Journal of Communication Disorders 46 (2013) 1–16

Contents lists available at SciVerse ScienceDirect

Journal of Communication Disorders

# The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores

Elin Thordardottir [a,b,*], Myrto Brandeker [a,b]

[a] McGill University, Canada
[b] Centre interdisciplinaire de recherche en réadaptation du Montréal métropolitain (CRIR), Canada

ARTICLE INFO

*Article history:*
Received 17 November 2011
Received in revised form 26 June 2012
Accepted 16 August 2012
Available online 6 September 2012

*Keywords:*
Nonword repetition
Sentence imitation
Bilingualism
Primary language impairment
Clinical marker
Diagnostic accuracy

ABSTRACT

*Purpose:* Nonword repetition (NWR) and sentence imitation (SI) are increasingly used as diagnostic tools for the identification of Primary Language Impairment (PLI). They may be particularly promising diagnostic tools for bilingual children if performance on them is not highly affected by bilingual exposure. Two studies were conducted which examined (1) the effect of amount of bilingual exposure on performance on French and English nonword repetition and sentence imitation in 5-year-old French-English bilingual children and (2) the diagnostic accuracy of the French versions of these measures and of receptive vocabulary in 5-year-old monolingual French-speakers and bilingual speakers with and without PLI, carefully matched on language exposure.
*Method:* Study 1 included 84 5-year-olds acquiring French and English simultaneously, differing in their amount of exposure to the two languages but equated on age, nonverbal cognition and socio-economic status. Children were administered French and English tests of NWR and SI. In Study 2, monolingual and bilingual children with and without PLI (four groups, $n = 14$ per group) were assessed for NWR, SI, and receptive vocabulary in French to determine diagnostic accuracy.
*Results:* Study 1: Both processing measures, but in particular NWR, were less affected by previous exposure than vocabulary measures. Bilingual children with varying levels of exposure were unaffected by the length of nonwords. Study 2: In contrast to receptive vocabulary, NWR and SI correctly distinguished children with PLI from children with typical development (TD) regardless of bilingualism. Sensitivity levels were acceptable, but specificity was lower.
*Conclusions:* Bilingual children perform differently than children with PLI on NWR and SI. In contrast to children with PLI, bilingual children with a large range of previous exposure levels achieve high NWR scores and are unaffected by the length of the nonwords.

**Learning outcomes:** Readers will recognize the effect of language input on the rate of language development, focusing specifically on how bilingual exposure affects the language learning of each language of bilingual children.



* Corresponding author at: School of Communication Sciences and Disorders, McGill University 1266, Pine Avenue West, Montreal, Quebec H3G 1A8, Canada. Tel.: +1 514 398 5356; fax: +1 514 398 8123.
*E-mail address:* elin.thordardottir@mcgill.ca (Elin Thordardottir).

0021-9924/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcomdis.2012.08.002

For over a decade, nonword repetition (NWR) and sentence imitation (SI) tasks have been increasingly frequent as a standard part of protocols for the assessment of children's language abilities. In 1998, Dollaghan and Campbell showed that an English NWR repetition test was highly sensitive to the presence of language impairment, but relatively insensitive to dialectal differences and differences in socio-economic status (SES), making it a promising tool across a wide population. Further studies have confirmed the diagnostic utility of several other NWR tasks (Archibald & Gathercole, 2006; Conti-Ramsden, 2003; Ellis Weismer et al., 2000). Whereas SI measures have long been part of many standardized tests, a study by Conti-Ramsden, Botting, and Faragher (2001) demonstrating its high degree of diagnostic accuracy heightened interest in the clinical interpretation and utility of this task. Both of these tasks are complex in terms of the abilities needed to complete them. However, a main interest in these tasks is that they focus relatively more on linguistic processing than on accumulated linguistic knowledge, in that nonwords involve non-meaningful stimuli and both tasks involve simple repetition and no linguistic formulation or demonstration of comprehension. These properties are likely to underlie the lack of sensitivity of NWR to dialectal differences and they may similarly make these tasks particularly useful for the assessment of bilingual children for the purpose of identifying Primary Language Impairment (PLI – the term PLI is preferred here over the term Specific Language Impairment (SLI), but it refers to the same group of children –those with an impairment *primarily* in the area of language, but who may have subtle deficits in other developmental domains not sufficiently severe to merit their own clinical diagnosis).

The identification of PLI in bilingual children is made remarkably difficult by the fact that bilingual exposure in general causes children to score lower than monolingual children on standardized tests of accumulated language knowledge – similarly to children with language impairment, albeit for a different reason. The extent of the difference is hard to predict and depends on a complex set of factors such as the quality, quantity and timing of previous exposure to each language (see e.g. Elin Thordardottir, Rothenberg, Rivard, & Naves, 2006). Measures of language processing have been seen as potentially offering a way around this dilemma. However, the extent to which performance on these measures depends on language exposure and linguistic knowledge is unclear. Presently, the effect of language exposure on NWR and SI performance has been largely inferred indirectly from studies focusing primarily on diagnostic accuracy in various populations of bilingual children. More direct examination of the effect of language exposure is required to better understand the extent to which it affects NWR and SI performance. This article presents two studies which aim to clarify this issue. The first examines the effect of bilingual exposure on the NWR and SI performance of bilingual children with typical development (TD) having been exposed to French and English for varying percentages of their total waking hours from birth. Having established this for the measures used in the first study, the second study assesses the diagnostic accuracy of the French NWR and SI measures in comparison to that of a standardized test of receptive vocabulary. This study provides novel insights in that it includes groups of monolingual and bilingual children with and without PLI and in that the bilingual groups have carefully documented exposure histories and have matched closely on this variable.

Nonword repetition has been shown to produce significant differences in performance between groups of monolingual children with PLI and typical development (TD) in a number of languages including English (Conti-Ramsden, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000), Spanish (Girbau & Schwartz, 2007, 2008; Gutiérrez-Clellen & Simon-Cereijido, 2010), Dutch (de Bree, Rispens, & Gerrits, 2007), Swedish (Hansson, Forsberg, Löfqvist, Mäki-Torkko, & Sahlén, 2004), Icelandic (Elin Thordardottir, 2008), and French (Elin Thordardottir et al., 2011). The only language published on to date in which significant group differences were not found is Cantonese (Stokes, Wong, Fletcher, & Leonard, 2006), with that finding possibly being attributable to the structure of Cantonese. Several studies have included the additional step of directly examining the sensitivity and specificity of NWR tests. These studies have supported the diagnostic utility of NWR tests in English (Conti-Ramsden, 2003; Ellis Weismer et al., 2000), Italian (Bortolini et al., 2006), Spanish (Girbau & Schwartz, 2007) and Quebec French (Elin Thordardottir et al., 2011). Results with bilingual children have been inconsistent. Girbau and Schwartz (2008) reported good diagnostic accuracy for children speaking Spanish and English, whereas Kohnert, Windsor, and Yim (2006) reported good specificity with low sensitivity and Gutiérrez-Clellen and Simon-Cereijido (2010), and Summers, Bohman, Gillam, Peña, and Bedore (2010) both found that their NWR measure was significantly affected by previous exposure pattern, as was also indicated by an early study by Thorn and Gathercole (1999). In contrast, a recent study of second-language (L2) learners of Icelandic showed uniformly high levels of NWR performance in children varying widely in the length of residence in Iceland and in performance on standardized tests of Icelandic (Elin Thordardottir & Anna Gudrun Juliusdottir, 2012).

Discrepancies between studies on bilingual children may stem from various factors. One of them is the background of the bilingual children. The US studies (Gutiérrez-Clellen & Simon-Cereijido, 2010; Kohnert, Windsor, & Yim, 2006; Summers et al., 2010) focused on children who were sequential bilinguals and who were of school-age at the time of testing. Although general exposure patterns in terms of home and school language are reported, studies vary in the level of detail of this documentation. Inportantly, although several of these studies suggest that previous exposure exerts an influence on NWR, their primary focus was on diagnostic accuracy rather than on exposure. Secondly, the studies employ several different nonword tests whose lexical and sublexical properties may affect performance as was highlighted in the meta-analysis by Graf Estes, Evans, and Else-Quest (2007) which found that the various tasks used in the studies surveyed were not interchangeable and elicited different effect sizes (see also Archibald & Gathercole, 2006).

Sentence imitation has been shown to have high diagnostic accuracy for the identification of PLI in English and in French (Conti-Ramsden et al., 2001; Elin Thordardottir et al., 2011). It seems intuitive that SI might be more highly affected by bilingual exposure than NWR given that it employs meaningful language as stimuli. The performance of L2 speakers of Icelandic in a longitudinal study confirms this: their performance on SI increased over time along with their progress in the Icelandic language, in contrast to their performance on NWR which was high at all test times (Elin Thordardottir & Anna Gudrun Juliusdottir, 2012).

A number of studies have shown that the rate of language acquisition is strongly affected by amount of language exposure in monolingual as well as bilingual children (de Houwer, 2007; Elin Thordardottir, 2011; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Pearson, 2007; Pearson, Fernandez, Lewedeg, & Oller, 1997). The effect is seen in vocabulary and in grammatical development (e.g. Elin Thordardottir et al., 2006; Jia & Fuse, 2007). As a result, many bilingual children obtain scores similar to those of children with PLI when tested in either one of their languages. In preschool children acquiring French and English simultaneously with carefully documented exposure histories from birth, a strong relationship was found between amount of exposure to each language and vocabulary scores in that language (Elin Thordardottir, 2011). Children with unequal exposure to the two languages evidenced a similarly unequal pattern of performance across the two languages in vocabulary and in morphosyntactic development (Elin Thordardottir, 2012). To the extent that the ability to repeat nonwords and real sentences depends on previous knowledge of the language in which the nonwords and sentences are presented, then children with different exposure histories can be expected to perform in a way that reflects this history.

The relationship between NWR and language knowledge is a complex one and still open to debate. Nonword repetition has been shown to correlate significantly with language proficiency, including both vocabulary and grammar (e.g. Archibald & Gathercole, 2006; Ellis Weismer et al., 2000; Elin Thordardottir, Kehayia, Lessard, Sutton, & Trudeau 2010a). Further, nonwords vary in the degree to which they resemble real words (referred to as word-likeness) and thus presumably in the extent to which real vocabulary knowledge and/or knowledge of frequent word patterns helps in completing the task. A correlational finding, however, does not provide conclusive evidence as to whether vocabulary knowledge aids NWR performance or whether the reverse is true. Longitudinal studies on nonword repetition and vocabulary in typically developing (TD) children suggest that the direction of the effect may change over time as children progress in their linguistic learning. Further, there have been indications that groups of children with TD and PLI may resort to different strategies in repeating nonwords, as shown by L2 learners being insensitive to the difference between word-like and non-wordlike items, unlike monolingual children with PLI (Elin Thordardottir & Anna Gudrun Juliusdottir, 2012).

In bilingual children, the effect of amount of exposure on NWR has mostly been inferred from studies that focus on the identification of PLI, but do not document bilingual exposure in great detail. To date, little research is available on the use of SI measures with bilingual children. More study is required isolating factors such as amount of exposure and item construction to better understand their effect on the performance of monolingual and bilingual children with and without PLI. The purpose of the studies presented here is first to directly assess the effect of previous bilingual exposure on NWR and SI performance of children with TD, and secondly, having clarified this aspect for our NWR and SI measures, to assess their diagnostic accuracy for the identification of PLI in monolingual and bilingual children carefully matched on language exposure.

## 1. Study 1. The relationship between amount of exposure and performance on NWR and SI

The purpose of this study was to examine the effect of varying degrees of bilingual exposure on the ability of simultaneously bilingual children to perform tasks of NWR and SI in each of their languages. For this same group of children, previous studies have shown a strong effect of amount of exposure and vocabulary and morphosyntactic development (Elin Thordardottir, 2011, 2012). It was hypothesized that performance on the two measures of language processing would be less strongly affected by amount of exposure because they do not require linguistic formulation or comprehension. However, SI employs real sentences, and nonwords respect the phonotactic rules of each language such that the French and English nonwords are clearly different and resemble their respective language. Therefore, amount of exposure to these languages was expected to exert some influence on performance. The study evaluated overall performance on SI and NWR in terms of the percentage of items repeated correctly, and also examined whether bilingual exposure has a differential effect on the ability to repeat nonwords of different length. Children with PLI have been shown to perform gradually more poorly as the length of nonword items increase (Ellis Weismer et al., 2000). This effect has been attributed to their deficient processing capacity in comparison to children with typical development. Bilingual children with typical development should not evidence a decrement in performance with increasing word length if this effect is related to poor processing skills. If they do, the effect is more likely to stem from low proficiency in the language of the nonwords.

### *1.1. Participants*

Participants were 84 5-year-old children (mean age 58.31 months, SD 3.91, range 52–69 months), all residents of Montreal, Quebec, and all developing normally according to parent report (no serious concerns in any developmental area, no major illnesses, hospitalizations, or prenatal complications). The children were part of a larger study in which other measures of language knowledge were administered as well. The children included 16 monolingual speakers of English, 19 monolingual speakers of French, and 49 bilingual children with varying relative amounts of exposure to French and English. The children were recruited so as to represent the continuum of relative amount of exposure to English and French and the entire group is included in the analysis on the effect of exposure on performance. In analyses of the effect of the length of

**Table 1**
Study 1 participant characteristics, means and (SD).

| | E | EEF | EF | FFE | F |
|---|---|---|---|---|---|
| *n* | 16 | 16 | 13 | 20 | 19 |
| Age in months | 58.0 | 56.4 | 60.5 | 58.0 | 59.0 |
| | (4.6) | (2.6) | (4.2) | (3.8) | (4.6) |
| Maternal education[a] | 17.9 | 16.9 | 17.0 | 18.2 | 16.3 |
| | (3.2) | (2.7) | (2.3) | (2.6) | (2.3) |
| Nonverbal cognition[b] | 103.9 | 104.2 | 101.3 | 108.3 | 105.6 |
| | (10.7) | (15.8) | (16.8) | (12.8) | (12.3) |
| % English[c] | 98.2* | 76.3* | 51.2* | 21.1* | 1.0* |
| | (2.4) | (9.9) | (7.5) | (8.3) | (0.0) |

E: 95–100% English exposure; EEF: 61–94% English exposure; EF: 40–60% English exposure; FFE: 6–39 English exposure; F: 0–5% English exposure.
* Significant group difference between all groups.
[a] number of years of education completed.
[b] Brief IQ scale of the Leiter International Performance Scale-Revised (Roid & Miller, 1997), standard score. The standard score of this test has a mean of 100 and standard deviation of 15.
[c] percentage of the child's waking hours since birth spent in English communicative environments (the remainder of the child's waking hours being spent in French communicative environments). Significant group difference found between all groups.

nonwords, the children are divided into subgroups of monolingual English-speakers (E, *n* = 16), children having receiving more exposure to English (EE, *n* = 16), children having received equal exposure to both languages (EF, *n* = 13), children having received more exposure to French (FF, *n* = 20) and monolingual French speakers (F, *n* = 19). This grouping is used as well in Table 1 showing background characteristics to demonstrate that participants included children with these different backgrounds, and that across these groups, children are matched on age and SES (maternal education) and nonverbal cognition. However, this grouping is not meant to imply that the continuum of bilingual exposure is made up of discrete groups.

Language exposure was calculated from a detailed parent questionnaire surveying language use in the home and in daycare for each year of the child's life. Questions on exposure ask the parents to list all people in the home, which language they speak and how much they interact with the child. For each year of life, the parents list daycare settings attended, the length of time attended and language(s) spoken. Other settings of significant exposure are reported as well (such as regular weekends with grand-parents). From these data, a single number was computed representing the percentage of the child's waking hours spent in exposure to English over their lifetime – the remaining waking hours being spent in exposure to French as no other languages were part of these children's environments on a regular basis. All children had started regular bilingual exposure before the age of 3 years and were all considered to be simultaneous bilinguals. The children varied in terms of whether they were exposed to both languages at home, or one at home and the other at daycare. The questionnaire and calculation procedure as well as the children's home and daycare language environments are presented in greater detail in Elin Thordardottir (2011). The children's hearing was screened using a portable audiometer.

### 1.2. Procedure

Children were tested individually by trained research assistants. Monolingual children were tested once in their language; bilingual children were tested twice: in English and French, by different examiners who were native speakers of the language they tested. The two sessions were spaced by approximately one week and the order of testing of the two languages was counterbalanced across participants. The children were administered a battery of measures; those of interest in this study are NWR and SI. The NWR tests chosen were existing tests in each language. They were administered from tape; SI was administered live. Scoring was done on-line and verified from recordings. Normative data for children age 4;6 to 5;6 for the French tests were reported by Elin Thordardottir et al. (2010a) and both were shown to have high diagnostic accuracy for PLI in monolingual French-speaking children age 5 (Elin Thordardottir et al., 2011). The diagnostic accuracy of the English NWR test for monolingual children has been reported by Gathercole, Willis, Baddeley, and Emslie (1994). All testing was audio and video-recorded.

(1) SI in French. The task employed is a French adaptation by Royle and Elin Thordardottir (2003) of the Recalling Sentences in Context subtest of the CELF-Preschool (Clinical Evaluation of Language Fundamentals-Preschool; Wiig, Secord, & Semel, 1992). A story supported by images is told and the child is asked to repeat selected sentences (''The moving scene – Le grand déménagement''). The sentences to be repeated increase in length, syntactic complexity and number of propositions as the story unfolds. The scoring procedure was modified and involves the percentage of words repeated correctly,
(2) SI in English – the task was analogous to that used for French. The story employed is from the Recalling Sentences in Context subtest of the CELF-Preschool 2 (''No Juice!''; Wiig, Secord, & Semel, 2004). The scoring procedure was the same as that used for the French SI test.

(3) French NWR test (Courcy, 2000; Elin Thordardottir et al., 2011). This test which was developed for Quebec French includes 40 items, ranging in length from two to five syllables. The test items contain no consonant clusters and no diphthongs and all syllables are CV. The resulting words have a uniform stress pattern with relatively equal stress on each syllable. The complete list of items appears in Elin Thordardottir et al. (2011). The scoring procedure followed that of Dollaghan and Campbell (1998). Words produced are scored in terms of the percentage of phonemes repeated correctly. Distortions and additions are not scored as incorrect, but omissions and substitutions are.

(4) English NWR test, the CNRep (Gathercole et al., 1994). This test is made of 40 items ranging in length from two to five syllables. It was constructed so as to minimize articulatory output demands. In spite of this, the structure of the items is more complex in several ways than that of the French words. Syllable shapes include CV, CCV and CVC. The words contain diphthongs and consonant clusters are found both in word initial, medial and final position. The items are constructed to reflect English dominant stress patterns (two syllable words are strong-weak, three-syllable words are strong-weak-weak, and longer words have variable patterns). The words contain common English derivational morphemes, and in that way resemble real English words. The scoring procedure was the same as that of the French test, and followed that of Dollaghan and Campbell (1998).

The tests were scored on-line. Reliability checks were conducted from videotape for the NWR measures. Fifteen percent of the NWR tests in French and English were selected randomly and were relistened to and rescored by independent scorers, resulting in a correlation between the original and rescored version of 0.914 ($p < 0.01$) for French tests and 0.915 ($p < 0.01$) for English tests.

## 2. Results

Polynomial curve estimation was used to examine the relationship between amount of exposure and performance on the two measures, in French and in English. This analysis includes the entire group of children. The resulting curves are displayed

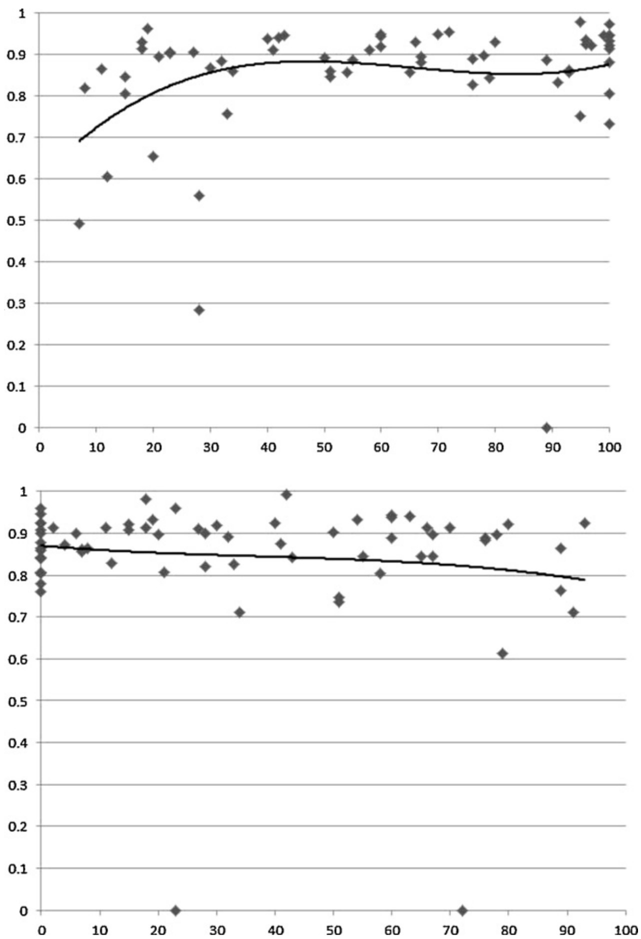


**Fig. 1.** Study 1: nonword repetition scores in English (top panel) and in French (bottom panel) as a function and line of best fit between these scores and percentage of time spent in English.

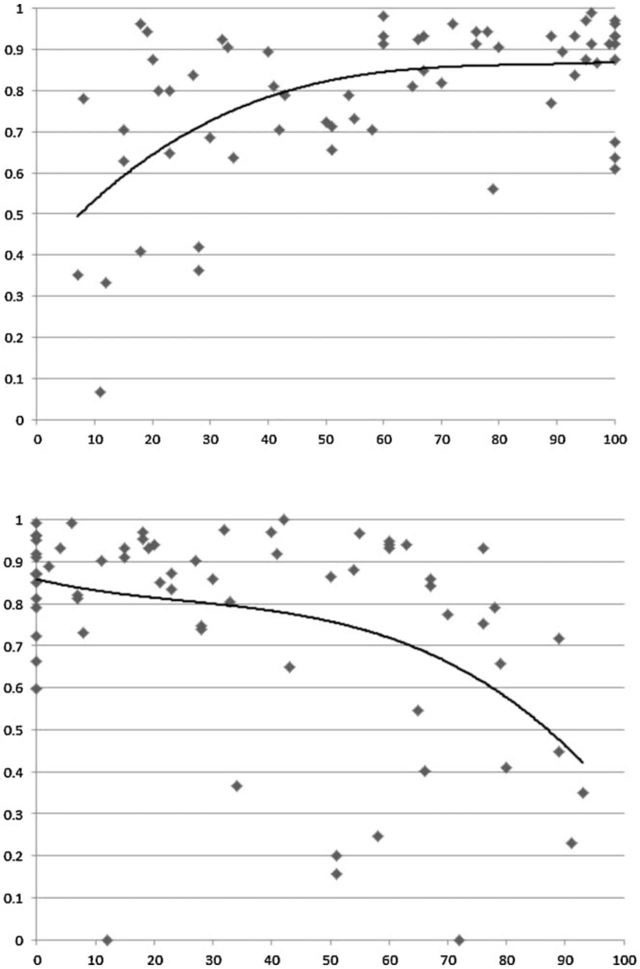


**Fig. 2.** Study 1: sentence imitation scores in English (top panel) and French (bottom panel) and line of best fit between these scores and percentage of time spent in English.

in Figs. 1 and 2. For SI in French, significant linear and nonlinear fits were obtained, with the nonlinear fit stronger (nonlinear: $r = 0.486$, $p = 0.000$; linear: $r = 0.455$, $p = 0.000$). Scores increased with increased exposure to the French language, with the fitted line ranging from approximately 60–90%. Some variability was found around the mean performance, but few children obtained scores much below the mean. The same pattern emerged for English SI (nonlinear: $r = 0.562$, $p = 0.000$; linear: $r = 0.505$, $p = 0.000$). As for the French sentences, performance increased with increased exposure to English with a similar range of scores and comparable variability.

For NWR in French, the association between amount of exposure to French and NWR performance was nonsignificant (nonlinear: $r^2 = 0.221$, $p = 0.205$; linear $r = 0.141$, $p = 0.258$). From the lowest to highest level of exposure to French, fitted scores varied only from approximately 85–88%. Some variability was found, but only very few children obtained scores below 80%. For the English nonwords, both linear and nonlinear fits were significant, but the nonlinear fit was stronger: nonlinear: $r = 0.411$, $p = 0.004$; linear: $r = 0.349$, $p = 0.005$. Across the range of exposure amounts, mean performance, however, varied only a little bit more than that of the French test, or from approximately 80–90% and only few children obtained scores lower than 80%. It is evident from Fig. 1 that there are outliers in the English data, in particular a child with an exposure of 28% to English who received a score of 28% on the nonword test. When this outlier is removed, a significant association between amount of exposure to English and performance on English NWR remains, with a similar strength of association as reported above. Two outliers are found as well in the French scores.

Fig. 3 displays the performance by word length of the bilingual children. For both English (top panel) and French (bottom panel) nonwords, mean performance is shown at each word length for groups of children having been exposed to English only, more English (61–95% English), to French and English equally (40–60% English), to more French (5–29% English) and French only. Word length effects were examined by ANOVAs for English and French words separately, with bilingual exposure group as the between subjects factor and nonword length as the within subjects factor, treated as repeated measures. For the English nonwords (Fig. 3, top panel), the ANOVA revealed a significant effect of group ($F(3, 50) = 3.240$,

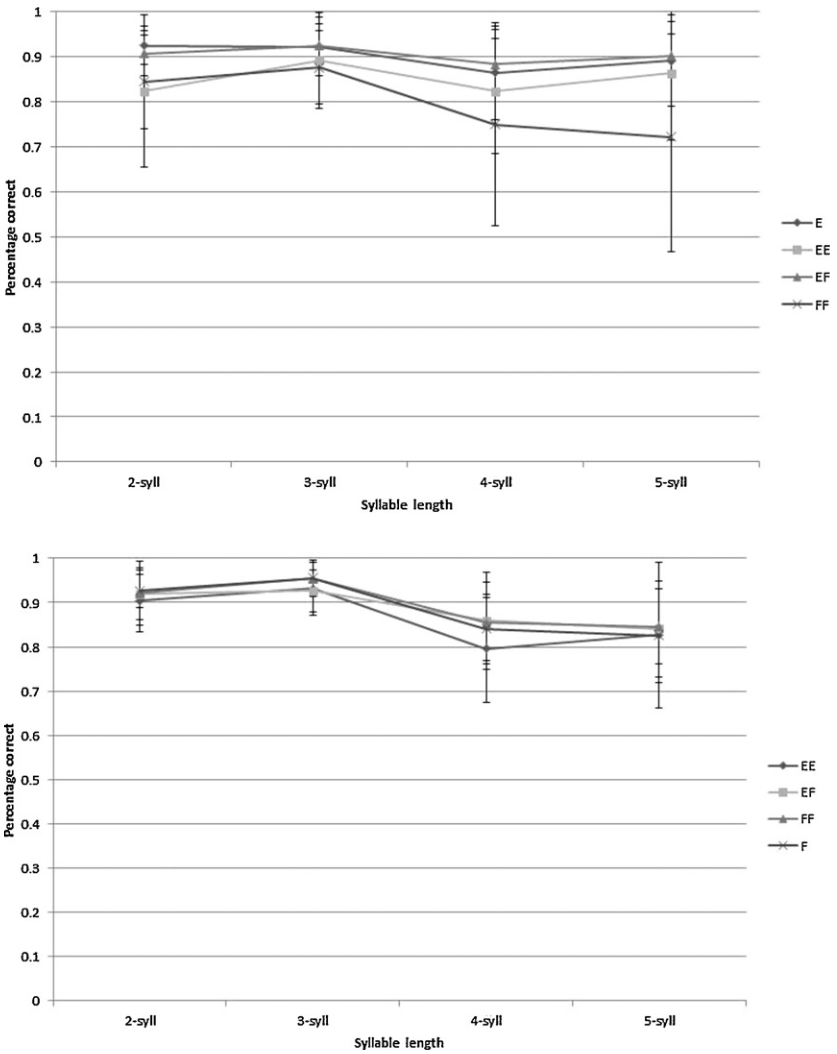


**Fig. 3.** Study 1: nonword repetition scores in English (top panel) and French (bottom panel) broken down by nonword length, for each of the bilingual exposure groups (E: English only, EE: more English than French, EF: equal English and French, FF: more French than English).

$p = 0.030$, $\eta^2 = 0.163$), a significant effect of length ($F(3,150) = 7.709$, $p = 0.000$, $\eta^2 = 0.113$) and a significant length × group interaction ($F(9,150) = 2.332$, $p = 0.017$, $\eta^2 = 0.184$). Post hoc tests on the interaction employing the Tukey–Kramer procedure because of unequal sample sizes revealed that no group differences existed at nonword lengths of two or three syllables. However, at four syllables, the FF group had a significantly lower performance than the EF group and at five syllables, the FF group performed significantly lower than the E, EF and EE groups. For the French nonwords (Fig. 3, bottom panel), the ANOVA revealed a significant effect of nonword length ($F(3,177) = 43.468$, $p = 0.000$, $\eta^2 = 0.424$. No other effect was significant (group $p = 0.598$; group × length interaction $p = 0.673$). Post hoc tests (Tukey–Kramer) revealed no significant difference between the performance of two- and three-syllable words or between four- and five- syllable words. All other pairwise comparisons between word lengths revealed a significant difference.

Lastly, correlations were run between the nonword and SI tests across the two languages for the bilingual children. No significant correlation was found between NWR performance in French and English ($r = 0.226$, $p = 0.130$), or between SI performance in French and English ($r = 0.015$, $p = 0.919$). Within French, SI and NWR were significantly correlated ($r = 0.624$, $p = 0.000$). The same was true for the English versions ($r = 0.555$, $p = 0.000$). Scores on receptive vocabulary in English and French were available for these children as well (see Elin Thordardottir, 2011). Raw scores on English receptive vocabulary (PPVT-III; Dunn & Dunn, 1997) correlated significantly with English NWR ($r = 0.575$, $p = 0.000$) and English SI ($r = 0.656$, $p = 0.000$) but not with the corresponding measures in French (nonwords: $r = 0.200$, $p = 0.177$; SI: $r = -0.046$, $p = 0.759$). Raw scores on receptive vocabulary in French (EVIP, Échelle de vocabulaire en images Peabody – a Canadian French measure of receptive vocabulary; Dunn, Thériault-Whalen, & Dunn, 1993) correlated significantly with NWR in French ($r = 0.483$, $p = 0.001$) and SI in French ($r = 0.752$, $p = 0.000$) but not with the corresponding measures in English (nonwords: $r = 0.032$, $p = 0.833$; SI: $r = -0.120$, $p = 0.422$).

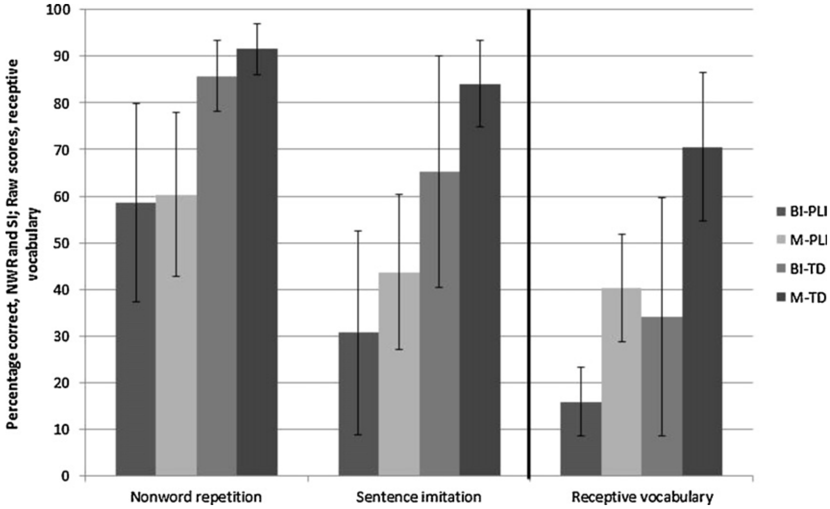


**Fig. 4.** Study 2: group means for each of the four groups of children on nonword repetition (percent phonemes correctly repeated), sentence imitation (percent correct words repeated), receptive vocabulary scores (raw score), and percent correct use of object clitits in spontaneous speech.

## 3. Discussion

A significant association with amount of input was found on three out of four of the language processing tasks: NWR in English and SI in both languages. The notable exception was NWR in French which was not significantly correlated with amount of previous exposure to French. However, the strength of association between NWR in English and exposure is considerably weaker than that reported previously for receptive, and even more so, expressive vocabulary: Elin Thordardottir (2011) reported *r* values of 0.480 and 0.628 for English and French receptive vocabulary, and 0.729 and 0.709 for English and French expressive vocabulary. In comparison, the *r* value for English NWR was 0.411. The association between exposure and SI is of similar strength as that for receptive vocabulary, but less strong than that for expressive vocabulary. Overall, these findings indicate that SI is more highly influenced by previous exposure than is NWR. This was expected given the fact that SI involves meaningful material. In conclusion, the expectation that the processing measures would be less influenced by exposure than vocabulary measures is clearly borne out by the curve estimation analysis for French NWR, but much less clearly so for English nonwords.

A few other observations lend support to a relatively greater immunity from the effects of previous exposure for nonwords than SI, and for nonwords in comparison to vocabulary. First, the range of scores for NWR is small in both languages. As is evident from Fig. 1, the great majority of the children obtain scores above 80%. In the case of English nonwords (top panel of Fig. 1), it is also evident that most of the children with lower performance have had low amounts of exposure to English. The figure suggests that at a critical exposure level of approximately 35–40% of waking hours since birth, 5-year-old children can be expected to perform similarly to native speakers on this particular English nonword repetition test. For French, the bottom panel of Fig. 1 shows an even more striking flat performance rate across levels of previous exposure to French. The steadily high performance across the bilingual continuum on both of these tests shows that high levels of performance can be expected even at fairly low exposure levels (with the above-mentioned difference between the English and French tests). Further, this indicates that these measures may be of high clinical utility as diagnostic measures. In a previous study on 5-year-old speakers of French, the cut-off value set to identify PLI was 82% for NWR and 74% for SI (Elin Thordardottir et al., 2011). In spite of some variation due to bilingualism, the great majority of the bilingual children scored above these values. This does suggest that the ability required to repeat nonwords presupposes some exposure to the language, but not as much exposure as is needed for a high level of vocabulary development.

Second, the level of previous bilingual exposure had little effect on children's ability to repeat nonwords of increasing length. In both English and French, performance decreased somewhat with increasing word length. In French, nonword length affected all groups in the same way. For the English nonwords, somewhat of an effect of previous exposure to English was seen at lengths of 4 and 5 syllables, but it affected only children with more exposure to French than English. Children who had spent up to half of their waking hours in English were not affected by length any more than monolingual speakers of French. These findings indicate bilingual children are not significantly affected by increasing length of nonwords. In this respect, they resemble monolingual children with typical development and differ from monolingual children with PLI (Ellis Weismer et al., 2000). Overall then, the results of this study show that performance on NWR is considerably less influenced by previous exposure than measures of language knowledge such as vocabulary, but that this depends also on the characteristics of the nonword items. Further, the findings show that nonword length does not tax the abilities of bilingual children in the way that taxes the abilities of children with PLI.

Nevertheless, there are also indications in the data from this study that performance on these linguistic processing tasks goes hand in hand with language knowledge. Both NWR and SI are strongly correlated with receptive vocabulary scores in the same language, and more strongly so than with amount of exposure to that language (however, receptive vocabulary is strongly correlated with amount of input). At the same time, nonword repetition and SI are not correlated with receptive vocabulary in the other language. Further, NWR skills in French are not correlated with NWR skills in English for the same children, nor are SI skills. This lack of within-subject correlation is no doubt related in part to the relatively low range of scores. However, taken together with other findings, it seems reasonable to conclude that these tasks do not appear to tap directly or exclusively into an underlying processing ability. They may tap into such an ability, but each test gets at this processing ability by means of stimuli that remain colored by the language in which they were constructed. The extent of this coloring varied, however, across the measures, being least for the French NWR task.

The relatively greater influence of amount of exposure on SI than NWR is intuitive: SI contains meaningful language, nonwords do not. The difference between the French and English nonword tasks in terms of their association with amount of input is likely to be related to the particulars of the item construction. Testing in French and English involved the same group of bilingual children. The difference between the French and English tests can, therefore, not be attributed to background characteristics of the children. The children varied in bilingual exposure, but were otherwise well matched in terms of age, nonverbal cognition and SES. The two languages being tested are both spoken by strong language groups in Montreal – they can both be viewed as having majority language status. The French nonword items are simpler than the English ones in several respects – they do not contain clusters or diphthongs, they contain CV syllables only, and they have a simpler stress pattern which is uniform across the words and which reflects the typical stress pattern of French. These words were not designed to be either word-like or non-wordlike. As a whole, they are far from representing the phonological complexity of the French language. However, in spite of their simple structure, individual words are rather word-like. One feature of the NRT (Dollaghan & Campbell, 1998) which has been found to be insensitive to dialectal differences within English, is an attempt to make the words nonwordlike by having none of the syllables correspond to real English words and by using uncharacteristic stress patterns. The findings suggest that simple nonwords are preferable to more complex words when testing bilingual children in that they are less affected by bilingual exposure. It should be noted, however, that the difference between the French and the English nonwords may not be as large as it appears. Even though performance on the English nonwords was significantly correlated with amount of previous exposure to English, the range of scores was small and the great majority of the children scored in the 80–90% range regardless of their bilingual status.

### *3.1. Study 2. Comparison of the diagnostic accuracy of French measures of language processing and French measures of language knowledge*

Study 1 indicated that the French NWR test stood out as the measure least affected by amount of bilingual exposure, indeed showing no significant correlation. This same test has been shown to have high diagnostic accuracy for the identification of PLI in monolingual French-speaking children (Elin Thordardottir et al., 2011). Taken together, the sensitivity to PLI and insensitivity to previous exposure should make this nonword test an ideal diagnostic measure for the identification of PLI in bilingual children. Given its greater susceptibility to exposure, the French SI test should be less accurate for this purpose, and might or might not be more accurate than a traditional measure of accumulated linguistic knowledge. Study 2 examines the diagnostic accuracy of the French NWR and SI tests in comparison to a test of receptive vocabulary by comparing groups of monolingual and bilingual children with and without PLI. The study addresses the effect of language exposure by closely matching the PLI and TD groups on this variable.

### *3.2. Participants*

Four groups of children participated, with 14 children in each group: (1) bilingual children with PLI (BI-PLI), (2) bilingual children with typical language development (BI-TD), (3) monolingual children with PLI (M-PLI), and (4) monolingual children with typical language development (M-TD). The language common to all children was French (Montreal, Quebec). Participant characteristics are shown in Table 2. The four groups were formed by selecting matched children from previous studies conducted in the lab of the first author within which they were all administered a series of common language measures using identical procedures (Elin Thordardottir, 2011; Elin Thordardottir et al., 2011; Elin Thordardottir, Rvachew, & Ménard, 2010b). All groups were equivalent in age, with mean age close to 5 years ($p = 0.344$). Nonverbal cognition was measured by the brief IQ scale of the Leiter International Performance Scale-Revised (Roid & Miller, 1997), and was well within normal limits for all groups, with no significant group differences ($p = 0.438$). A significant group difference was found for maternal education, with mothers of children in the two PLI groups having, on the average, significantly fewer years of formal education than mothers of children in the two TD groups.

The monolingual children were required to have had no significant regular exposure to any language other than French, defined as less than 5 h per week. The history of language exposure since birth was surveyed by the same parent questionnaire as that used in Study 1. Children in the two bilingual groups (typical development and PLI) were matched pairwise on amount of exposure to French over their lifetime as well as on their age of onset of bilingualism. A close pairwise match was available for both amount of exposure and age of onset in all cases except for two children in the BI-TD group whose age of onset was later than 36 months. All the children in the BI-PLI group and all but two children in the BI-TD group

**Table 2**
Study 2 participant characteristics, means and (SD).

| | BI-PLI | M-PLI | BI-TD | M-TD |
|---|---|---|---|---|
| Age in months | 59.8 | 61.5 | 57.4 | 59.9 |
| | (5.7) | (7.2) | (4.6) | (6.3) |
| Maternal education | 14.7* | 13.3* | 17.3 | 17.7 |
| Years | (4.6) | (3.0) | (3.1) | (1.1) |
| Nonverbal cognition | 91.9 | 98.1 | 101.2 | 90.4 |
| | (11.6) | (18.4) | (8.1) | (17.5) |
| % exposure to French | 22.8 | 26.7 | | |
| | (16.8) | (20.7) | | |
| Range | 5–66 | 4–70 | | |
| Age of onset of bilingual exposure | 17.9 | 20.5 | | |
| | (14.1) | (18.5) | | |
| Range | 0–36 | 0–55[a] | | |

* A significant group difference was found for maternal education.
[a] All children in this group except two were in the range of 0–30 months.

had an age of onset of regular bilingual exposure between 0 and 36 months. The bilingual children were classified as simultaneous bilinguals based on their age at first regular exposure to French occurring before 3 years of age. As Table 2 shows, their average exposure to French was 25% of their waking hours over their lifetime indicating that French was their weaker language. The bilingual groups did not differ significantly on either overall quantity of exposure to French ($p = 0.595$) or age of onset of bilingualism ($p = 0.674$).

The bilingual children all spoke French but differed in their other language. Children in the TD group had English as their primary language, whereas for the PLI group, the primary language included a number of minority languages (Tamil, Urdu, Spanish, Punjabi, Japanese, Singhalese, Arabic, Dutch, and Russian). The reason for this discrepancy was the availability of these groups of children with BI-PLI and M-PLI from previous studies which could be matched closely on other crucial variables and the fact that bilingual children with PLI who speak French and English are extremely hard to recruit in Montreal as it is customary for these children to be channeled towards either one language when they start showing signs of impaired language development. This is possible for these children because schools are in place in both French and English from preschool and through university. In that sense, the bilingual group with PLI is quite representative of the population of bilingual children with PLI, which in fact comprises fairly few children who speak English and French. A TD comparison group reflecting the same variety of languages would have been more ideal. However, a number of previous studies on the effects of bilingualism on children's cognitive abilities and linguistic processing include children with diverse language combinations (e.g. Bialystok, 2010; Bialystok & Feng, 2009; Moreno, Bialystok, Wodniecka, & Alain, 2010). This is not ideal in that it introduces a confound of linguistic variability whose effects are not well known. At the same time, homogeneous language groups are not necessarily reflective of other bilingual populations. Because this confound exists in this study for the B-PLI group, comparisons involving this group need to be interpreted with greater caution than comparisons involving the remaining three groups.

Children in the TD groups were reported by their parents to have had normal development in language and other developmental areas. Their parents reported no concerns about their developmental milestones, no major illnesses or hospitalizations. Children in the PLI groups were identified as having PLI as part of the original studies in which they took part. Children in the M-PLI group were referred to the study in which they participated based on a previous clinical evaluation done by certified Speech-Language Pathologists at a major Montreal hospital (Elin Thordardottir et al., 2011). Children in the BI-PLI group participated in a study on the efficacy of language intervention for bilingual children (Elin Thordardottir et al., 2010b). The data reported on here for these children are those from their initial intake evaluation for the intervention study. These children were referred based on previous evaluation by certified Speech-Language Pathologists at a major Montreal hospital and in Montreal French public schools. In both studies, diagnostic status was verified as part of the studies in which the children participated and included testing of language, nonverbal cognition, hearing and developmental history. Evaluation of children in the BI-PLI group included spontaneous language sampling in their home language conducted by an interpreter who transcribed the sample and computed mean length of utterance in words (MLUw). Based on this evaluation, the French evaluation, and clinical history, the presence of PLI was confirmed in these children.

### *3.3. Procedure*

Children were tested individually by trained research assistants who were native speakers of Quebec French and who spoke only French during the test session and gave no indication of knowing other languages. The following measures were administered to all the children in this study:

(1) French NWR test (Courcy, 2000). The same test items, administration and scoring procedure as that employed in Study 1.
(2) SI using the same test adaptation and scoring procedure as that used in Study 1 (Royle and Elin Thordardottir's adaptation of a subtest of the CELF-P; Wiig et al., 1992).

(3) The Échelle de vocabulaire en images Peabody (EVIP; Dunn et al., 1993). This is a test of comprehension of single words in which the child selects from four pictures the one that corresponds to the word spoken by the examiner. The EVIP is a standardized Canadian French version of the Peabody Picture Vocabulary Tests-Revised (PPVT; Dunn & Dunn, 1997).

## 4. Results

### 4.1. Group means

Means and standard deviations for each of the groups on each of the measures are displayed in Fig. 4. Group differences were examined by one-way ANOVA analysis with Tukey post hoc tests. For NWR, a significant group effect was found: $F(3,50) = 18.995$, $p = 0.000$, $\eta^2 = 0.53$. Post hoc tests revealed that the PLI groups each differed significantly from both TD groups ($p = 0.000$ for each comparison). The two PLI groups, BI-PLI and M-PLI, did not differ significantly from each other ($p = 0.990$) and the two TD groups, BI-TD and M-TD, did not differ significantly from each other ($p = 0.704$). NWR thus yielded a significant difference between PLI and TD groups regardless of bilingual exposure. The group effect was also significant for SI $F(3,51) = 20.890$, $p = 0.000$, $\eta^2 = 0.55$. Post hoc testing revealed the same pattern as for NWR. The BI-PLI group differed significantly from each TD group ($p = 0.000$). The M-PLI group also differed significantly from each TD group ($p = 0.021$). The two PLI groups did not differ significantly from each other ($p = 0.301$), nor did the two TD groups ($p = 0.055$), although the comparison just failed to reach significance. For the EVIP, the group effect was again significant: $F(3,52) = 26.449$, $p = 0.000$, $\eta^2 = 0.60$. Post hoc tests revealed that the BI-PLI group differed from each of the other three groups (BI-TD, $p = 0.026$; M-TD, $p = 0.000$; M-PLI, $p = 0.002$). The M-PLI group differed significantly from the BI-PLI group and the M-TD group ($p = 0.000$), but not from the BI-TD group ($p = 0.778$). The two TD groups differed significantly from each other ($p = 0.000$). These results show that the EVIP distinguishes the BI-PLI group from the TD groups. However, the EVIP does not distinguish the M-PLI group from both TD groups. Further, the two PLI groups differed from each other, as did the two TD groups.

### 4.2. Sensitivity and specificity

To examine the sensitivity and specificity of these measures for the bilingual children, we employed the cut-off scores established previously for the identification of PLI in 5-year-old francophone children (Elin Thordardottir et al., 2011). Accordingly, children were labeled as having PLI if they scored lower than 82% on NWR, lower than 74% on SI, or obtained a standard score lower than 100 on the EVIP relative to the published norms. It should be noted that both monolingual groups in the present study participated in the previous study examining sensitivity and specificity. This study extends the examination to bilingual children with and without PLI.

The sensitivity and specificity of each measure individually and for combinations of measures for the monolingual and bilingual children is displayed in Table 3. Both NWR and SI achieved a high level of sensitivity, correctly detecting PLI in 85 and 92% respectively, of the children previously identified as having PLI, regardless of whether the children were bilingual or monolingual, and using the same cut-off score for both groups. Specificity was high for both of these measures for monolingual children (100 and 86% for nonwords and SI respectively), but dropping to only fair to poor for bilingual children (79 and 57% respectively). Specificity was better for NWR than SI for both monolingual and bilingual children, but SI was the measure that yielded the best sensitivity by itself. By itself, the EVIP yielded good levels of sensitivity and perfect specificity for the monolingual group. The EVIP had poor specificity for the bilingual group. In our previous study of monolingual French-speaking children, the combination of NWR and the EVIP produced higher precision than either measure alone (Elin Thordardottir et al., 2011). In the present study, any combination of two measures including NWR, SI and the EVIP brought sensitivity to 100% for both monolingual and bilingual children, and specificity to 100% for monolingual children. No combination of two measures improved the specificity for bilingual children, with the best result obtained by NWR alone, at 79%.

**Table 3**
Sensitivity and specificity of each measure: EVIP, nonword repetition, sentence imitation, and for combinations of measures (nonword repetition and sentence imitation combined with each of the other measures). Sensitivity refers to the proportion of children with PLI (shown for bilingual and monolingual groups separately) who were correctly identified by each measure as having PLI. Specificity refers to the proportion of children with TD (shown for bilingual and monolingual groups separately) who was correctly identified by each measure as not having PLI.

| Group | BI-PLI | M-PLI | BI-TD | M-TD |
|---|---|---|---|---|
| | Sensitivity | Sensitivity | Specificity | Specificity |
| Nonword repetition % correct[a] | 85% | 92% | 79% | 100% |
| Sentence imitation[b] | 92% | 93% | 57% | 86% |
| EVIP[c] | 100% | 86% | 43% | 86% |
| Nonword repetition or EVIP | 100% | 100% | 79% | 100% |
| Nonword repetition, or sentence imitation | 100% | 100% | 79% | 100% |
| Sentence imitation or EVIP | 100% | 100% | 57% | 100% |

[a] Cut-off score: 82%. Nonword repetition scores were missing for one child in group BI-PLI and one child in group M-PLI.
[b] Cut-off score 74%. Sentence imitation scores were missing for one child in group BI-PLI.
[c] Cut-off score: standard score of 100 relative to published norms.

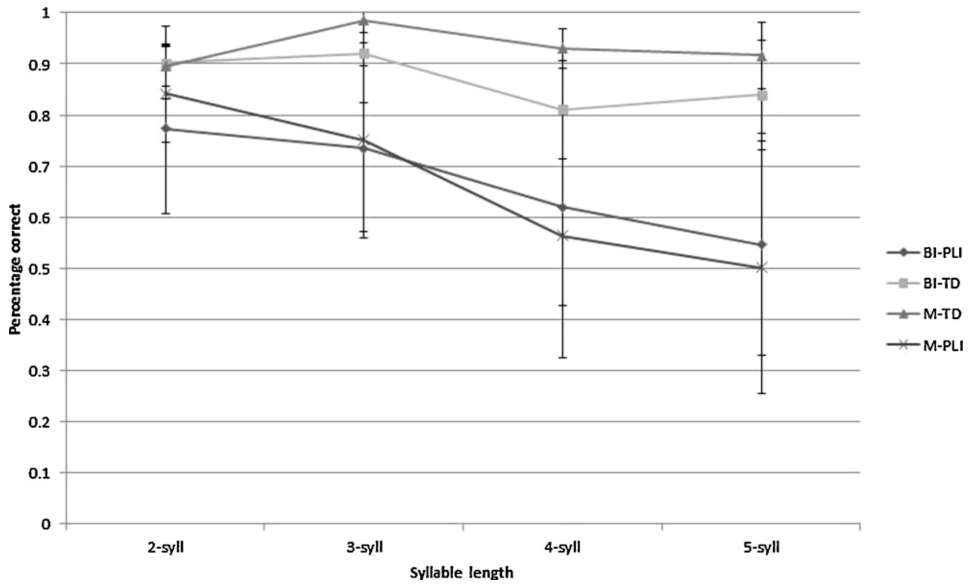


**Fig. 5.** Study 2: performance on the French nonword repetition test as a function of word length by the four groups of children.

### *4.3. Effect of the length of nonwords*

The NWR performance of each group of children is broken down by word length in Fig. 5, showing the mean performance of each group at words of two, three, four, and five syllables. The groups were compared in a mixed-model ANOVA with Group as the between-subjects factor and word length as the within-subjects factor treated as repeated measures. This analysis produced a significant effect of Group ($F(3,47) = 16.101$, $p = 0.000$, partial $\eta^2 = 0.507$), a significant effect of Word Length ($F(3,47) = 49.076$, $p = 0.000$, partial $\eta^2 = 0.511$), and a significant Group $\times$ Word Length interaction ($F(9,47) = 10.989$, $p = 0.000$, partial $\eta^2 = 0.412$). Pairwise post hoc comparisons using Tukey's method revealed the source of the interaction to be that increasing length essentially affected only the PLI groups significantly. For both PLI groups, performance decreased significantly with each increase in length of one syllable except between two and three syllables. In contrast, no significant differences were found between means at different word lengths for the M-TD group, and the only significant difference for the B-TD was between nonwords of three and four syllables (note that the mean for this group is higher at three syllables than two syllables). It is noted also that for the PLI groups, performance decreased systematically from two to five syllables. For the TD groups, performance fluctuated somewhat, but did not show a systematic decrease with increasing length. Comparing the groups at each word length, the PLI groups scored significantly lower than the two TD groups at each length except at two syllables. At each length, the two PLI groups did not differ significantly from each other, and the two TD groups did not differ significantly from each other, with the exception of a significant difference between the M-TD and B-TD groups at four syllables. Overall, therefore, a pattern of a clear separation between PLI and TD groups was maintained across word lengths past two syllables, and was more pronounced as nonwords got longer.

## 5. Discussion

The inclusion of all four groups, including monolingual and bilingual children, each with and without PLI, allows the separate examination of the effect of bilingualism versus language impairment on performance on the tasks involved. Comparing the four groups, it is seen that group means of both NWR and SI accurately distinguish PLI groups from TD groups, regardless of bilingual exposure. In contrast, on the EVIP measure of receptive vocabulary, bilingual children with PLI score significantly differently from the monolingual TD group, but the two PLI groups also score significantly differently from each other, whereas the M-PLI and B-TD groups are not significantly different from each other. The EVIP, therefore, correctly separates only the BI-PLI group from the M-TD group and does not distinguish M-PLI from BI-TD. In fact, the M-PLI group has a higher mean performance than the BI-TD group.

Alternatively, a more conservative interpretation of the data would include only the monolingual children and the bilingual group that was homogeneous in terms of language combination. This includes the M-TD, M-PLI and B-TD groups. Comparison of these groups clearly shows that all four measures distinguished the M-PLI group from the M-TD group. However, the EVIP did not distinguish the M-PLI group from the B-TD group. In contrast, both NWR and SI distinguished the M-PLI group from both TD groups, thus distinguishing PLI from TD regardless of bilingual status. As a result, the superiority of the language processing measures is evident even if the BI-PLI group is not considered.

Going beyond group means, the study also analyzed the sensitivity and specificity of each of the measures in order to ascertain the proportion of children correctly identified by each measure as having PLI or TD. Sensitivity ranged from fairly good to very good for all the measures, using the criteria proposed by Plante and Vance (1995). SI had the single best sensitivity at 92 and 93% for bilingual and monolingual children and NWR was a bit lower at 85 and 92% respectively. By combining either of these measures with each other or with the EVIP, sensitivity was brought to 100%. All the measures had specificity in the acceptable range for the monolingual children. For the bilingual children, specificity was unacceptably low for the EVIP and SI. The highest specificity level was obtained for NWR, at 79%, which is considered fairly good. No combination of measures could bring the specificity rate to a higher level than NWR alone. Again, a more conservative interpretation would exclude the B-PLI group. It is not clear how the results for this group would change if they all had an English-French background, or whether they would. It is possible that this group did worse on average than French-English peers would have because their home language is a minority language. Speaking a minority language versus a majority language does affect performance level in that language, as shown by Gathercole and Thomas (2009) for children speaking Welsh versus English. However, the language being tested here is not the minority language itself but French which is the official language of Quebec. It is thus hard to say how performance would be affected by these circumstances. However, it is possible that the fact that the B-PLI children speak different home languages did not have a significant effect on the data. First, Study 1 indicated that performance on the French NWR test was not significantly influenced by bilingual exposure. Further, a recent study showed the performance of L2 speakers of various origins on an Icelandic NWR test was uniformly high (Elin Thordardottir & Anna Gudrun Juliusdottir, 2012). These questions await further research.

The sensitivity and specificity results of this study support the use of processing measures for the identification of PLI in bilingual children, with various combinations of measures yielding perfect sensitivity and with NWR providing a much better specificity than any of the other measures alone or in combination. The better specificity of NWR compared to the other measures is in line with its relative immunity to the effects of previous amount of exposure to French (see Study 1) and its sensitivity to language impairment (see Elin Thordardottir et al., 2011). The combination of these two characteristics allows NWR to distinguish between the presence and absence of PLI in monolingual children and a large number of bilingual children.

The results suggest that the inclusion of the French NWR test should be strongly considered in the evaluation of bilingual children for the purpose of identifying PLI. In this study, however, specificity remained a challenge compared to sensitivity, with 79% still leaving 21% of children with TD misidentified as having PLI. In general, these results are consistent with those of Girbau and Schwartz (2008) who reported high levels of diagnostic accuracy for Spanish-English bilingual children on NWR. However, they are inconsistent with those of Kohnert et al. (2006) who reported higher specificity than sensitivity for bilingual children, and those of Gutiérrez-Clellen and Simon-Cereijido (2010), who found NWR to be more highly influenced by previous bilingual exposure than was the case here in Study 1 or Study 2. These differences may be due to several factors. The first one is the test items. The French NWR task used in this study was simple in terms of phonological complexity, syllable structure and stress pattern. This may have contributed to its relative immunity to amount of exposure to French. At the same time, this task is sensitive to some particular difficulty that children with PLI have but which children with TD do not have. These two features are what give it its high degree of accuracy with bilingual children.

Another factor that may contribute to discrepancies between our studies and previous ones is the age of the children, with our participants being younger than those in many of the previous studies who included school-age children. Yet another aspect is the previous linguistic exposure of the participating children. Studies vary in whether this is documented in terms of exposure pattern (home language-school language), amount of exposure in more detailed form, or in terms of language dominance. In the present studies, this documentation was based on a very detailed history form and children in both bilingual groups were closely matched on both amount of exposure and age of onset. Finally, unlike previous studies, this study focused on simultaneous bilinguals rather than sequential bilinguals. It is possible that school-age children who are proficient readers and therefore have greater phonological and metalinguistic awareness, approach the task of repeating nonwords or sentences in a different way than younger children. Further study is required to clarify whether these factors are responsible for the differences between studies.

Gutiérrez-Clellen and Simon-Cereijido (2010) as well as Summers et al. (2010) concluded from their studies that, if used with bilingual children, NWR should be administered in both languages of the child. This follows the general recommendation that only by assessing both languages can a complete picture of a bilingual child's linguistic proficiency be obtained (e.g. ASHA, 1985; CASLPA, 1997; Fredman, 2006). We fully agree with this general view. However, our results indicate that NWR scores in only one language gave sufficient information to indicate a strong likelihood of the presence or absence of PLI. What is even more noteworthy is that such a conclusion could be reached with a fairly high degree of accuracy from assessment only in French, which was for all the bilingual children the language to which they had received less exposure. If these results are replicated by other studies, perhaps by the use of test items with simple structure, the use of NWR tests and SI test could prove to be a valuable contribution to increasing the accuracy of identification of PLI in bilingual children and in reducing the levels of misdiagnosis. For a more comprehensive documentation of language level, pattern of strengths and weaknesses across language domains, which are required for intervention decisions, a more elaborate set of language tests in both languages would still be required.

## 6. Conclusions

The two studies reported on here have examined the effect of bilingual exposure on the performance of children with and without language impairment on two measures focusing on language processing, NWR and SI and compared this with their performance on measures of accumulated language knowledge. The results of this study suggest that language processing measures are less affected by amount of exposure to a language than are measures of language knowledge and that they do hold promise as diagnostic measures for the identification of PLI in bilingual children. However, the findings also show that particular measures differ in this respect and that more research is needed on the influence of the characteristics of particular nonword items on performance.

Further, the results shed some light on the direction of causality in the low processing scores typical of children with PLI. A large number of studies published over the years have shown that children with PLI have significant deficits in NWR in comparison to TD peers (e.g. see reviews in Coady & Evans, 2008; Graf Estes et al., 2007). Many of these studies present correlational results and there have been contradictory results regarding the direction of the relationship between vocabulary size and the ability to repeat nonwords (see Coady & Evans, 2008). Other findings have also been advanced that support a causal relationship between the manipulation of processing load and linguistic performance (e.g. Elin Thordardottir, 2008; Grela & Leonard, 2000; Montgomery, 2000; Montgomery & Evans, 2009). To the extent that NWR taps into linguistic processing, the results of the present studies clearly indicate that deficient processing skills are a characteristic of PLI, but not a typical characteristic of children of the same age whose language skills are distributed over two languages. That is, bilingual children and children with PLI each obtained low scores of language knowledge in French compared to norms for their age, but for different reasons. Bilingual children have spent considerably less time than monolingual counterparts in French environments, leading to only part of their total language proficiency being in French. The fact that the bilingual children with TD did not differ significantly from their monolingual TD counterparts on language processing skills (NWR and SI) shows that the reason for their low language scores (receptive vocabulary) is not attributable to poor processing skills, but rather to their more limited previous exposure to the language being assessed. Importantly, the pattern of scores of the bilingual children with TD indicates that the fact that they have significantly low language scores in French does not result in low performance on French language processing measures. In contrast, both groups of children with PLI evidence deficient language processing scores, whether or not they are bilingual. This dissociation of processing skills and accumulated language knowledge is seen clearly in Fig. 5, which shows that word length exerted a strong effect on children with PLI (monolingual and bilingual) whereas both groups of children with TD were essentially unaffected, as revealed statistically by the significant Group x Word Length interaction. Thus, high scores on NWR were associated with TD status, and low scores are associated with PLI status, in each case regardless of bilingual status. These results for the TD children are consistent with those of the larger group of bilingual children with TD in Study 1, which also showed that bilingual exposure did not lead to poorer performance on longer French nonwords.

The findings presented here help clarify the role of processing limitations tapped by NWR in PLI in that they suggest that such limitations are present fairly independently of accumulated language knowledge given that some threshold level of knowledge has been reached. In spite of significantly lower French language scores compared to the M-TD group, the BI-TD group possesses sufficient language processing ability to score similarly to the M-TD group on French language processing, even for the longer nonwords. In contrast, the M-PLI group does significantly more poorly on French language processing than the BI-TD group, in spite of scoring similarly and even higher than the BI-TD group on the French language measures. Overall, the results support a causative role of language processing in PLI, suggesting that such limitations are one of the factors that underlie these children's low attainment of language knowledge.

### Acknowledgments

The studies reported on in this article were funded by a grant from the Canadian Language and Literacy Research Network (CLLRNet) awarded to Elin Thordardottir (PI), Susan Rvachew and Mela Sarkar, a grant from the Ministère de la santé et des services sociaux, Québec (MSSS) awarded to Elin Thordardottir and Eva Kehayia as PIs, and a grant from the Social Sciences and Humanities Research Council of Canada, awarded to Elin Thordardottir. The authors thank the children and parents who participated in this study.

### Conflict of Interest Statement

The authors reported no conflicts of interest with the research described in this article.

## Appendix A. Continuing education

1. Amount of input was observed to have a different effect on performance on language knowledge measures and language processing measures
   a) amount of input affected performance on processing measures more
   b) amount of input affected performance on language knowledge measures more
   c) amount of had the least effect on nonword repetition scores
   d) b) and c)

2. True or false: Nonword repetition and sentence imitation are processing measures that are free of language
3. The repetition of nonwords taxes children with PLI more than bilingual children
   a) children with PLI perform overall worse than bilingual children on NWR
   b) children with PLI perform worse as nonword length increases whereas bilingual children do not
   c) bilingual children perform significantly worse than monolingual children on NWR
   d) and b)
4. True or false: Nonword length has a greater effect on children with PLI than children with typical language development.
5. Why is it more difficult to identify PLI in bilingual children than in monolingual children?
   a) Bilingual children score higher on standardized tasks of language knowledge since they have had more language input.
   b) Bilingual children score lower on standardized tasks of language knowledge despite having typical language development.
   c) Language development of bilingual children differs to a very large extent to that of monolingual children.
   d) Due to cultural differences in the assessment situation.

## References


Archibald, L., & Gathercole, S. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research, 49*, 970–983.

American Speech-Language-Hearing Association (ASHA) (1985). Clinical management of communicatively handicapped minority language populations [Position Statement]. Available from: www.asha.org/policy.

Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology, 46*, 93–105.

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*, 93–100.

Bortolini, U., Arfé, B., Caselli, M. C., Degasperi, L., Deevy, P., & Leonard, L. (2006). Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders, 41*, 695–712.

Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA). (1997). Position paper on speech-language pathology and audiology in the multicultural, multilingual context. *JSLPA* September.

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

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment. *Journal of Speech, Language and Research, 46*, 1029–1037.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*, 741–748.

Courcy, A. (2000). Conscience phonologique de apprentissage de la lecture. Unpublished doctoral dissertation. Montréal: Université de Montréal.

de Bree, E., Rispens, J., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics & Phonetics, 21*, 935–944.

de Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics, 28*, 411–424.

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

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Dunn, L., Thériault-Whalen, C., & Dunn, L. (1993). *Échelle de vocabulaire en images Peabody: Adaption française du Peabody Picture Vocabulary Test*. Toronto, ON: PsyCan.

Elin Thordardottir. (2008). Language-specific effects of task demands on the manifestation of specific language impairment: A comparison of English and Icelandic. *Journal of Speech, Language, and Hearing Research, 51*, 922–937.

Elin Thordardottir. (2011). The relationship between amount of bilingual exposure and vocabulary development. *International Journal of Bilingualism* http://dx.doi.org/10.1177/1367006911403202.

Elin Thordardottir. (2012). Use of grammatical morphology by simultaneously bilingual 3-and 5-year olds. *Talk presented at the Bilingual and Multilingual Interaction conference*. Bangor, Wales: ESRC Research Centre on Bilingualism. March–April.

Elin Thordardottir, & Anna Gudrun Juliusdottir (2012). Icelandic as a second language: A longitudinal study of language knowledge and processing by school-age children. International Journal of Bilingual Education and Bilingualism, 1–25, http://dx.doi.org/10.1080/13670050.2012.693062.

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-olds. *Canadian Journal of Speech Pathology and Audiology, 34*, 5–16.

Elin Thordardottir, Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., et al. (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.

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*, 1–21.

Elin Thordardottir, Rvachew, S., & Ménard, (2010). Efficacy of language intervention for bilingual children. *Poster presented at the annual ASHA conference*.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, 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*, 865–878.

Fredman, M. (2006). Recommendations for working with bilingual children – Prepared by the Multilingual Affairs Committee of IALP. *Folia Phoniatrica et Logopaedica, 58*, 458–464.

Gathercole, S., Willis, C., Baddeley, A., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological memory. *Memory, 2*, 103–127.

Gathercole, V., & Thomas, E. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition, 12*, 213–237.

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

Girbau, D., & Schwartz, R. G. (2008). Phonological working memory in Spanish–English children with and without specific language impairment. *Journal of Communication Disorders, 41*, 124–145.

Graf Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*, 177–195.

Grela, B., & Leonard, L. (2000). The influence of argument–structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech, Language, and Hearing Research, 43*, 1115–1125.

Gutiérrez-Clellen, V., & Simon-Cereijido, 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.

Hansson, K., Forsberg, J., Löfqvist, A., Mäki-Torkko, E., & Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with Specific language impairment. *International Journal of Language and Communication Disorders, 39*, 401–422.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language and Hearing Research, 50*, 1280–1299.

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

Montgomery, J. (2000). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 43*, 293–308.

Montgomery, J., & Evans, J. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 52*, 269–288.

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics, 23*, 564–579.

Pearson, B. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics, 28*, 399–410.

Pearson, B., Fernandez, S., Lewedeg, V., & Oller, K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18*, 41–58.

Plante, E., & Vance, R. (1995). Diagnostic accuracy of two tests of preschool language. *American Journal of Speech-Language Pathology, 4*, 70–76.

Roid, G., & Miller, L. (1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting Co.

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.

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

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.

Thorn, A., & Gathercole, S. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology, 52A*, 303–324.

Wiig, E., Secord, W., & Semel, E. (1992). *Clinical evaluation of language fundamentals-preschool*. San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovitch Inc.

Wiig, E., Secord, W., & Semel, E. (2004). *Clinical evaluation of language fundamentals-preschool* (2nd ed.). San Antonio, TX: The Psychological Corporation.