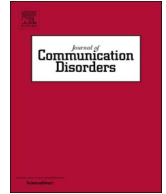


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis

Assessment of bilingual children: What if testing both languages is not possible?



Tessel Boerma*, Elma Blom

Utrecht University, The Netherlands

ARTICLE INFO

Keywords:

Language impairment
Bilingualism
Assessment
Parental questionnaire

ABSTRACT

Language delays of bilingual children can arise from language impairment (LI) but also from insufficient exposure to the target language. A reliable diagnosis of LI in bilingual children is therefore ideally based on the evaluation of both languages, as LI affects each language that is learned. However, due to the multitude of language combinations that are encountered in clinical practice, this is often not feasible. Bilingual norm-referencing may offer a solution, but the heterogeneity within the bilingual population makes it difficult to determine appropriate standards for every child. The present study examined an alternative approach to assessing both languages or using bilingual norm-referencing, aiming to assemble instruments that reduce bias against bilingual children. We used a four-group design, including monolingual and bilingual children with and without LI ($N = 132$), to first investigate the effects of LI and bilingualism on risks associated with a child's early language development and the prevalence of language problems in the family, as reported by parents. Second, we evaluated the diagnostic validity of these two indices, and, in addition, combined these with two unbiased language measures which we previously examined in isolation: a quasi-universal nonword repetition task and a narrative task. Results showed that the index of Early Language Development was a strong predictor of LI. In combination with the two direct language measures, it excellently identified the presence or absence of LI in and across monolingual and bilingual learning contexts.

Learning outcomes: As a result of this study, the reader will learn about an alternative approach to testing a bilingual child in both languages. The reader will recognize the importance of using unbiased measures for the identification of LI in a bilingual context, and, in addition, will appreciate the value of combining parental report with direct language measures.

1. Introduction

Bilingualism can be a complicating factor when diagnosing a child with language impairment (LI), potentially leading to misdiagnosis (Bedore & Peña, 2008; Grimm & Schulz, 2014; Salameh, Nettelbladt, Håkansson, & Gullberg, 2002; Smeets, Driessen, Elfering, & Hovius, 2009). When tested in one language, a bilingual child may score substantially lower than a monolingual peer (Thordardottir, Rothenberg, Rivard, & Naves, 2006), and the cause of this delay, e.g. an inborn LI or insufficient exposure to the language of testing, is difficult to determine (Kohnert, 2010). The International Association of Logopedics and Phoniatrics (IALP, 2011) and the American Speech-Language-Hearing Association (ASHA, 2004) therefore recommend that a bilingual child is assessed in both languages. This provides invaluable comprehensive information beneficial to a reliable diagnosis of LI (Paradis, 2016), as LI affects each language that is learned. Bilingual children with LI learn both their languages at a slower pace than children with a

* Corresponding author at: Department of Special Education, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands.
E-mail address: T.D.Boerma@uu.nl (T. Boerma).

typical language development (TD). In contrast, bilingual children with TD, especially children from a cultural minority, often show an unbalanced profile with a higher proficiency level in at least one language (Håkansson, Salameh, & Nettelbladt, 2003).

Collecting data on the first (L1) and second (L2) language of a child is feasible in situations involving a homogeneous bilingual population and well-documented languages, such as the population of Spanish-English speakers in the USA for whom bilingual assessment procedures have been developed (e.g., Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2014). However, due to the linguistic diversity in many countries, including the Netherlands where the present study was situated, assessing both languages of a bilingual child can often not be realized. Time restrictions, insufficient financial resources, and the lack of (culturally) appropriate instruments, bilingual speech-language pathologists and skilled interpreters are just a number of obstacles to overcome.

Several alternatives to assessing a bilingual child in both languages have been explored, for example using bilingual norm-referencing and adapting the norms of standardized tests that are typically used to diagnose LI in monolingual children. When relying on the evaluation of only one language of a bilingual child, Gathercole, Thomas, Roberts, Hughes, and Hughes (2013) emphasize the importance of defining appropriate bilingual norm categories which not only distinguish between monolinguals and bilinguals, but also acknowledge the variation within the bilingual population in terms of amount of exposure. In her guideline for the assessment of bilingual children, Thordardottir (2015) similarly argues to take amount of exposure into consideration, proposing to lower the cut-off criteria of standardized tests to varying extents, depending on whether a bilingual child's dominant or weaker language is tested.

Next to standardized language tests with adapted norm-referencing, Thordardottir (2015) also recommends to use other measures, such as a detailed case history and a nonword repetition task, to support the identification of LI in bilingual children. Paradis, Schneider and Sorenson Duncan (2013) tested the clinical value of such a combination of instruments with English L2 learners, including English standardized tests of morphology and vocabulary, an English nonword repetition task, and a case history based on parental report. As expected, the use of monolingual norm-referencing on the English measures resulted in a high percentage of bilingual TD children scoring below age expectations, indicating that these children were disadvantaged by these tests and were at risk for overidentification of LI. In contrast, a reliable diagnosis of LI was supported by the use of appropriate bilingual norm-referencing. The combination of instruments accurately differentiated the bilingual children with LI from a large group of bilingual children with TD, who were comparable in terms of amount of English input.

Standardized tests with appropriate norm adjustments may thus, in combination with other measures, offer a solution to the problem of assessing a bilingual child in both languages. However, given the heterogeneity within a bilingual population (e.g., in terms of linguistic background, language use, age of onset, amount of exposure, and language status), it remains challenging to determine which adjustment to the norm is most appropriate for which bilingual child. Especially in the case of sequential bilinguals, the search for a fitting standard is complex (Thordardottir, 2015). An alternative approach that eliminates the need of adapted norm-referencing could entail obtaining information on a child's language development with instruments that do not draw on language-specific abilities or knowledge, but, e.g., allow children to use skills that they have acquired in any language. In comparison with typical standardized language measures, this method of language assessment reduces the bias against children with differing language experiences. The present study evaluated the clinical value of such an alternative method. In accordance with Thordardottir (2015) and Paradis et al. (2013), we used a combination of instruments, including two indices based on a parental questionnaire and two direct language measures. The two direct language measures have previously been investigated in isolation and have already proven to be sensitive to effects of LI, but insensitive to effects of bilingualism (Boerma et al., 2015; Boerma, Leseman, Timmermeister, Wijnen, & Blom, 2016). However, less is known about the effects of LI and bilingualism on parental report, which is therefore the first focus of study.

1.1. The use of a parental questionnaire

Information about the quantity and quality of language input is of vital importance in determining whether the language problems of a bilingual child are caused by LI or by environmental factors, enabling a clinician to interpret a child's scores on direct language measures (Paradis, 2011; Tuller, 2015). A parental questionnaire or bilingual anamnesis can provide such information. Moreover, parental report is also shown to be a valid method to obtain information on the language development of a bilingual child, which is especially valuable when direct assessment in the child's both languages is not possible (Paradis, Emmerzael, & Sorenson Duncan, 2010; Tuller, 2015). Tuller (2015) developed a parental questionnaire in collaboration with members of the COST Action IS0804 *Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment*¹. This questionnaire can provide information on both the quantity and quality of language input as well as on potential risk factors of LI. Three risk factors, late language emergence, the prevalence of language problems in the family and poor current L1 skills, are discussed below.

Late language emergence, often indicated by the late production of a child's first word and word combination, can be the first diagnostic symptom of LI (Rice, 2007). It is therefore important to obtain information on when a child reached these basic milestones, and parents appear to be a valuable source for this. Paradis et al. (2010) showed that the timing of early milestones, as reported by parents, was the strongest indicator of LI in a sample of bilingual children. Moreover, the authors note that the early milestones of bilingual children were comparable to what had been previously documented for monolingual children. This finding is in line with other work that suggests that the timing of early milestones, in at least one language, is not affected by early exposure to two languages (De Houwer, 2009; Hoff et al., 2012). Nevertheless, a systematic investigation of the effects of LI and bilingualism on a

¹ A network that was set up thanks to funding of the European Cooperation in Science and Technology (COST) with the aim of coordinating research on linguistic and cognitive abilities of bilingual children with LI across different migrant communities (www.bi-sli.org).

child's early language development, as reported by parents, has not yet been carried out and is needed to confirm these findings.

The prevalence of speech, language, learning or literacy problems in a child's family is a second factor that has proven to be effective in differentiating TD and LI. For example, in the parental questionnaire data of Restrepo (1998), who studied Spanish-English speaking children, family history was one of the strongest discriminators. On the contrary, in the data of Paradis et al. (2010), family history did not contribute much to the identification of LI, even though significant group differences emerged. This discrepancy in the literature, next to a lack of research that examined whether monolingual and bilingual parents report differently on the prevalence of language problems in the family, calls for further research.

A third factor that is mentioned by Tuller (2015) is parents' evaluation of their child's current L1 skills, which may potentially be an alternative for a more direct measure of L1 abilities. Both Restrepo (1998) and Paradis et al. (2010) found that current L1 skills, as reported by parents, discriminated well between bilingual children with TD and LI. However, for the current study, this index is problematic. Paradis et al. (2010: 487) point out that scores on this index are more likely to be sensitive to differences in bilingual experiences than, e.g., early milestones in any language or family history, as the L1 skills of children are highly influenced by factors such as language status, language exposure and language use. The current study aimed to examine instruments that would be suited for clinical purposes irrespective of a child's language learning context, being relatively insensitive to differing language experiences. We therefore limited our scope to parental report on the early language development of a child and the prevalence of language problems in the family.

1.2. Unbiased language measures

Like the parental questionnaire, the two direct language measures included in the current study are both recently developed within the framework of the COST Action IS0804.

1.2.1. The Quasi-Universal Nonword Repetition Task

Following the recommendations of Thordardottir (2015), a nonword repetition task (NWRT), testing phonological short-term memory, was included in our study. It is well-known that children with LI have difficulty with repeating nonwords (for a meta-analysis, see Graf-Estes, Evans, & Else-Quest, 2007) and nonword repetition is therefore often suggested to be a clinical marker of LI in monolingual children (e.g., Conti-Ramsden, Botting, & Faragher, 2001). It is however questionable whether NWRTs are also appropriate to identify LI in a bilingual learning context. A NWRT may on the one hand be less biased against bilingual children than traditional standardized measures of vocabulary or grammar that tap into language knowledge, as nonwords are by definition not lexicalized (Campbell, Dollaghan, Needleman, & Janosky, 1997; Ellis Weismer et al., 2000; Thordardottir & Juliusdottir, 2013). On the other hand, nonwords commonly conform to the phonotactics of a specific language, often benefitting children with more experience in that particular language (e.g., Engel de Abreu, 2011; Sharp & Gathercole, 2013; Summers, Bohman, Gillam, Peña, & Bedore, 2010).

Research evaluating the clinical validity of NWRTs in bilingual contexts have produced mixed findings. Most studies reported inadequate diagnostic accuracy when a language-specific NWRT was administered in one of the child's languages (e.g., Boerma et al., 2015; Gutiérrez-Clellen & Simon-Cerejido, 2010; Kohnert, Windsor, & Yim, 2006; Windsor, Kohnert, Lobitz, & Pham, 2010), but Girbau and Schwartz (2008) presented levels of sensitivity and specificity above 80% for their Spanish NWRT, which is considered clinically accurate (Plante & Vance, 1994). In addition, the French NWRT of Thordardottir and Brandeker (2013) showed high sensitivity and nearly adequate specificity (79%) when monolingual cut-off scores were used with French-English bilingual children, while Armon-Lotem and Meir (2016) only arrived at these values for their Hebrew task with bilingual cut-off points. Thus, the use of a language-specific NWRT in only one language of a bilingual child gave valuable information about the child's language development in some cases, but not in others.

These mixed results may partly stem from differences in the characteristics of the nonwords across various NWRTs, as pointed out by Thordardottir and Brandeker (2013) and Graf-Estes et al. (2007) (for an elaborate discussion, see Boerma et al., 2015). While particular properties such as syllable length may enlarge differences in NWRT performance between children with LI and TD (Graf-Estes et al., 2007), other properties such as a complex syllable structure or high wordlikeness may enlarge differences between monolingual and bilingual children due to their reliance on experience with the target language (Engel de Abreu, Baldassi, Puglisi, & Befi-Lopez, 2013; Thordardottir & Brandeker, 2013). To maximally reduce the influence of language-specific experience on NWRT performance, a quasi-universal (Q-U) NWRT was developed by Chiat (2015) in collaboration with members of the COST Action IS0804. The Q-U NWRT avoids segmental and phonotactic structures that vary substantially across languages and instead incorporates features that are very common across languages, allowing children to use acquired knowledge from any language learned. Boerma et al. (2015) evaluated the clinical value of the Q-U NWRT and their findings were promising. Bilingualism did not affect performance on the Q-U NWRT and the instrument accurately identified LI in a monolingual and bilingual context, supporting the further use of this tool.

1.2.2. The Multilingual Assessment Instrument for Narratives

The third measure included in the present study is a narrative task that was analyzed on a macrostructural level. Macrostructure refers to the structural organization of a story consisting of several plot elements, such as a setting and a goal. Contrary to the linguistic expression of these plot elements, referred to as microstructure, it is assumed that the macrostructure of a story is universal and not language-specific (Berman & Slobin, 1994). Research suggests that bilingual children are able to apply knowledge about narrative macrostructure acquired in their L1 to their L2 (e.g., Pearson, 2002; Squires et al., 2014). Moreover, expressing a narrative's

macrostructure does not appear to be influenced by the amount of language-specific exposure that a child received, in contrast to narrative microstructure (Hipfner-Boucher et al., 2015). These findings may explain why equal performance of monolingual and bilingual children has been found on a narrative task analyzed on the macrolevel (Boerma et al., 2016; Hipfner-Boucher et al., 2015; Pearson, 2002). On the other hand, both monolingual and bilingual children with LI are known to have significant problems with understanding and expressing a narrative's macrostructure (e.g., Bishop & Donlan, 2005; Boerma et al., 2016; Dodwell & Bavin, 2008; Squires et al., 2014). These differential effects of LI and bilingualism on children's macrostructural narrative skills are promising in light of the challenging identification of LI in bilingual learning contexts.

Boerma et al. (2016) used the *Multilingual Assessment Instrument for Narratives* (MAIN; Gagarina et al., 2012) to assess macrostructural narrative comprehension and production skills of monolingual and bilingual children with and without LI. Like the Q-U NWRT, this instrument is developed within the COST Action IS0804 and is specifically aimed at bilingual children. Next to investigating the effects of LI and bilingualism on children's narrative skills, as reviewed above, Boerma et al. (2016) also examined the diagnostic validity of the MAIN. Results corresponded to previous work using other narrative tasks in monolingual samples (e.g., Pankratz, Plante, Vance, & Insalaco, 2007; Schneider, Hayward, & Dubé, 2006) and showed that the MAIN accurately identified a large number of children with TD and LI, irrespective of children's linguistic background. Due to heterogeneity within the population of children with LI, recommendations of previous work (Schneider et al., 2006) to combine a narrative task with other measures were reiterated by Boerma et al. (2016). In fact, it is often emphasized that a diagnosis of LI should never be solely based on individually administered language tests (Bishop & McDonald, 2009), and the use of multiple instruments is advocated (e.g., Thordardottir, 2015). The present study therefore examined the diagnostic validity of a combination of instruments, including the parental questionnaire, Q-U NWRT, and MAIN.

1.3. The present study

The current study evaluated an alternative approach to assessing a bilingual child in both languages or using bilingual norm-referencing. To this end, we combined parental report with direct language measures, as it was previously recommended and demonstrated that this combination may support the identification of LI in bilingual children (Paradis et al., 2013; Thordardottir, 2015). We extended the work of Paradis et al. (2010, 2013) and used a four-group design, including monolingual and bilingual children with and without LI, which allowed for a systematic examination of the effects of LI and bilingualism on the instruments and enabled us to investigate the diagnostic validity of the instruments within and across monolingual and bilingual contexts. Furthermore, unlike Paradis et al. (2013), the instruments included in the present study were specifically developed for use with bilingual children, aiming to reduce the bias against children with differing language experiences.

Previous work with the same four groups of children evaluated two direct language measures, namely the Q-U NWRT (Boerma et al., 2015) and the MAIN (Boerma et al., 2016), and demonstrated that these tasks did not disadvantage bilingual children. It is, however, yet unknown to what extent parental report on the early language development of a child and on the prevalence of language problems in the family is influenced by the lingual status of either the children or the parents. The present research is the first to systematically investigate the effects of LI and bilingualism on these indices based on a parental questionnaire, and, in addition, consider their diagnostic validity in both a monolingual and a bilingual learning context. These were the first two aims of the current study, which thus focused on parental report. We expected that the early language development of a child would not be influenced by bilingualism (De Houwer, 2009; Hoff et al., 2012) and would be a strong predictor of LI (Paradis et al., 2010). Family history was also hypothesized to discriminate between LI and TD (Restrepo, 1998), but may be a less reliable index of LI in bilingual children than in monolingual children due to interfering cultural factors (Paradis et al., 2010). For our final aim, following Paradis et al. (2013), we combined the parental questionnaire with the two direct language measures that have previously shown their potential for use with bilingual children (Boerma et al., 2015, 2016). We anticipated that the three instruments would adequately complement each other, resulting in a combination that could be highly successful in identifying LI in children, irrespective of children's linguistic background.

2. Methods

2.1. Participants

The present study used the same participant sample as Boerma et al. (2016), including four groups of 5 and 6 year old children ($N = 33$ per group). Half of the children in the sample had been previously diagnosed with LI according to a standardized protocol in the Netherlands (see Stichting Siméa, 2014), based on extensive clinical and psychometric evaluations by licensed professionals. These children either scored at least 2 standard deviations (SD) below the mean on a language assessment test battery or obtained a score of at least 1.5 SD below the mean on two out of four predefined subscales (speech production, auditory processing, grammatical knowledge and lexical-semantic knowledge) that were assessed with at least two appropriate measures. In the Netherlands, it varies from child to child which tasks are administered. The most commonly used instruments are the Dutch version of the *Clinical Evaluation of Language Fundamentals* (CELF-4-NL; Kort, Schittekatte, & Compaan, 2008), the *Schlichting Test for Language Production and Comprehension* (Schlichting & Lutje Spelberg, 2010a,b), the *Peabody Picture Vocabulary Test* (PPVT-III-NL; Schlichting, 2005), and the *Dutch Language Proficiency Test for All Children* which has bilingual norms (*Taaltoets Alle Kinderen* (TAK); Verhoeven & Vermeer, 2001). These instruments have adequate psychometric properties, as evaluated by the Committee On Test Affairs Netherlands (COTAN). In addition, a guideline on the assessment of bilingual children is provided by Stichting Siméa (2016), emphasizing the

Table 1
Characteristics of the participants (from Boerma et al., 2016).

	N	Age in months		Nonverbal intelligence		Gender	% Exposure to Dutch before age of 4		% Current exposure	
		Mean (SD)	Range	Mean (SD)	Range	Nr. of boys (%)	Mean (SD)	Range	Mean (SD)	Range
MOTD	33	71.3 (6.7)	59-84	102.9 (13.6)	81-128	21 (64%)	n/a	n/a	n/a	n/a
MOLI	33	71.8 (7.3)	59-87	97.0 (14.5)	72-127	22 (67%)	n/a	n/a	n/a	n/a
BITD	33 ^a	71.7 (7.1)	58-83	94.9 (14.1)	70-126	17 (52%)	42.1 (8.7)	25–60	52.3 (12.0)	29–79
BILI	33	72.8 (8.7)	58-86	94.5 (15.1)	71-124	23 (70%)	40.7 (11.0)	20–67	45.6 (17.1)	14–100 ^b

© 2016 Royal College of Speech and Language Therapists

Note: MOTD = monolingual typically developing; MOLI = monolingual language impaired; BITD = bilingual typically developing; BILI = bilingual language impaired.

^a Information on exposure to Dutch was missing for two bilingual TD children.

^b Due to severe difficulties learning the native tongue, parents of one child with LI decided to consistently speak Dutch to the child when he entered elementary school (explaining the 100% current exposure to Dutch at home). Before this, he was exposed to Dutch 50% of the time.

need for a bilingual anamnesis and the evaluation of the child's L1 and L2. The participants in the current study did not experience hearing loss or severe articulation problems and had a nonverbal intelligence score of 70 or above. The children with LI attended special education ($N = 64$) or received ambulatory care at a regular elementary school ($N = 2$). The second half of the sample included children with TD who attended regular education and did not have documented language difficulties.

Half of the children with TD and LI was monolingual (MOTD and MOLI, respectively), while the other half was raised bilingually (BITD and BILI, respectively). Following Kohnert's (2010: 456) definition of bilingualism, children were considered bilingual if they received 'regular input in two or more languages during the most dynamic period of communication development'. At least one parent of the child had to have a mother tongue other than Dutch, the majority language, and interact with the child in this language for a substantial period of the child's life. The children that met these criteria differed in the extent to which they had been exposed to Dutch, which is representative for the diverse groups of immigrants in the Netherlands (Centraal Bureau voor de Statistiek, 2015). The BITD and BILI groups were therefore matched on exposure to Dutch (Table 1) based on a parental questionnaire (PaBiQ; Tuller, 2015). The bilingual groups did not differ on exposure to Dutch before the age of 4 ($F(1,62) = 0.31, p = 0.58, \eta_p^2 = 0.01$) or current exposure to Dutch at home ($F(1,61) = 3.2, p = 0.08, \eta_p^2 = 0.05$), although there were indications that the bilingual TD children received slightly more Dutch input at home at the time of testing than the bilingual children with LI. The bilingual children were heterogeneous with respect to their first language background, but the majority of children were of Turkish (36%) or Moroccan (42%) descent (for more detailed information, see Boerma et al., 2016).

The four groups of children were matched on age in months, gender and nonverbal intelligence, measured with the short version of the *Wechsler Nonverbal-NL* (Wechsler & Naglieri, 2008). There were no significant differences between any of the four groups of children in age ($F(3,128) = 0.22, p = 0.88, \eta_p^2 = 0.01$), gender ($\chi^2(3, N = 132) = 2.7, p = 0.44$) or nonverbal intelligence ($F(3,128) = 2.4, p = 0.07, \eta_p^2 = 0.05$). Slightly higher scores of the MOTD group compared with the other groups resulted in the sub-optimal match on nonverbal intelligence. Table 1 presents the group characteristics. For background information on the Dutch language abilities of the children, we refer to Boerma et al. (2016).

2.2. Instruments and procedures

The Standing Ethical Assessment Committee of the Faculty of Social and Behavioral Sciences at Utrecht University approved this research. An informed consent was signed by parents of children who participated. A female native speaker of Dutch tested all participants individually at their school and administered tasks tapping into language, attention and working memory (not all relevant for the current study). Each participant completed two test sessions that lasted 45–60 minutes. Three instruments which were developed within the framework of the COST Action IS0804 *Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment* were used for the present study.

2.2.1. PaBiQ

The *Questionnaire for Parents of Bilingual Children* (PaBiQ; Tuller, 2015) is the short version of a longer questionnaire piloted by research groups in several countries within the COST Action IS0804, which is in part based on the *Alberta Language Environment Questionnaire* (ALEQ; Paradis, 2011) and the *Alberta Language and Development Questionnaire* (ALDeQ; Paradis et al., 2010). The questionnaire is intended to obtain information on possible developmental risk factors as well as quantity and quality of a child's language input. It hereby aims to enable professionals to better interpret children's performance on standardized tests, especially when assessment in both languages of a bilingual child is not possible. For the current study, the PaBiQ was administered during a telephone interview with one of the child's parents. The interview was conducted by bilingual speakers of both Dutch and Turkish/Berber/Moroccan Arabic and could therefore be carried out in the preferred language of the parent. For parents with a different native tongue, the interview was carried out in Dutch, if needed with the help of a family member or friend who could act as interpreter.

The literature reviewed in the introduction showed that a delayed early language development and the prevalence of language

problems in the family are strong risk factors of LI. These two factors were therefore included in the present study. Following [Tuller \(2015\)](#), the factor ‘Early Language Development’ consisted of: 1) Production of the first word (< 15 months = 6 points; 16–24 months = 4 points; > 24 months = 0 points), 2) Production of the first sentence (< 24 months = 6 points; 25–30 months = 4 points; > 31 months = 0 points), 3) Early parental concerns (no = 2 points; yes = 0 points). A high score on this variable, with a maximum of 14 points, indicated that a child experienced a positive early language development. For the bilinguals, the production of the first word and sentence was either in the L1 or the L2, whichever came first. The second factor, ‘Family History’, was derived from the question asking whether the child’s mother, father and siblings ever had experienced problems with reading/spelling, understanding others, and expressing themselves. With three family members and three types of language problems (3 × 3), a child was awarded a maximum of 9 points if language problems were not prevalent in the family. For each problem that each family member (had) experienced, one point was deducted from this maximum score.

2.2.2. Q-U NWRT

The *Quasi-Universal Nonword Repetition Task* (Q-U NWRT; [Chiat, 2015](#)), testing phonological short-term memory, comprises 16 items equally divided between items of two to five syllables. The items are compatible with cross-linguistically diverse constraints on lexical phonology, to maximize the task’s applicability across languages ([Chiat, 2015: 15](#)). For example, items do not include consonant clusters or final coda’s. Instead, they are composed of consonant-vowel sequences that occur frequently in many languages. The present study used the Dutch version of the task that was created by [Boerma et al. \(2015\)](#) and followed the same procedures. The authors of this study ensured that the items were not existing words in either the majority language Dutch or the most common minority languages (Turkish, Tarifit-Berber and Moroccan Arabic), covering 79% of the bilingual children’s languages. In addition, the items did not contain phonemes that did not originally occur in any of these languages. A female native speaker of Dutch recorded the 16 items. She produced the consonant-vowel sequences with their Dutch phonetic qualities, but avoided using language-specific Dutch prosody. With the aim of reducing a possible effect of language-specific prosodic knowledge, all syllables were stressed equally (see [Chiat, 2015](#)).

The Q-U NWRT was presented on a laptop screen. Following the procedure of [Engel de Abreu et al. \(2013\)](#), a cartoon ‘alien’ was introduced that wanted to teach a foreign language to the children. The task began with two practice items which were repeated until the child understood the procedure. Test items were only presented once. Children’s repetitions were scored offline by a native speaker of Dutch according to the procedure described by [Dollaghan and Campbell \(1998\)](#). Repetitions that included omissions were scored incorrect, in contrast to repetitions with only additions which do not indicate loss of information. Substitutions were also considered incorrect, unless they were systematic, reflecting articulation ability. The percentage of items correct was chosen as the dependent variable for the current study, as this scoring method is more practical for clinicians than the percentage of phonemes correct. A second independent rater scored over 50% of the data and scores overlapped in 98% of the cases. The intra-class correlation coefficient (ICC; absolute) was excellent (0.97) ([Boerma et al., 2015](#)).

2.2.3. MAIN

The *Multilingual Assessment Instrument for Narratives* (MAIN; [Gagarina et al., 2012](#)) was developed as a tool to evaluate narrative comprehension and production skills of bilingual children. The MAIN comprises four comparable stories that are illustrated by six pictures. Each story is composed of three episodes for which specific macrostructural story elements can be expressed. These elements include the introduction of a setting (place and time), and a basic episode structure consisting of a goal (e.g. Fox wanted to catch the baby goat), an attempt (e.g. Fox jumped up) and an outcome (e.g., Fox grabbed a baby goat). In addition, each episode allowed for the description of the internal states of the protagonist(s) (e.g., happy or scared).

The current study used the Dutch version of the MAIN and followed the same procedure as [Boerma et al. \(2016\)](#). Children were first presented with a model story (Dog or Cat) followed by ten questions that aimed to assess narrative comprehension skills. These comprehension questions tapped into the goals of the episodes ($N = 3$; ‘Why did the cat jump forwards?’), the internal states of the main characters ($N = 6$; ‘How does the cat feel?’), and the ability to make inferences about consequences of the events in the story ($N = 1$; ‘Will the boy become friends with the cat?’). Subsequently, children were asked to generate their own story based on a different picture sequence (Baby Birds or Baby Goats). The specific story elements described above (setting, goal, attempt, outcome and internal states) constituted the measure for narrative production. The groups of children were matched on which version of the model and production story they saw. The four groups did not differ with respect to version of the model story ($\chi^2(3, N = 132) = 1.4, p = 0.70$) or production story ($\chi^2(3, N = 132) = 0.12, p = 0.99$). Possible differences in degree of difficulty between the versions could thus not have influenced the outcomes. Children’s narratives were recorded and scored offline by a native speaker of Dutch. A second independent rater scored over 40% of the data. There was complete agreement between the two raters in 96% of the cases for narrative comprehension. The ICC was 0.93. For narrative production, the two raters agreed in 84% of the cases and the ICC was good (0.82) ([Boerma et al., 2016](#)).

2.3. Data-analysis

SPSS 22 (IBM Corp., 2013) was used to conduct all statistical analyses. The indices Early Language Development and Family History violated the assumption of normality. Nonparametric Mann-Whitney U tests were therefore conducted to examine the effects of LI and bilingualism, which was the first aim of the study. Secondly, we investigated to what extent the indices of the PaBiQ could predict the absence or presence of LI, and, finally, whether the predictive accuracy improved when the Q-U NWRT and MAIN were added to the model. To this end, a hierarchical binary logistic regression was carried out with Early Language Development and

Family History as fixed factors in the first block, scores of the Q-U NWRT in the second block, and the comprehension and production measures of the MAIN in the third block. This order was based on previous work indicating that the Q-U NWRT attained higher classification accuracy than the MAIN (Boerma et al., 2015, 2016). Starting with the full model with all predictors, backward elimination was used to determine the optimal model that maximized the classification accuracy but minimized the number of predictors. The logistic regression was first carried out in the monolingual and bilingual group separately, after which the same analysis including both groups was conducted. The predicted group membership was saved in both cases, allowing us to compare children's classification in the separate analyses with their classification in the combined analysis. If scores on the instruments are, as we predicted, not influenced by differing language experiences, individual classification of children as TD and LI is not expected to be altered by the composition of the reference group. On the other hand, differences in the classification between the separate and combined analyses are expected to emerge when the instruments are biased against bilingual children. In that case, the inclusion of a monolingual reference group could require a higher cut-off criteria and result in the overidentification of bilingual children with TD as having LI.

In the present study, sensitivity of the instruments is calculated based on the proportion of children who were correctly identified as having LI, while specificity is determined by the proportion of children correctly identified as having TD. Following Plante and Vance (1994), we assumed that sensitivity and specificity above 90% were good and rates between 80% and 89% were fair. The diagnostic utility of the instruments is furthermore indicated by likelihood ratios, which were derived from the sensitivity and specificity levels. A positive likelihood ratio (LR+) demonstrates how likely it is that a child with LI scores within the impaired range on the instruments, and a negative likelihood ratio (LR-) shows how likely it is that a child with TD scores within the normal range. Following Dollaghan (2004), likelihood ratios of 1 were considered clinically uninformative, while a LR+ of 10 or more and a LR- of 0.10 or less was assumed to indicate a large and convincing increase or decrease in the likelihood of LI, respectively.

3. Results

3.1. Effects of LI and bilingualism on two indices of the PaBiQ

Table 2 presents the scores of the four groups of children on Early Language Development and Family History based on the PaBiQ. A nonparametric Mann-Whitney U test revealed a negative effect of LI on Early Language Development in the monolingual ($U = 65.00, z = 6.05, p < 0.001, r = 0.77$) and bilingual ($U = 137.50, z = 4.94, p < 0.001, r = 0.63$) groups of children. Parents of children with LI more often indicated that their child had a late first word onset (monolingual: $U = 298.50, z = 3.21, p < 0.01, r = 0.40$; bilingual: $U = 218.50, z = 4.01, p < 0.001, r = 0.51$), and late first sentence onset (monolingual: $U = 118.00, z = 5.71, p < 0.001, r = 0.73$; bilingual: $U = 165.00, z = 4.86, p < 0.001, r = 0.62$) than parents of children with TD. In addition, parents of children with LI were more often concerned about the language development of their child (monolingual: $\chi^2(1, N = 66) = 44.3, p < 0.001, \phi = 0.82$; bilingual: $\chi^2(1, N = 63) = 19.4, p < 0.001, \phi = 0.55$). A negative effect of LI also emerged on Family History in the group of monolingual children ($U = 336.00, z = 2.93, p < 0.01, r = 0.36$). Family members of monolingual children with LI were more likely to have problems with reading, understanding others or expressing themselves than family members of monolingual children with TD. No effect of LI on Family History was found in the bilingual group.

Another Mann-Whitney U test with Language Group as grouping variable revealed a marginally significant negative effect of bilingualism on Early Language Development in the TD group ($U = 334.00, z = 2.00, p = 0.05, r = 0.26$). Bilingual TD children did not differ from monolingual TD children in reaching the basic early milestones (first word and sentence). However, a marginally significant difference emerged on early parental concerns ($\chi^2(1, N = 64) = 3.6, p = 0.06, \phi = 0.24$), suggesting that parents of bilingual TD children were more often concerned about the language development of their child than parents of monolingual TD children. There were no differences between the monolingual and bilingual children with LI on any of the components of the index of Early Language Development, nor on the composite score. On Family History, no effect of bilingualism emerged in the TD group. On the other hand, parents of monolingual children with LI more often reported the occurrence of language problems in the family than parents of bilingual children with LI ($U = 366.00, z = 2.30, p = 0.02, r = 0.29$).

Table 2

Scores of the four groups of children on Early Language Development and Family History based on the PaBiQ.

	Early Language Development (max. 14)						Family History (max. 9)			
			First Word (max .6)		First Sentence (max. 6)		Early Concerns (max. 2)			
	N^a	Mean (SD)	N^a	Mean (SD)	N^a	Mean (SD)	N^a	Mean (SD)	N^a	Mean (SD)
MOTD	30	13.07 (2.39)	31	5.55 (1.23)	30	5.67 (1.18)	33	1.88 (0.49)	33	8.48 (1.06)
MOLI	32	5.69 (4.34)	32	3.81 (2.56)	32	1.63 (2.51)	33	0.24 (0.67)	33	7.94 (1.00)
BITD	30	12.07 (3.04)	30	5.33 (1.32)	30	5.20 (1.86)	31	1.55 (0.85)	31	8.16 (1.19)
BILI	32	5.13 (5.25)	32	2.81 (2.63)	32	1.88 (2.54)	32	0.44 (0.84)	32	8.41 (0.98)

Note: The scoring scheme of the PaBiQ (Tuller, 2015) is followed to obtain the scores presented in this table (see Methods section for further details). MOTD = monolingual typically developing; MOLI = monolingual language impaired; BITD = bilingual typically developing; BILI = bilingual language impaired.

^a Information on Early Language Development and Family History was missing for two bilingual TD children and one bilingual child with LI. Moreover, some parents did not recall when their child produced his/her first word and/or sentence.

Table 3

Hierarchical binary logistic regression in the monolingual group with outcome measures of the PaBiQ, Q-U NWRT and MAIN as predictors of Impairment Status.

Block	Model ^a	Predictors included	B	SE	Wald	p	Sn	Sp	LR+	LR-
1	Full	Early Development	-.518	.140	13.728	.000	81%	90%	7.85	.21
		Family History	-.381	.440	.746	.388				
	Optimal	Early Development	-.545	.141	14.962	.000	81%	90%	7.85	.21
2	Full & Optimal	Early Development	-.340	.125	7.412	.006	94%	93%	13.6	.07
		Q-U NWRT	-.113	.039	8.585	.030				
3	Full	Early Development	-.300	.129	5.416	.020	94%	93%	13.6	.07
		Q-U NWRT	-.076	.039	3.775	.052				
		MAIN Production	-.113	.080	2.005	.157				
		MAIN Comprehension	-.046	.045	1.032	.310				
	Optimal	Early Development	-.296	.126	5.563	.018	94%	93%	13.6	.07
		Q-U NWRT	-.085	.038	5.089	.024				
MAIN Production		-.108	.072	2.230	.135					

Note: Sn = Sensitivity; Sp = Specificity; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio.

^a The full model is the model with all predictors and the optimal model is the minimal model that generated the highest sensitivity and specificity.

3.2. Diagnostic validity

A hierarchical binary logistic regression was carried out to examine the diagnostic validity of the PaBiQ indices separately as well as in combination with the Q-U NWRT and the MAIN. Early Language Development and Family History were used as predictors of Impairment Status in the first block of the regression. Scores on the Q-U NWRT and MAIN (Production and Comprehension) were added in the second and third block, respectively (see Appendix A for their descriptive statistics). Results for the monolingual and bilingual group are presented in Tables 3 and 4, respectively.

In the monolingual group, the first block of the binary logistic regression showed that the model including Early Language Development and Family History as predictors was statistically significant from the intercept only model ($\chi^2(2, N = 61) = 42.32, p < 0.001$, Nagelkerke $R^2 = 0.67$). Over 85% of the monolingual children were correctly identified as having TD or LI. Sensitivity was 81% and specificity was 90%, indicating fair to good discriminant accuracy (Plante & Vance, 1994). Model reduction resulted in the exclusion of Family History. This optimal model remained significant ($\chi^2(1, N = 61) = 41.52, p < 0.001$, Nagelkerke $R^2 = 0.66$) and did not show poorer fit compared to the full model ($\Delta\chi^2(1, N = 61) = -0.80, p = 0.37$). The addition of the Q-U NWRT as predictor in block 2 resulted in a significant model ($\chi^2(2, N = 61) = 58.64, p < 0.001$, Nagelkerke $R^2 = 0.82$) with improved model fit ($\Delta\chi^2(1, N = 61) = 17.12, p < 0.001$). Sensitivity and specificity increased to 94% and 93%, respectively. The

Table 4

Hierarchical binary logistic regression in the bilingual group with outcome measures of the PaBiQ, Q-U NWRT and MAIN as predictors of Impairment Status.

Block	Model ^a	Predictors included	B	SE	Wald	p	Sn	Sp	LR+	LR-
1	Full	Early Development	-.324	.081	15.860	.000	70%	90%	6.77	.33
		Family History	.060	.292	.042	.838				
	Optimal	Early Development	-.326	.081	16.053	.000	70%	90%	6.77	.33
2	Full & Optimal	Early Development	-.360	.112	10.342	.001	90%	87%	6.95	.12
		Q-U NWRT	-.108	.033	10.587	.010				
3	Full & Optimal	Early Development	-.462	.150	9.493	.002	97%	97%	28.0	.03
		Q-U NWRT	-.071	.034	4.312	.038				
		MAIN Comprehension	-.065	.039	2.794	.095				
		MAIN Production	-.072	.048	2.259	.133				

Note: Sn = Sensitivity; Sp = Specificity; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio.

^a The full model is the model with all predictors and the optimal model is the minimal model that generated the highest sensitivity and specificity.

classification accuracy did not further increase when MAIN Production and Comprehension were added in block 3, nor did the model fit improve ($\Delta\chi^2(2, N = 61) = 4.22, p = 0.12$). When the model in block 3 was further reduced, excluding MAIN Comprehension, a marginally significant improvement in model fit emerged ($\Delta\chi^2(1, N = 61) = 3.16, p = 0.08$), resulting in a final model including Early Language Development, the Q-U NWRT and MAIN Production ($\chi^2(3, N = 61) = 61.80, p < 0.001$, Nagelkerke $R^2 = 0.85$) (Table 3).

In the bilingual group, the model with the PaBiQ indices of Early Language Development and Family History was also statistically significant compared to the intercept only model ($\chi^2(2, N = 60) = 28.03, p < 0.001$, Nagelkerke $R^2 = 0.50$). Sensitivity was 70% and specificity was 90%, indicating that a fewer number of bilingual children with LI were correctly classified by the PaBiQ indices compared with monolingual children with LI. The discriminant accuracy remained the same when the predictor Family History was excluded. This optimal model did not decrease in fit from the full model ($\Delta\chi^2(1, N = 60) = -0.04, p = 0.84$) and was again significant ($\chi^2(1, N = 60) = 27.99, p < 0.001$, Nagelkerke $R^2 = 0.50$). Model fit improved in block 2 when the Q-U NWRT was added ($\Delta\chi^2(1, N = 60) = 19.04, p < 0.001$), resulting in a significant model with two predictors ($\chi^2(2, N = 60) = 47.03, p < 0.001$, Nagelkerke $R^2 = 0.72$). Sensitivity (90%) increased and specificity slightly decreased (86%). The model was further strengthened by the addition of the outcome measures of the MAIN in block 3 ($\Delta\chi^2(2, N = 60) = 7.78, p = 0.02$). Together, Early Language Development, Q-U NWRT, MAIN Production and MAIN Comprehension classified 97% of the bilingual children with TD and LI correctly. The final model with four predictors was significant ($\chi^2(4, N = 60) = 54.81, p < 0.001$, Nagelkerke $R^2 = 0.80$) (Table 4).

Another binary logistic regression was conducted with the data of all children together to check whether the classification of children with TD and LI was affected by the inclusion of both monolingual and bilingual children. Results showed a sensitivity and a specificity of 95%, misclassifying three children with TD and three children with LI. The predicted group membership was saved and was compared with the predicted group membership of the analyses in the monolingual and bilingual group separately. Exactly the same children were misclassified in the combined analysis as in the separate analyses, indicating that the individual classification was not affected by using the instruments with a mixed group including both monolingual and bilingual children.

4. Discussion

It is highly recommended that bilingual children are assessed in both languages when their language development causes concerns (ASHA, 2004; IALP, 2011). This is, however, not always feasible due to the multitude of language combinations that are encountered in clinical practice. Adapted bilingual norm-referencing only partly solves the problem, since it is challenging to determine the appropriate standard for every child, especially for sequential bilingual children (Thordardottir, 2015). The present study investigated the clinical value of an alternative approach to assessing both languages, combining data from a parental questionnaire with two direct language measures that allow children to use acquired skills in any language, instead of tapping into language-specific abilities. This approach was aimed to reduce the bias against bilingual children, while remaining informative about the presence or absence of LI.

We first examined the effects of LI and bilingualism on two well-known developmental risk factors of LI, about which information was obtained through parental report via a questionnaire. Risks associated with a child's early language development were indexed by the timing of basic milestones and parental concerns, and risks associated with family history were indexed by the prevalence of language problems in a child's close family members. The results showed large negative effects of LI on the index of Early Language Development, both in the monolingual and bilingual group, confirming the status of late language emergence as risk factor of LI (Tuller, 2015). In addition, a marginally significant negative effect of bilingualism emerged on Early Language Development in the typically developing (TD) group of children. Monolingual and bilingual children did not differ in terms of the timing of their basic early milestones, which is in line with previous work (Hoff et al., 2012; Paradis et al., 2010). However, there were indications that parents of bilingual TD children had been more often concerned about the language development of their child than parents of monolingual TD children. Together with the smaller effect size of LI in the bilingual group ($\phi = 0.55$) compared with the monolingual group ($\phi = 0.82$), this suggests that parental concerns may thus be a stronger indicator of LI in monolingual children than in bilingual children.

With respect to the index of Family History, results showed an expected negative effect of LI in the group of monolingual children, whereas this effect was not observed in the bilingual group. Although Paradis et al. (2010) did report a significant difference between their bilingual groups with TD and LI, they mention that the willingness and ability to report family history may be complicating cultural factors when investigating the prevalence of language problems in culturally diverse settings (p. 486). The observation that parents of bilingual children with LI were less likely to report the occurrence of language problems in the family than parents of monolingual children with LI reinforces the notion that such cultural factors may have played a role in the questionnaire data from the current study. Family History as reported by parents may thus be less reliable as an index of LI in bilingual children than in monolingual children.

In a next step, we examined to what extent the indices of the parental questionnaire could predict the absence or presence of LI. Whereas Family History did not significantly predict LI in either the monolingual or bilingual group, the index of Early Language Development was a strong predictor in both groups of children, classifying many of the children with LI and TD correctly. This corresponds to the findings of Paradis et al. (2010), who initially questioned the reliability of retrospective parental report, but concluded that salient elements of a child's early language development, such as basic milestones, may not be difficult to recall for parents. However, the question remains to what extent retrospective reports of parents of children with LI are influenced by the diagnosis that was given to these children. Parents may have unconsciously underestimated the timing of their child's first word or

word combination, or overestimated their own concerns. As Leonard (2014: 158) pointed out, it may not necessarily be the case that all children with LI start out with a small and late-developing lexicon. Thus, even though late language emergence appears a strong risk factor of LI irrespective of a child's linguistic background, this emphasizes the importance of using information about the early language development of a child in combination with other instruments.

We therefore combined the index of Early Language Development with two instruments that have proven to be sensitive to effects of LI, but insensitive to effects of bilingualism (Boerma et al., 2015, 2016). The addition of the quasi-universal nonword repetition task (Chiat, 2015) significantly improved the predictive accuracy in both the monolingual and bilingual group of children. Furthermore, the narrative task (MAIN; Gagarina et al., 2012) was of significant added value to the parental questionnaire and quasi-universal nonword repetition task in the bilingual group, resulting in correct classification of 97% of the bilingual children with LI and TD. In the monolingual group, classification accuracy did not change when the narrative task was included, although narrative production scores marginally improved the model fit. Sensitivity and specificity remained 94% and 93%, respectively. Narrative comprehension was the only component that produced a meaningful difference between the analyses in the monolingual and bilingual group of children. The classification accuracy in the bilingual group benefitted from the weak performance of the bilingual children with LI on the comprehension questions of the MAIN. However, narrative comprehension did not emerge as a significant predictor of LI in the monolingual group of children, as the difference in the comprehension scores of the children with and without LI in the monolingual group was relatively small compared with the difference in the bilingual group. Apart from narrative comprehension, results from the separate analyses in the monolingual and bilingual group were highly similar. A combined analysis that included all participants, making no distinction between monolinguals and bilinguals, even classified the exact same children as having LI and TD as separate analyses of the two groups, showing that a change in reference group did not affect the individual classification.

The present study demonstrated that a combination of instruments consisting of a parental questionnaire, a quasi-universal nonword repetition task, and a narrative task testing macrostructure can reliably identify LI in monolingual and bilingual learning contexts. The tools are quick and easy to administer, and provide valuable information about the language development of a child. Moreover, the classification of children as TD or LI was not affected by the composition of the reference group. This suggests that bilingual norm-referencing may not be necessary when using this combination of instruments. However, subsequent steps have to be taken before these instruments can be implemented in clinical practice. The sample in the current study was limited in size and, moreover, may have overrepresented children with LI in special education schools. Based on their broad educational needs (e.g., in the areas of language, reading, math, socio-emotional development, and learning abilities), children with LI in the Netherlands either attend special education schools or regular education with ambulatory care (Stichting Siméa, 2013). Including mostly children in special education, possibly with relatively great educational needs, may thus have limited the generalizability of our results. Future work within a large population-based sample is needed to validate our findings, and, in addition, would allow for the standardization of the instruments and the development of clinically relevant index scores. The present research made important headway for such future work, evaluating an unbiased method of language assessment for bilingual children and offering a promising alternative to testing both languages or using bilingual norm-referencing when such methods are not within reach.

Acknowledgements

This work is part of the research program 'Cognitive development in the context of emerging bilingualism: Cultural minority children in the Netherlands' which is financed by a VIDI-grant awarded to dr. Elma Blom (PI) by the Netherlands Organization for Scientific Research (NWO; grant number 016.124.369). We thank the children, parents and schools that participated in the study. We give special thanks to the members of the COST Action IS0804 who developed the instruments used in the current study.

Appendix A. Percentage of items correct on the Quasi-Universal Nonword Repetition Task (Q-U NWRT – Boerma et al., 2015) and narrative performance on the MAIN (Boerma et al., 2016).

Percentage of items correct on the Quasi-Universal Nonword Repetition Task (Q-U NWRT – Boerma et al., 2015) and narrative performance on the MAIN (Boerma et al., 2016).

	N	Q-U NWRT %		MAIN	
		Mean (SD)	N	Comprehension % Mean (SD)	Production % Mean (SD)
MOTD	32 ^a	58.95 (15.0)	33	91.14 (9.3)	45.27 (11.9)
MOLI	33	27.41 (13.9)	33	81.2 (15.6)	29.77 (7.9)
BITD	33	55.15 (14.6)	33	90.24 (9.3)	44.03 (11.3)
BILI	31 ^a	28.10 (17.1)	33	71.82 (23.2)	29.77 (14.0)

Note: MOTD = monolingual typically developing; MOLI = monolingual language impaired; BITD = bilingual typically developing; BILI = bilingual language impaired.

^aThe Quasi-Universal NWRT of three children could not be scored properly due to unintelligibility.

Appendix B. Continuing education questions

- 1 The index of Early Language Development is a significant predictor of LI in:
 - A Monolingual children.
 - B Bilingual children.
 - C Monolingual and bilingual children.
 - D The index of Early Language development is not a significant predictor of LI.

- 2 The prevalence of language problems in the family as reported by parents may be a less reliable indicator of LI in bilingual children than in monolingual children. **TRUE/FALSE?**

- 3 Nonword repetition tasks never disadvantage bilingual children. **TRUE/FALSE?**

- 4 In the current study:
 - A Narrative comprehension was a stronger predictor of LI in monolingual children than in bilingual children.
 - B Narrative comprehension did not improve model fit in either the group of monolingual or the group of bilingual children.
 - C Narrative comprehension and production were excluded from the model in the group of monolingual children.
 - D Narrative comprehension and production improved the classification accuracy in the group of bilingual children.

- 5 When a bilingual child's language development causes concerns, a promising alternative to assessing both languages or using bilingual norm-referencing may be:
 - A Testing one language with standardized measures tapping into vocabulary and grammar.
 - B Analyzing the spontaneous speech of the child in one language.
 - C Combining parental report with unbiased language measures.
 - D Waiting until the child gets older, having had more exposure to the target language.

References

- American Speech-Language-Hearing Association (2004). *Knowledge and skills needed by speech-language pathologists and audiologists to provide culturally and linguistically appropriate services [Knowledge and Skills]*. Retrieved from: www.asha.org/policy.
- Armon-Lotem, S., & Meir, N. (2016). Sensitivity and specificity of repetition tasks for diagnosing SLI in monolingual and bilingual children. *International Journal of Language and Communication Disorders*. <http://dx.doi.org/10.1111/1460-6984.12242>.
- Bedore, L., & Peña, E. (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.
- Berman, R., & Slobin, D. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bishop, D., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology*, 23(1), 25–46.
- Bishop, D., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language & Communication Disorders*, 44(5), 600–615.
- Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research*, 58(6), 1747–1760.
- Boerma, T., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2016). Narrative abilities of monolingual and bilingual children with and without language impairment: Implications for clinical practice. *International Journal of Language and Communication Disorders*, 51(6), 626–638.
- Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research*, 40(3), 519–525.
- Centraal Bureau voor de Statistiek [CBS; Statistics Netherlands] (2015). *Bevolking: Generatie, geslacht, leeftijd en herkomstgroepering, 1 januari*. Den Haag. Retrieved from: www.cbs.nl.
- Chiat, S. (2015). Nonword repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 125–150). Bristol: Multilingual Matters.
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, 42(6), 741–748.
- De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.
- Dodwell, K., & Bavin, E. (2008). Children with specific language impairment: An investigation of their narratives and memory. *International Journal of Language and Communication Disorders*, 43(2), 201–218.
- Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, 41(5), 1136–1146.
- Dollaghan, C. (2004). Evidence-based practice in communication disorders: What do we know, and when do we know it? *Journal of Communication Disorders*, 37(5), 391–400.
- Ellis Weismer, S., Tomblin, J., Zhang, X., Buchwalter, P., Chynoweth, J., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language and Hearing Research*, 43(4), 865–878.
- Engel de Abreu, P., Baldassi, M., Puglisi, M., & Bepi-Lopes, D. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language minority children with an immigrant background. *Journal of Speech, Language and Hearing Research*, 56(2), 630–642.
- Engel de Abreu, P. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*, 19(5), 529–537.
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., et al. (2012). MAIN – multilingual assessment instrument for narratives. *ZAS Papers in Linguistics*, 56, 1–155.
- Gathercole, V., Thomas, E., Roberts, E., Hughes, C., & Hughes, E. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 20–55). Bristol: Multilingual Matters.

- Girbau, D., & Schwartz, R. (2008). Phonological working memory in Spanish–English bilingual children with and without specific language impairment. *Journal of Communication Disorders*, 41(2), 124–145.
- Graf-Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in the nonword repetition performance of children with and without Specific Language Impairment: A meta-analysis. *Journal of Speech, Language and Hearing Research*, 50(1), 177–195.
- Grimm, A., & Schulz, P. (2014). Specific language impairment and early second language acquisition: The risk of over- and underdiagnosis. *Child Indicators Research*, 7(4), 821–841.
- Gutiérrez-Clellen, V., & Simon-Cerejido, G. (2010). Using nonword repetition tasks for the identification of language impairment in Spanish-English-speaking children: Does the language of assessment matter? *Learning Disabilities Research & Practice*, 25(1), 48–58.
- Håkansson, G., Salameh, E., & Nettelbladt, U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics*, 41(2), 255–288.
- Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015). Narrative abilities in subgroups of English language learners and monolingual peers. *International Journal of Bilingualism*, 19(6), 677–692.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(1), 1–27.
- International Association of Logopedics and Phoniatrics (IALP) (2011). *Recommendations for working with bilingual children*. Retrieved from: http://www.linguistics.uaberta.ca/en/CHESL_Centre/?/media/linguistics/Media/CHESL/Documents/WorkingWithBilingualChildren-May2011.pdf.
- Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with primary language impairment from typical bilinguals? *Journal of Learning Disabilities Research & Practice*, 21(1), 19–29.
- Kohnert, K. (2010). Bilingual children with primary language impairment: Issues evidence and implication for clinical actions. *Journal of Communication Disorders*, 43(6), 456–473.
- Kort, W., Schittekatte, M., & Compaan, E. (2008). *CELF-4-NL: Clinical Evaluation of Language Fundamentals*. Amsterdam: Pearson Assessment and Information B.V.
- Leonard, L. (2014). *Children with specific language impairment* (2nd ed.). Cambridge: MIT Press.
- Pankratz, M., Plante, E., Vance, R., & Insalaco, D. (2007). The diagnostic and predictive validity of The Renfrew Bus Story. *Language, Speech, and Hearing Services in Schools*, 38(4), 390–399.
- Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, 43(6), 474–497.
- Paradis, J., Schneider, P., & Sorenson Duncan, T. (2013). Discriminating children with language impairment among English language learners from diverse first language backgrounds. *Journal of Speech, Language and Hearing Research*, 56(3), 971–981.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213–237.
- Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research*, 59(1), 171–182.
- Peña, E., Gutiérrez-Clellen, V., Iglesias, A., Goldstein, B., & Bedore, L. (2014). *Bilingual English Spanish Assessment (BESA)*. San Rafael, CA: AR Clinical Publication.
- Pearson, B. (2002). Narrative competence among monolingual and bilingual school children in Miami. In D. K. Oller, & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 135–174). Clevedon, UK: Multilingual Matters.
- 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.
- Restrepo, A. (1998). Identifiers of predominately Spanish-speaking children with language impairment. *Journal of Speech, Language and Hearing Research*, 41(6), 1398–1411.
- Rice, M. (2007). Children with specific language impairment: Bridging the genetic and developmental perspectives. In E. Hoff, & M. Shatz (Eds.), *Handbook of language development* (pp. 411–431). Oxford, UK: Blackwell Publishers.
- Salameh, E., Nettelbladt, U., Håkansson, G., & Gullberg, B. (2002). Language impairment in Swedish bilingual children. A comparison between bilingual and monolingual children. *Acta Paediatrica*, 91(2), 22–234.
- Schlichting, J., & Lutje Spelberg, H. (2010a). *Schlichting Test voor Ttest voor taalbegrip: Manual*. Houten: Bohn, Stafleu, Van Loghum.
- Schlichting, J., & Lutje Spelberg, H. (2010b). *Schlichting Test voor Ttest voor taalproductie-II: Manual*. Houten: Bohn, Stafleu, Van Loghum.
- Schlichting, L. (2005). *Peabody Picture Vocabulary Tpicture vocabulary test-III-NL. Dutch version*. Amsterdam: Harcourt Assessment B.V.
- Schneider, P., Hayward, D., & Dubé, R. (2006). Storytelling from pictures using the edmonton narrative norms instrument. *Journal of Speech-Language Pathology and Audiology*, 30(4), 224–238.
- Sharp, K., & Gathercole, V. (2013). Can a novel word repetition task be a language neutral assessment tool? Evidence from Welsh–English bilingual children. *Child Language Teaching and Therapy*, 29(1), 77–89.
- Smeets, E., Driessen, G., Elfering, S., & Hovius, M. (2009). *Allochtone leerlingen en speciale onderwijsvoorzieningen*. Nijmegen: ITS.
- Squires, K., Lugo-Neris, M., Peña, E., Bedore, L., Bohman, T., & Gillam, R. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language and Communication Disorders*, 49(1), 60–74.
- Stichting Siméa (2013). *Handreiking toeleiding naar onderwijsarrangementen*. Retrieved from: <http://www.simea.nl/dossiers/passend-onderwijs/onderwijsarrangementen/handreiking-toeleiding-onderwijsarrangementen-mei-2013.pdf>.
- Stichting Siméa (2014). *Indicatiecriteria: Auditief en/of communicatief beperkte leerlingen*. Retrieved from: <http://www.simea.nl/dossiers/passend-onderwijs/brochures-po/14-simea-brochure-indicatiecriteria-digitaal.pdf>.
- Stichting Siméa (2016). *Handreiking meertaligheid TOS*. Retrieved from: <http://www.simea.nl/dossiers/si166289-simea-handreiking-meertaligheid-tos.pdf>.
- Summers, C., Bohman, T., Gillam, R., Peña, E., & Bedore, L. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders*, 45(4), 480–493.
- Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*, 46(1), 1–16.
- Thordardottir, E., & Juliusdottir, A. (2013). Icelandic as a second language: A longitudinal study of language knowledge and processing by school-age children. *International Journal of Bilingual Education and Bilingualism*, 16(4), 411–435.
- Thordardottir, E., Rothenberg, A., Rivard, M., & 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.
- Thordardottir, E. (2015). Proposed diagnostic procedures for use in bilingual and cross-linguistic contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 331–358). Bristol: Multilingual Matters.
- Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 301–330). Bristol: Multilingual Matters.
- Verhoeven, L., & Vermeer, A. (2001). *Taaltoets Alle Kinderen (TAK)*. Arnhem: Cito.
- Wechsler, D., & Naglieri, J. (2008). *Wechsler Nonverbal Scale of Ability*. Purchased from: <http://www.pearsonclinical.nl/wvn-nl-wechsler-non-verbal>.
- Windsor, J., Kohnert, K., Lobitz, K., & Pham, G. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology*, 19(4), 298–310.