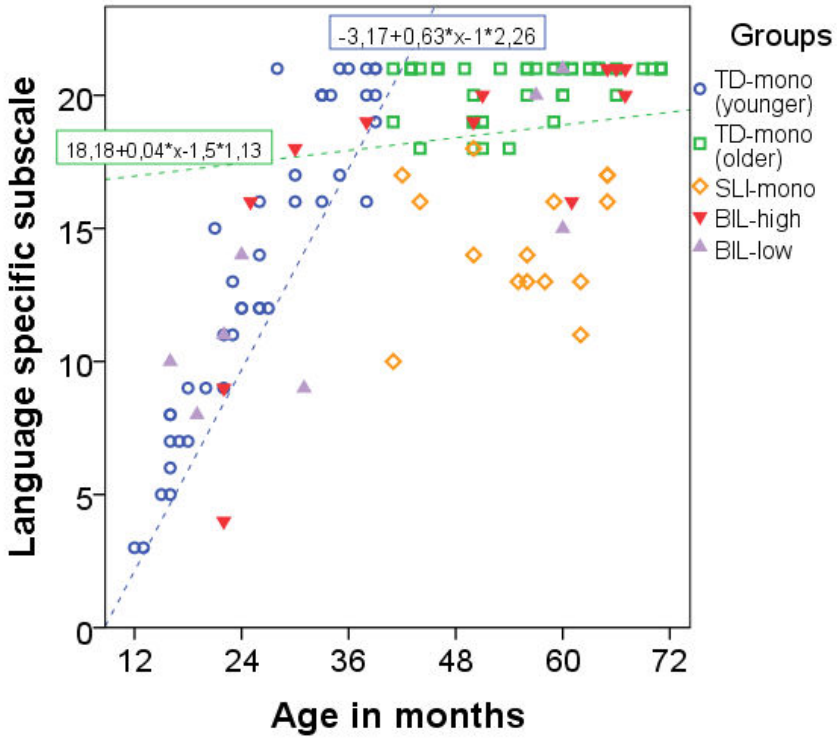


Figure 1. Scatterplot of the score as a function of age in months for TD monolingual children between 12 and 39 months and between 40 and 71 months, for a) the language specific subscale, and b) the language general subscale.

The linear regression lines and its formula are shown for each age group in full lines; the -1 SD and -2 SD reference lines based on the average standard deviation of each group in dotted lines.

a)



b)

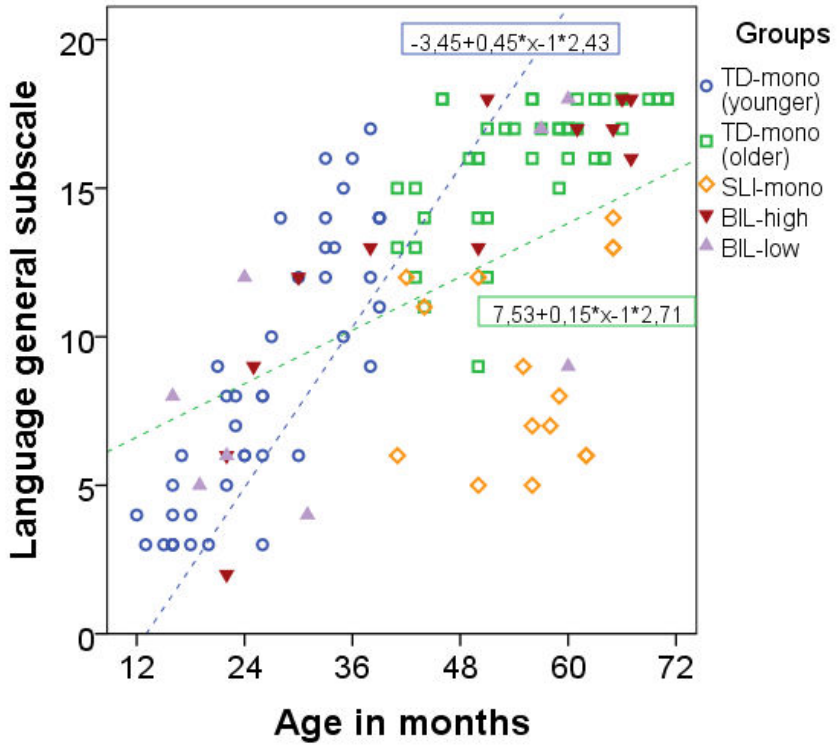


Figure 2. Scatterplot of the scores as a function of age in months for the four groups, for a) the language specific subscale, and b) the language general subscale.

The equations of each reference line used as cut-offs are shown beside its corresponding doted

lines. The four groups are: TD monolingual children (TD-mono), separated into two age groups, monolingual children with SLI (SLI-mono), TD with a dominant exposure to French (BIL-high), TD children with a dominant exposure to a language other than French (BIL-low)

Table 1: List of items in the MilBec per language domain, divided into the language-specific and language-general subscales.

Domains	Number of items: Item number in the MilBec
Language-specific items: 21 items	
Vocabulary	8 expressive items: #5-#6-#10-#20-#21-#27-#29-#31
	2 receptive items: #3-#9
vocabulary and syntax	4 expressive items: #15-#25-#28-#30
	3 receptive items: #4-#7-#22
Syntax	4 expressive items: #11-#17-#18-#23
Language-general items: 18 items	
narrative skills (excluding microstructure)	3 items: #24-#35-#36
use of gesture	1 item: #8
language use	2 items: #16-#19
phonology / articulation	5 items: #2-#12-#13-#14-#26
meta-linguistics	2 items on meta-phonology: #33-#34 1 item on meta-morphology: #32
pre-literacy (print-knowledge)	3 items: #37-#38-#39
judgment of normality	1 item: #1

Table 2: Background characteristics of the participants for each group.

The number of participants, mean, standard deviation (SD), and range for the participants' background characteristics age, maternal education and percentage of exposure to French for the TD monolingual children group (TD-mono), monolingual children with SLI group (SLI-mono), TD with a dominant exposure to French group (BIL-high), TD children with a dominant exposure to a language other than French group (BIL-low).

	Nb of participants	Age in months: mean (SD), range	Maternal education: mean (SD), range	French exposure: mean (SD), range
TD-mono	85	41.1 (17.2) 12-71	16.3 (2.5) 11-24	99.7% (1.0%) 95%-100%
SLI-mono	15	55.3 (8.3) 41-65	12.2 (2.7) 8-18	99.7% (1.3%) 95%-100%
BIL-high	12	47.0 (18.6) 22-67	18.08 (4.7) 11-25	73.4% (9.1%) 60%-85%
BIL-low	8	36.1 (19.4) 16-60	14.3 (1.8) 11-17	35% (18.9%) 8%-54%

Table 3: The average standard deviation for the younger group and the older group of monolingual TD children for the language specific subscale and the language general subscale.

	Younger group	Older group
Language specific subscale	2.26	1.13
Language general subscale	2.43	2.71

Note: The averages for the younger groups are based on the 5 groups of 6-month intervals between 12 and 41 months. For the older groups, they are based on the 2 groups of 6-month intervals between 42 and 53 months.

Table 4: The number of children having passed or failed each subscale, per age groups and language groups, using the cut-off line presented in Figure 2.

The children are separated by language groups: TD monolingual children (TD-mono), monolingual children with SLI (SLI-mono), TD with a dominant exposure to French (BIL-high), TD children with a dominant exposure to a language other than French (BIL-low).

	Group	Nb of participant	Language specific subscale results		Language general subscale results	
			Failed	Passed	Failed	Passed
12-39 months	TD-mono	42	2	40	5	37
	SLI-mono	0	-	-	-	-
	BIL-high	5	1	4	1	4
	BIL-low	5	1	4	1	4
40-71 months	TD-mono	43	4	39	3	41
	SLI-mono	15	15	0	14	1
	BIL-high	7	1	6	0	7
	BIL-low	3	1	2	1	2

Table 5: The percentage of all children having failed the screening, for the language specific and the language general subscales.

The nonparametric khi-square test used test-values corresponding to the percentage of TD monolingual children (TD-mono) having passed/failed the screening to determine if the proportion is different for monolingual children with SLI (SLI-mono), TD with a dominant exposure to French (BIL-high), TD children with a dominant exposure to a language other than French (BIL-low).

		Failed screening		
	Groups	Nb of participant	Percentage	Nonparametric khi-square
Language specific subscale	TD-mono	6/85	7.06%	-
	SLI-mono	15/15	100%	Can't be performed
	BIL-high	2/12	16.67%	$\chi2(1) = 1,688$ at $p = .194$
	BIL-low	2/8	25.00%	$\chi2(1) = 3,924$ at $p = .048$
Language general subscale	TD-mono	8/85	9.41%	-
	SLI-mono	14/15	93.33%	$\chi2(1) = 123,933$ at $p = .000$
	BIL-high	1/12	8.33%	$\chi2(1) = .016$ at $p = .898$
	BIL-low	2/8	25.00%	$\chi2(1) = 2,281$ at $p = .131$

Table 6: The percentage of all children having failed both subscales

The nonparametric khi-square test used test-values corresponding to the percentage of TD monolingual children (TD-mono) having passed/failed the screening to determine if the proportion is different for monolingual children with SLI (SLI-mono), TD with a dominant exposure to French (BIL-high), TD children with a dominant exposure to a language other than French (BIL-low).

Groups	Nb of participant	Percentage	Nonparametric khi-square
TD-mono	5/85	5.88%	-
SLI-mono	14/15	93.33%	$\chi 2(1) = 207{,}293$ at $p = .000$
BIL-high	1/12	8.33%	$\chi 2(1) = .131$ at $p = .718$
BIL-low	2/8	25.00%	$\chi 2(1) = 5{,}285$ at $p = .022$

General Conclusion

Typical language development is characterized by a high variability between children, particularly when the children are raised in a bilingual environment. However, about 7% of children (Tomblin et al, 1997) fail to acquire language normally, even though there is no known cause for these difficulties; these children are said to have specific language impairment (SLI). There are many long-term negative effects associated with this impairment, notably emotional and behavioural problems, as well as difficulties with peer relationships (Conti-Ramsden et al., 2013). Children with SLI should thus receive appropriate help from an early age, since such interventions have been shown to be effective (Nelson et al. 2006; Wallace et al., 2015) and can provide long-term benefits (Vlassopoulos et al., 2014).

The use of parent-questionnaires to screen for language difficulties could be an efficient and cost-effective way to identify children, since they can be used without special training and require only a limited amount of time to administer and score. Furthermore, parent questionnaires provide insight into the child's language abilities in a different way than direct testing can, as they are based on the parent's daily observations of the child's skills in a variety of contexts. However, for such tests to be trustworthy, their ability to correctly identify typically developing children as "not at risk" and children with a specific language impairment as "at risk" have to be verified. As these screeners are used to identify children in need of a more complete language assessment, it is particularly important that they have a high diagnostic accuracy.

Early identification of children with SLI is a hard thing to accomplish. Firstly, the early manifestations of persistent SLI are far from being fully known. Indeed, recent findings indicated that vocabulary and first word combinations, the elements traditionally considered for early referral to language services, identified only a limited number of children who will have a persistent SLI (Rescorla, 2011; Poll & Miller, 2013). Secondly, for French-speaking children in

Quebec (Canada), there are few validated tests, which means that speech-language pathologists are limited in the information they can gather regarding the child's skills, as both informal and formal testing should be performed to obtain a clearer view of the child's strengths and weaknesses (Paul, R. & Norbury, 2012). Furthermore, ruling out language impairment is particularly challenging for children raised in a bilingual environment. Indeed, these children receive less input in each language than age-matched monolingual peers, because their language exposure is distributed between their two languages. This may cause them to have language skills that are similar to what is observed for younger children (for a review, see Kohnert & Medina, 2009) when the level of exposure in the language is low (Elin Thordardottir, 2015a). For an altogether different reason, children with SLI also exhibit language skills that resemble the skills of younger peers: for them, this similarity is related to difficulties in acquiring language (e.g. Elin Thordardottir & Namazi, 2007).

The first goal of this thesis was to document the psychometric properties of a new parent-questionnaire to screen for SLI, the Milestones en Français du Québec (MilBec: Paul & Elin Thordardottir, 2010). The second one was to determine its ability to rule in or rule out moderate to severe SLI in monolingual children. The third goal was to investigate the performance of bilingual children with either a high or low level of exposure to French, when compared to the performance of monolingual children with and without SLI. Together, these goals also have theoretical importance, as the findings can be used to inform on the nature of SLI and the role of input in language acquisition.

The first two chapters documented the MilBec's characteristics as a language screening tool, as its adequacy had to be established before exploring how it can be used to answer the main question. More specifically, Chapter 1 presented the adaptation procedures of the parent-questionnaire by Luinge et al. (2006) leading to the creation of the MilBec, as well as a

longitudinal study presenting the first support of the MilBec's developmental sensitivity, concurrent validity and reliability. Chapter 2 provided additional support for the test's concurrent validity and preliminary norms for children between 12 and 71 months. It also presented an evaluation of the MilBec's diagnostic accuracy, confirming that the MilBec can adequately rule in or rule out moderate to severe language impairment in monolingual children between 36 and 71 months. An investigation of the usefulness of two items in particular considered potential clinical markers of SLI, namely subject and oblecjt clitic use, was also investigated. This study pointed out the limited value of these markers at ruling out a moderate-to-severe SLI. Indeed, the results showed that only a subgroup of monolingual children with SLI was reported to not have acquired these elements although they should have based on their chronological age. The lack of subject use after age 27 months or the lack of object clitic use after the age of 31 months were strong indicators of a moderate-to-severe SLI for monolingual children. However, when a child was reported using them, it was not informative of his or her language status, as a high proportion of children with SLI were reported using them.

These findings on the performance of children with SLI for the two clinical markers highlighted the heterogeneity of the manifestations of SLI in French-speaking children. From a theoretical point of view, these findings could be used against the deviant view of language impairment, since only a small portion of children would be identified using them. If children with SLI lacked the innate knowledge necessary to acquire these elements, all children with the disorder should experience these difficulties. However, it could be that these markers identify only a subgroup of the children with SLI; other markers would have to be found to identify the other subgroups. This idea of separating children with SLI into subgroups has been suggested by different authors, such as Rapin and Allen (1983, 1987) who proposed six subgroups. Similar subgroups were found by Conti-Ramsden and Botting (1999) using a cluster analysis based on

the scores of 242 children, and the same subgroups were found in the same children a year later. However, 45% of the children were reported not to be classified in the same subgroup at time 1 and time 2. Such a variation in the manifestations of SLI within a child is not coherent with the innate view of language associate with the deviant view of language impairment.

Finally, Chapter 3 explored how the MilBec can inform on the particularity of language acquisition of bilingual children, having either a low level or a high level of exposure to French, when compared to monolingual children with and without SLI. To do so, the items of the MilBec were separated into two subscales based on a literature review of the effect of bilingualism on the children's performance on the various language domains included in the MilBec. The language-specific subscale, contains items targeting a language domain expected to be affected by a reduced exposure to language caused by bilingualism, namely vocabulary and syntax. The second one, the language-general subscale, contains the items targeting a language domain not expected to be affected by bilingualism, namely microstructure of narration, articulation/phonology, and metalinguistic skills. In general, bilingual children receiving either a high or a low exposure to French performed similarly to TD monolingual children. However, the results indicated that some bilingual children with an exposure to French inferior to 55% should be expected to obtain a lower performance on the MilBec, but only on the language-specific subscale. In contrast, monolingual children with SLI would obtain low scores on both subscales. Finally, an unexpectedly high variability in the children's performance on the language-general subscale, which was observed for all groups, had been found. This finding pointed to the possible influence of other factors other than age, language exposure and the presence or not of a language impairment on the children's performance. Additional studies should be performed to investigate what these factors are.

These results added to the growing body of evidence showing that amount of exposure

plays a role in bilingual children's performance on language tests: only those who received a low level of exposure to French obtained lower scores than monolingual peers. Furthermore, lower scores obtained by these children occurred mainly on vocabulary and syntax, not necessarily on all of the language domains. This knowledge is important for clinicians working with bilingual children, as a more precise view of the effect of bilingualism on language acquisition should help them set expectations that are more appropriate. In contrast, monolingual children with SLI have lower skills in various language domains, as shown with their lower performance on both subscales of the MilBec.

The work presented in this thesis provided validation data and preliminary norms for a new screening tool for language impairment in French-speaking children. Further studies should be performed to collect additional normative data to provide norms based on much larger groups; the ability of the MilBec to identify children with a milder form of language impairment should also be investigated. Further investigation of the performance of bilingual children with different levels of exposure to French should also be performed to determine how to best interpret their scores on the MilBec.

References

Attia, J. (2003). Moving beyond sensitivity and specificity: Using likelihood ratios to help interpret diagnostic tests. *Australian Prescriber, 26*(5), 111-113. doi: 1,018,773/austprescr.2003.082

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language, 27*, 521-559. Retrieved from http://journals.cambridge.org/action/displayJournal?jid=JCL

Bates, E., Bretherton, I., & Snyder, L. S. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge, United Kingdom: Cambridge University Press.

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. doi: 102,167/beb392.0

Beitchman, J. H. & Brownlie, E. (2005-2010). Language development and its impact on children psychosocial and emotional development. In *Encyclopedia on Early Childhood Development*. Retrieved from http://www.child-encyclopedia.com/Pages/PDF/Beitchman-BrownlieANGxp_rev.pdf

Bialystok, E. (2012). The impact of bilingualism on language and literacy development. In T. K. Bhatia, and W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (2nd ed.) (pp. 624-648). Malden, MA: Blackwell Publishing. doi: 101,002/9781118332382

Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Development Perspectives, 9*(2), 117-121. doi: : 10.1111/cdep.12116

Bialystok , E. (2016). Bilingual education for young children: review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*, online.

http://dx.doi.org/10.1080/13670050.2016.1203859

Bialystok, E. & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language, 109*(2), 93-100. doi:10.1016/j.bandl.2008.09.001

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*(4), 525-531. doi: 101,017/S1366728909990423

Billard, C., Livet, M. O., Motte, J., Vallée, L., Gillet, P., Galloux, A., Piller, A. G., & Vol, S. (2000). BREV: A new scale for the evaluation of cognitive function in school-age and preschool-age children. *Archives de Pediatrie, 8*(5), 545-552. doi: 101,016/S0929-693X(00)00262-1

Bishop, D. V., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37*(4), 391-403. doi: 101,111/j.1469-7610.1996.tb01420.x

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., ... & Vyt, A. (2004). Cross-Linguistic Analysis of Vocabulary in Young Children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75*(4), 1115-1139. doi: 101,111/j.1467-8624.2004.00729.x

Bouchard, C., Trudeau, N., Sutton, A., Boudreault, M. C., & Deneault, J. (2009). Gender differences in language development in French Canadian children between 8 and 30 months of age. *Applied Psycholinguistics, 30*(4), 685-707. http://dx.doi.org/10.1017/S0142716409990075

Boudreault, M.-C.,Cabirol, É.-A. Trudeau, N. Poulin-Dubois, D., & Sutton, A. (2007). Les Inventaires MacArthur du développement de la communication: Validité et données

normatives préliminaires. *Canadian Journal of Speech-Language Pathology and Audiology, 31*(1), 27-37. Retrieved from http://cjslpa.ca/detail.php?ID=984

Boudreau, D. M. & Hedberg, N. L. (1999). A comparison of early literacy skills in children with specific language impairment and their typically developing peers. *American Journal of Speech-Language Pathology, 8*(3), 249-260. doi:10.1044/1058-0360.0803.249

Brevet, C. & Chambournac, M.-L. (2009). *Validation de l'échelle de compréhension RDLS-III auprès d'enfants de maternelle : Intérêt pour l'enfant autiste* (Unpublished Thesis). Université Paris VI : France.

Brunet, G., Cossette, G., Cousineau, D., & Lemieux, D. (2015). *L'ABCdaire du suivi prériodique de l'enfant de 0 à 5 ans.* Centre de Pédagogie Appliquée aux Science de la Santé et la Faculté de Médecine de l'Université de Montréal. Retrieved from www.cpass.umontreal.ca/formation/auto-apprentissage/outils-de-soutien-a-la-pratique

Buschmann, A., Jooss, B., Rupp, A., Dockter, S., Blaschtikowitz, H., … & Pietz, J. (2008). Children with developmental language delay at 24 months of age: Results of a diagnostic work-up. *Developmental Medicine & Child Neurology, 50*(3), 223-229. doi: 101,111/j.1469-8749.2008.02034.x

Cabell, S. Q., Lomax, R. G., Justice, L. M., Breit-Smith, A., Skibbe, L. E., & McGinty, A. S. (2010). Emergent literacy profiles of preschool-age children with Specific Language Impairment. *International Journal of Speech-Language Pathology, 12*(6), 472-482. doi: 103,109/17549507.2011.492874

Campbell, J. M., Brown, R. T., Cavanagh, S. E., Vess, S. F., & Segall, M. J. (2008). Evidence-based assessment of cognitive functioning in pediatric psychology. *Journal of Pediatric Psychology, 33*(9), 999-1014. doi: 101,093/jpepsy/jsm138

Carrow-Woolfolk (1985). *Test for Auditory Comprehension of Language-Revised (TACL-R).*

Austin, TX: Pro-Ed.

Carscadden, J., Corsiatto, P., Ericson, L., Illchuk, R., Esopenko, C., … & Oddie, S. D. (2010). A pilot study to evaluate a new early screening instrument for speech and language delays. *Canadian Journal of Speech-Language Pathology and Audiology, 34*(2), 86-95. Retrieved from http://cjslpa.ca/detail.php?ID=1017

Catts, H. W. (1993). The relationship between speech language impairments and reading disabilities. *Journal of Speech and Hearing Research, 36*, 948–958. doi:10.1044/jshr.3605.948

Chevrie-Muller, C., Goujard, J., & Goujard, A.-M. (1994). *Questionnaire langage et comportement – 3 ans ½ : Observation par l'enseignant en petite section de maternelle*. Paris, France : PDG communication.

Chevrie-Muller, C. & Plaza, M. (2001). N*ouvelles Épreuves pour l'Examen du Langage (N-EEL)*. Paris, France : Les Éditions du Centre de Psychologie Appliquée.

Chiang, P. & Rvachew, S. (2007). English-French bilingual children's phonological awareness and vocabulary skills. *Canadian Journal of Applied Linguistics, 10*(3), 293-308. doi : 101,044/0161-1461(2007/006)

Childers, J. B., Vaughan, J., & Burquest, D. A. (2007). Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *Journal of Child Language, 34*(02), 199-225. http://dx.doi.org/10.1017/S0305000906007835

Cohen, J. (1992). *A power primer. Psychological Bulletin, 112*(1), 155-159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Coquet, F. & Maetz, B. (1997). Le DPL 3 dépistage langage et prévention à 3 ans : Un nouvel outil de repérage des troubles du langage oral chez l'enfant de 3 ans 6 mois. *Rééducation Orthophonique, 35*(190), 153-162.

Conti-Ramsden, G. & Botting, N. (1999). Classification of children with specific language

impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42*, 1195-1204. doi:10.1044/jslhr.4205.1195

Conti-Ramsden, G., Mok, P. L., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities, 34*(11), 4161-4169. doi:10.1016/j.ridd.2013.08.043

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon, United Kindom: Multilingual Matters.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294-304. doi: 101,037/0893-3200.19.2.294

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders, 43*(4), 361-389. doi: 101,080/13682820701546854

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2010). Three profiles of language abilities in toddlers with an expressive vocabulary delay: Variations on a theme. *Journal of Speech, Language, and Hearing Research, 53*(3), 699-709. doi:10.1044/1092-4388(2009/07-0245)

Dickinson, D. K., McCabe, A., Clark–Chiarelli, N., & Wolf, A. (2004). Cross-language transfer of phonological awareness in low-income Spanish and English bilingual preschool children. *Applied Psycholinguistics, 25*(3), 323-347. doi: 10.1017.S0142716404001158

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848-856. doi: 101,111/jcpp.12378

Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders, 47*(5), 542-555.doi: 101,111/j.1460-6984.2012.00164.x

Dunn, L. & Dunn, L. (1981). *Peabody Picture Vocabulary Test- Revised (PPVT-R)*. Circle Pines, MN: Amercian Guidance Service.

Dunn, L.M. & Dunn, D.M. (1997). *Peabody picture vocabulary test, third edition*. Circle Pines, MN: American Guidance Service.

Dunn, L.M. & Dunn, D.M. (2007). *Peabody picture vocabulary test, fourth edition*. Minneapolis, MN: NCS Pearson.

Dunn, L.M., Theriault-Whalen, C.M., & Dunn, L.M. (1993). *Échelle de vocabulaire en image Peabody (EVIP)*. Canada : PsychCorp.

Duyme, M., Capron, C., & Zorman, M., (2010). *IDE forme langage: Inventaire du développement de l'enfant, pour les enfants agés de 15 mois à 5 ans 11 mois*, *Adaptation française du Child Development Inventory (CDI) d'Ireton, H.*. Available online at http:://www.cognisciences.com

Edwards, S., Fletcher, P., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1997). *Reynell developmental language scales III*. London: NFER-Nelson.

Elin Thordardottir (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language & Communication Disorders, 40*(3), 243-278. doi: 101,080/13682820410001729655

Elin Thordardottir (2010, May). *Identification et prévalence du trouble spécifique du langage chez les enfants francophones de cinq ans*. Oral presentation at the Congrès annuel de l'Association francophone pour le savoir-ACFAS, Université de Montréal, Québec, Canada.

Elin Thordardottir (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism, 15*(4), 426-445. doi: 101,177/1367006911403202

Elin Thordardottir (2014). The typical development of simultaneous bilinguals. In T. Grüter, and J. Paradis (Eds.), *Input and Experience in Bilingual Development* (pp. 141-160). Amsterdam, Netherlands: John Benjamins Publishing Company.

Elin Thordardottir (2015a). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology, 17*(2), 97-114. doi: 103,109/17549507.2014.923509

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

Elin Thordardottir (2016). Typical language development and Primary Language Impairment in French-speaking children. In J. Patterson and B. Rodriguez, (Eds.), *Multilingual perspectives on child language disorders*. Bristol, United Kingdom: Multilingual Matters.

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

Elin Thordardottir & Ellis Weismer, S. (1996). Language assessment via parent report: Development of a screening instrument for Icelandic children. *First Language, 16*(48), 265-285. doi: 101,177/014272379601604801

Elin Thordardottir & Ellis Weismer, S. (2002). Verb argument structure weakness in specific language impairment in relation to age and utterance length. *Clinical Linguistics & Phonetics,*

16(4), 233-250. doi: 101,080/02699200110116462

Elin Thordardottir, Kehayia, E., Courcy, A., Lessard, N., Majnemer, A., Mazer, B., Sutton, A., & Trudeau, N. (2003-2008). *Troubles spécifiques du langage: Développement d'outils d'évaluation et étude préliminaire de prévalence provinciale*. [Research project].

Elin Thordardottir, Kehayia, E., Lessard, N., Sutton, A., & Trudeau, N. (2010). Typical performance on tests of language knowledge and language processing of French-speaking 5-year-old. *Canadian Journal of Speech Pathology and Audiology, 34*(1), 5–16. Retrieved at http://cjslpa.ca/detail.php?ID=996

Elin Thordardottir, 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, and Hearing Research, 54*, 580-597. doi:10.1044/1092-4388(2010/09-0196)

Elin Thordardottir & Namazi, M. (2007). Specific language impairment in French-speaking children: Beyond grammatical morphology. *Journal of Speech, Language, and Hearing Research, 50*, 698-715. doi: 101,044/1092-4388(2007/049)

Elin Thordardottir, Rothenberg, A., Rivard, M. E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders, 4*(1), 1-21. doi: 101,080/14769670500215647

Evans, G.W., Maxwell, L.E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology, 35*, 1020–1023. http://dx.doi.org/10.1037/0012-1649.35.4.1020

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development, 71*(2), 310-322. doi: 101,111/1467-

8624.00146

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur communicative development inventories: User's guide and technical manual*. San Diego, CA : Singular Publishing Group.

Ferrand, P. (2000). *Protocole d'évaluation rapide (enfants de 3 ans 6 mois à 5 ans 6 mois)*. France: Ortho Édition.

Frank, I., Poulin-Dubois, D., & Trudeau, N. (1997). *Inventaires MacArthur du développement de la communication : Mots et énoncés*. Monréal, QC: Concordia University.

Frisk, V., Montgomery, L., Boychyn, E., Young, R., van Ryn, E., McLachlan, D., & Neufeld, J. (2009). Why screening Canadian preschoolers for language delays is more difficult than it should be. *Infants & Young Children, 22*(4), 290-308. doi: 10.1097/IYC.0b013e3181bc4db6

Gallinat, E., & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 57*(4), 1363-1382. doi:10.1044/2014_JSLHR-L-12-0363

Garcia, L. D., Paradis, J., Sénécal, I., & Laroche, C. (2006) Utilisation et satisfaction à l'égard des outils en français évaluant les troubles de la communication. *Revue Canadienne d'Orthophonie et d'Audiologie*, *30*(4), 239-249. Retrieved from http://cjslpa.ca/archive.php

Gathercole, V. C. (2002). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In D. K. Oller, and R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 220–254). Clevedon: Multilingual Matters.

Gaul Bouchard, M.-E., Fitzpatrick, E.M., & Olds, J. (2009). Analyse psychométrique d'outils d'évaluation utilisés auprès des enfants francophones. *Revue Canadienne d'Orthophonie et d'Audiologie, 33*(3), 129-139. Retrieved from http://cjslpa.ca/archive.php

Genesee, F. (2015). Myths about early childhood bilingualism. *Canadian Psychology, 56*(1), 6-15. http://dx.doi.org/10.1037/a0038599

Glascoe, F. P. (1991). Can clinical judgment detect children with speech-language problems? *Pediatrics, 87*(3), 317-322. Retrieved from http://pediatrics.aappublications.org/content/by/year

Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech, and Hearing Services in Schools, 30*(2), 196-206. doi:10.1044/0161-1461.3002.196

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test-Preschool, 2nd edition. *Language, Speech & Hearing Services in Schools, 40*(2), 150-160. doi:10.1044/0161-1461(2008/07-0049)

Gregoire, J. (1993). Le dépistage des troubles de langage en période préscolaire. *Canadian Family Physician, 39*, 856-863. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2379844/

Groupe coopératif en orthophonie région Laval, Laurentides, Lanaudière (1999). *Épreuve de compréhension de Carrow-Woolfolk, adaptation du TACL-R*. (Unpublished). Ordre des orthophonistes et audiologistes du Québec.

Grüter, T. (2005). Comprehension and production of French object clitics by children second language learners and children with specific language impairment. *Applied Psycholinguistics, 26*(03), 363-391. http://dx.doi.org/10.1017/S0142716405050216

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian Journal of Internal Medicine, 4(2), 627-635. Retrieved at http://www.caspjim.com/

Hamann, C. (2003). Phenomena in French normal and impaired language acquisition and their implications for hypotheses on language development. *Probus, 15*(1), 91-122. doi: 101,515/prbs.2003.006

Hamann, C. & Belletti, A. (2006). Developmental patterns in the acquisition of the complement clitic pronouns. *Rivista di Grammatica Generativa, 31*, 39-78. Retrieved from http://arcaold.unive.it/handle/10278/2487

Hamann, C., Ohayon, S., Dubé, S., Frauenfelder, U. H., Rizzi, L., Starke, M., & Zesiger, P. (2003). Aspects of grammatical development in young French children with SLI. *Developmental Science, 6*(2), 151-158. doi: 101,111/1467-7687.00265

Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing, 27*(4), 603-626. 101,177/0265532209355669

Heilmann, J., Ellis Weismer, S., Evans, J., & Hollar, C. (2005). Utility of the MacArthur—Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology, 14*(1), 40-51. doi:10.1044/1058-0360(2005/006)

Hemsley, G., Holm, A., & Dodd, B. (2006). Diverse but not different: The lexical skills of two primary age bilingual groups in comparison to monolingual peers. *International Journal of Bilingualism, 10*(4), 453-476. doi: 101,177/13670069060100040401

Hickmann, M. (1997). The acquisition of French as a native language: Structural and functional determinants in a crosslinguistic perspective. *Journal of Speech-Language Pathology and Audiology, 21*(4), 236-257. Retrieved from http://cjslpa.ca/archive.php

Hilaire-Debove, G., & Kern, S. (2013). Évaluation and développement de la macrostructure du récit oral chez les enfants avec ou sans troubles du langage. *Approche Neuropsychologique*

des Apprentissages chez l'Enfant, 124, 306-315. Available at http://www.anae-revue.com/

Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2014). Narrative abilities in subgroups of English language learners and monolingual peers. *International Journal of Bilingualism, 19*(6), 677-692. doi: 101,177/1367006914534330

Jacoby, W. G. (2000). Loess: a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies, 19*, 577-613. doi:10.1016/S0261-3794(99)00028-1

Jakubowicz, C., Nash, L., Rigaut, C., & Gerard, C. L. (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition, 7*(2-4), 113-160. doi : 101,207/s15327817la0702-4_3

Jaeschke, R., Guyatt, G. H., Sackett, D. L., Guyatt, G., Bass, E., Brill-Edwards, P., ... & Haynes, B. (1994). Users' guides to the medical literature: III. How to use an article about a diagnostic test B. What are the results and will they help me in caring for my patients? *Journal of the American Medical Association, 271*(9), 703-707. doi:10.1001/jama.1994.03510330081039

Justice, L. M, Bowles, R. P., & Skibbe, L. E. (2006). Measuring preschool attainment of print-concept knowledge: A study of typical and at-risk 3- to 5-year-old children using item response theory. *Language, Speech, and Hearing Services in Schools, 37*, 224–235. doi:10.1044/0161-1461(2006/024)

Justice, L. M. & Ezell, H. K. (2004). Print referencing: An emergent literacy enhancement strategy and its clinical applications. *Language, Speech, and Hearing Services in Schools, 35*(2), 185-193. doi: 101,044/0161-1461(2004/018)

Kaderavek, J. N. & Sulzby, E. (2000). Narrative production by children with and without specific language impairment: Oral narratives and emergent readings. *Journal of Speech, Language, and Hearing Research, 43*(1), 34-49.doi:10.1044/jslhr.4301.34

Kennison, S. M. (2014). *Introduction to language development*. Thousand Oaks, CA: SAGE

Publications.

King, T. M. & Glascoe, F. P. (2003). Developmental surveillance of infants and young children in pediatric primary care. *Current Opinion in Pediatrics, 15*, 624-629. Retrieved from http://journals.lww.com/co-pediatrics/toc/2003/12000

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41*(3), 627-641. doi:10.1044/jslhr.4103.627

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster,, S-A. H., Gustafsson, J-E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18*(1), 146-154. doi: 101,111/desc.12202

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. & Medina, A. (2009). Bilingual children and communication disorders: A 30-year research retrospective. *Seminars in Speech and Language, 30*(4), 219-233. doi: 101,055/s-0029-1241721

Kuder, S. J. (2013). *Teaching students with language and communication disabilities* (4e ed.). Upper Saddle River, NJ: Pearson. Education.

Landry, S. H., Smith, K. E., Miller-Loncar, C. L., & Swank, P. R. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Developmental Psychology, 33*(6), 1040-1053. doi 1040. 101,037//0012-1649.33.6.1040

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for speech and language delay: A systematic review of the literature. *Health Technology Assessment, 2*(9), 1-184.

http://dx.doi.org/10.3310/hta2090

Lefebvre, P., Girard, C., Desrosiers, K., Trudeau, N., & Sutton, A. (2008). Phonological awareness tasks for French-speaking preschoolers. *Canadian Journal of Speech-Language Pathology and Audiology, 32*(4), 158-167. Retrieved from http://www.cjslpa.ca/archive.php

Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, England: MIT press.

Luinge, M. R. (2005). *The language-screening instrument SNEL* (Unpublished doctoral dissertation). University Hospital Groningen, Groningen, Netherlands.

Luinge, M. R., Post, W. J., Wit, H. P., & Goorhuis-Brouwer, S. (2006). The ordering of milestones in language development of children from 1 and 6 years of age. *Journal of Speech, Language and Hearing Research, 49*, 923-940. doi:10.1044/1092-4388(2006/067)

Luk, G. & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605-621. http://dx.doi.org/io.1080/20445911.2013.795574

Maeder, C. & Roy, B. (2000). Detecting language disorders in 4-year-old French children. An application of the ERTL-4. *Child: Care, Health and Development, 26*(2), 129-136. doi: 101,046/j.1365-2214.2000.00174.x

Maillart, C. & Parisse, C. (2006). Phonological deficits in French speaking children with SLI. *International Journal of Language & Communication Disorders, 41*(3), 253-274. doi: 101,080/13682820500221667

Marchman, V. A. & Martinez-Sussmann, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research, 45*(5), 983-997. doi:10.1044/1092-4388(2002/080)

Marchman, V. A., Martínez-Sussmann, C., & Dale, P. S. (2004). The language-specific nature of grammatical development: evidence from bilingual language learners. *Developmental Science, 7*, 212–224. doi: 101,111/j.1467-7687.2004.00340.x

McGinty, A. S. & Justice, L. M. (2009). Predictors of print knowledge in children with specific language impairment: Experiential and developmental factors. *Journal of Speech, Language, and Hearing Research, 52*(1), 81-97. doi: 101,044/1092-4388(2008/07-0279)

Melhuish, E. D., Phan, M.B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggard, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*(1), 95-114. doi: 101,111/j.1540-4560.2008.00550.x

Mercy, J. A. & Steelman, L. C. (1982). Familial influence on the intellectual attainment of children. *American Sociological Review, 47*, 532-542. http://www.jstor.org/stable/2095197

Monetta, L. et tous les membres du REPAR – regroupement stratégique #4. (n.d.). *Fiches descriptives des outils validés et/ou normés en franco-québécois pour l'évaluation du langage et de la parole, de 1980 à 2014*. Retrieved from http://www.repar.veille.qc.ca/fr/outils/Orthophonie.html

Moyle, M. J., Ellis Weismer, S., Evans, J. L., & Linstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research, 50*, 508-528. doi:10.1044/1092-4388(2007/035)

Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). Screening for speech and language delay in preschool children: Systematic evidence review for the US Preventive Services Task Force. *Pediatrics, 117*(2), e298-e319. Retrieved from http://pediatrics.aappublications.org/

Olswang, L.B., Rodriguez, B., & Timler, G. (1998). Recommending intervention for toddlers with specific language learning difficulties: We may not have all the answers, but we know a lot. *American Journal of Speech-Language Pathology, 7*, 23-32. doi: 101,044/1058-0360.0701.23

OOAQ (2004). Guides et outils cliniques – Trouble primaire du langage/Dysphasie.

Orgassa, A. & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research, 24*(3), 333-364. doi: 101,177/0267658308090184

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 227-252. doi:10.1017/S0142716409990373

Paradis, J., Crago, M., & Genesee, F. (2003). Object clitics as a clinical marker of SLI in French: Evidence from French-English bilingual children. *Proceedings of the 27th Annual Boston University Conference on Language Development, 2*, 638-649. Retrieved from http://www.ualberta.ca/~jparadis/BilSLI_pro_bucld.pdf

Paradis, J., Tremblay, A., & Graco, M. (2014). French-English bilingual children's sensitivity to child-level and language-level input factors in morphosyntactic acquisition. In T. Grüter, and J. Paradis (Eds.), *Input and Experience in Bilingual Development* (pp. 161-180). John Benjamins Publishing Company: Amsterdam/Philadelphia.

Parisse, C. & Le Normand, M. T. (2001). Local and global characteristics in the development of morphosyntax by French children. *First Language, 21*(62), 187-203. doi: 101,177/014272370102106203

Paul, M. & Elin Thordardottir (2010). *Milestones en français du Québec (MilBec)* (Unpublished independent study). McGill University, Montreal, Canada.

Paul, R., & Norbury, C. (2012). *Language disorders from infancy through adolescence:*

Listening, speaking, reading, writing, and communicating (4e éd.). St-Louis, MI: Elsevier.

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: 101,111/j.1467-1770.1993.tb00174.x

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. doi:10.1044/1058-0360(2006/023)

Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., ... & Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology, 122*(2), 542-557. doi: 10.1037/a0031963

Piérart, B. (2014). Compétences langagières lexicales et phonologiques des dysphasiques: délai ou différence? *Développements, 15*(2), 4-15. doi: 10.3917/devel.015.0004

Plante, E. & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25*(1), 15-24. doi:10.1044/0161-1461.2501.15

Poll, G. H. & Miller, C. A. (2013). Late talking, typical talking, and weak language skills at middle childhood. *Learning and Individual Differences, 26*, 177-184. doi:10.1016/j.lindif.2013.01.008

Prathanee, B., Thinkhamrop, B., & Dechongkit, S. (2007). Factors associated with specific language impairment and later language development during early life: A literature review. *Clinical Pediatrics, 46*(1), 22-29. doi: 101,177/0009922806297153

Rapin, I. & Allen, D. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling* (pp. 155-184). New York, NY: Academic Press.

Rapin, I. & Allen, D. A. (1987). Syndromes in developmental dysphasia and adult aphasia. *Research Publications-Association for Research in Nervous and Mental Disease, 66*, 57-75.

Rescorla, L. (1989). The Language Development Survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders, 54,* 587-599. doi:10.1044/jshd.5404.587

Rescorla, L. (2011). Late talkers: Do good predictors of outcome exist? *Developmental Disabilities Research Reviews, 17*(2), 141-150. doi: 101,002/ddrr.1108

Rice, M. L. (2000). Grammatical symptoms of specific language impairment. In V. M. Bishop and L. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 17-34). New York, NY: Psychology Press. (Reprinted in 2014)

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., & Wexler, K. (1996). A phenotype of specific language impairment: Extended optional infinitives. In M.L. Rice (Ed.), *Toward a genetics of language* (pp. 215-237). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research, 38*(4), 850-863. doi:10.1044/jshr.3804.850

Rindermann, H. & Baumeister, A. E. (2015). Parents' SES vs. parental educational behavior and children's development: A reanalysis of the Hart and Risley study. *Learning and Individual Differences, 37*, 133-138. doi:10.1016/j.lindif.2014.12.005

Roeper, T. (2011). Strict Interfaces and three kinds of Multiple Grammar. In T. Kupisch, an E. Rinke (Eds.), *The Development of Grammar: Language acquisition and diachronic change: In*

honour of Jürgen M. Meisel (pp. 205-228). Amsterdam, Netherlands: John Benjamins Publishing.

Roid, G. & Miller, L. (1997). *Leiter international performance scale – Revised.* Wood Dale, IL: Stoelting.

Rondal, J.-A. (2001). Votre enfant apprend à parler, (4th edition). Hayen (Sprimont) : Margada.

Roulet, L. (2007). L'accord grammatical de genre dans la dysphasie de développement. *Psychologie Française, 52*(1), 55-70. doi:10.1016/j.psfr.2006.08.002

Rourke, L., Leduc, D., & Rourke, J. (2014). *Rourke baby record: Evidence-based infant/child health maintenance*. Retrieved at www.rourkebabyrecord.ca

Roy, B. & Maeder, C. (1996). *Épreuve de repérage des troubles du langage utilisables lors du bilan médical de l'enfant de 4 ans (ERTL-4)*. Nancy: Ortho Édition.

Royle, P. & Elin Thordardottir (2003). *Le grand déménagement* [French adaptation of the Recalling Sentences in Context subtest of the CELF–P] (Unpublished research tool). McGill University, Montreal, Quebec, Canada.

Rvachew, S., Marquis, A., Brosseau-Lapré, F., Paul, M., Royle, P., & Gonnerman, L. M. (2013). Speech articulation performance of francophone children in the early school years: Norming of the Test de Dépistage Francophone de Phonologie. *Clinical Linguistics & Phonetics, 27*(12), 950-968. doi: 103,109/02699206.2013.830149

Schmitt, C., Simoës, A., & Laloi, A. (2015). L'évaluation du langage oral chez les enfants bilingues au moyen d'une batterie standardisée pour monolingues. *Neuropsychiatrie de l'Enfance et de l'Adolescence, 67*(7), 431-436. doi:10.1016/j.neurenf.2015.04.005

Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Retrieved from University of Alberta Faculty of Rehabilitation Medicine website: www.rehabresearch.ualberta.ca/enni.

Schwartz, R. G. (2009). Specific language impairment. In R.G. Schwartz (Ed.) *Handbook of Child Language Disorders* (pp. 3-43). New York: Psychology Press.

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals* (4th Ed.). San Antonio, TX: The Psychological Corporation.

Slobin, D. I. (1969). *Report WP-22: Universals of grammatical development in children*. Language, Behavior Research Laboratory, University of California. Retrieved from files. eric.ed.gouv/fulltext/ED032543.pdf

Spaulding, T. J., Hosmer, S., & Schechtman, C. (2013). Investigating the interchangeability and diagnostic utility of the PPVT-III an PPVT-IV for children with and without SLI. *International Journal of Speech-Language Pathology, 15*(5), 435-462. http://dx.doi.org/10.3109/17549507.2012.762042

Sprenger-Charolles, L. & Serniclaes, W. (2003). Acquisition de la lecture et de l'écriture et dyslexie : revue de la littérature. *Revue Française de Linguistique Appliquée, 8*(1), 63-90. Retrieved from www.cairn.info/revue-francaise-de-linguistique-appliquee-2003-1-page-63.htm

Statistique Canada (2011a). *Caractéristiques linguistiques des Canadiens – Langue, recensement de la population de 2011*. Retrieved from www12.statcan.gc.ca

Statistique Canada (2011b). *Population selon la connaissance des langues officielles, groupe d'âge (total), répartition en pourcentage (2011), pour le Canada, les provinces et les territoires, et les régions métropolitaines de recensement et les agglomérations de recensement*. Retrieved from www12.statcan.gc.ca

Tervo, R. C. & Balaton, J. (1980). Developmental screening, well baby care in Saskatchewan family practice. *Canadian Family Physician, 26*, 1136-1140. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2383714/

Thal, D. J. & Tobias, S. (1992). Communicative gestures in children with delayed onset of oral expressive vocabulary. *Journal of Speech, Language, and Hearing Research, 35*(6), 1281-1289. doi:10.1044/jshr.3506.1289

Tomasello, M. (2000). Acquiring syntax is not what you think. In V. M. Bishop and L. Leonard, *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp.1-15). New York, NY: Psychology Press. (Reprinted in 2014)

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research, 40*(6), 1245-1260. doi: 101,044/jslhr.4006.1245

Trudeau, N., Frank, I., & Poulin-Dubois, D. (1999). *Inventaire MacArthur du développement e la communication : Mots et gestes*. Montréal.

Twisk, J. R. (2013). *Applied longitudinal data analysis for epidemiology: A practical guide*. Cambridge: Cambridge University Press.

Uccelli, P. & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*(3), 225-236. doi:10.1044/0161-1461(2007/024).

Vallée, L. & Dellatolas, G. (2005). *Recommandations sur les outils de repérage, dépistage et diagnostic pour les enfants atteints d'un trouble spécifique du langage*. Ministère de la Santé et des Solidatrité. Retrieved from www.sante.gouv.fr

Van Noort-van der Spek, I. L., Franken, M. C. J., & Weisglas-Kuperus, N. (2012). Language functions in preterm-born children: A systematic review and meta-analysis. *Pediatrics, 129*(4), 745-754. Retrieved from http://pediatrics.aappublications.org/content/129/4/745.short

Vlassopoulos, M., Petraki, E., Rotsika, V., Tsipra, I., Legaki, L., Richardson, C., … & Lazaratou,

H. (2014). A follow-up study of treated and untreated Greek adolescents with a history of specific developmental language disorder. *International Journal of Speech & Language Pathology and Audiology, 2*, 1-10. Retrieved from http://www.synergypublishers.com/

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., … & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics, 136*(2), e448-e462. Retrieved from http://pediatrics.aappublications.org/

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language, 108*(3), 175-183. doi:10.1016/j.bandl.2008.07.001

Webster, R. I., Majnemer, A., Platt, R. W., & Shevell, M. I. (2004). The predictive value of a preschool diagnosis of developmental language impairment. *Neurology, 63*(12), 2327-2331. http://dx.doi.org/10.1212/01.WNL.0000147472.33,670.B6

Weiss, D. & Paul, R. (2010). Delayed language development in preschool children. In J.S. Damico, N. Müller, and M.J. Ball (Eds.), *The Handbook of Language and Speech Disorders* (pp. 178-209). Oxford, UK: Blackwell Publishing Ltd. doi: 101,002/9781444318975.ch8

Wiig, E.H., Secord, W., & Semel, E.M. (1992). *Clinical Evaluation of Language Fundamentals - Preschool*. San Antonio, TX: The Psychological Corporation.

Wray, C., Norbury, C. F., & Alcock, K. (2016). Gestural abilities of children with specific language impairment. *International Journal of Language & Communication Disorders, 51*, 174–182. doi: 101,111/1460-6984.12196

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*(5), 516-524. doi: 101,111/jcpp.12009

Zourou, F., Ecalle, J., Magnan, A., & Sanchez, M. (2010). The fragile nature of phonological awareness in children with specific language impairment: Evidence from literacy development. *Child Language Teaching and Therapy, 26*(3), 347-358. doi: 101,177/0265659010369288

Zorman, M. & Jacquier-Roux, M (2002). BSEDS 5-6 : Un dépistage des difficultés de langage oral et des risques de dyslexie qui ne fait pas l'économie de la réflexion clinique. *Approche Neuropsychologique des Apprentissages chez l'Enfant, 66*, 48-55.