Applied Psycholinguistics (2020), **41**, 107–131
doi:10.1017/S0142716419000420

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Sentence repetition and non-word repetition in early total French immersion

Maureen Scheidnes*

Memorial University of Newfoundland
*Corresponding author. Email: mscheidnes@mun.ca

(Received 13 September 2018; revised 19 July 2019; accepted 06 August 2019; first published online 25 November 2019)

Abstract

Recent research has focused on bilingual children's performance on non-word repetition (NWR) and sentence repetition (SR) tasks, but it remains unclear how their scores can be expected to vary as a function of language exposure, which creates challenges for developing age-appropriate performance expectations. With the goal of examining the impact of limited language exposure on these tasks, French NWR and SR performance from 33 first graders (mean age 6 years, 10 months) in early total French immersion in English-speaking Canada was compared to prior work on bilinguals acquiring French in France. With a mean length of exposure of 1 year, 7 months, but a mean cumulative length of exposure of only 3 months, the children in immersion have much less daily exposure to French than the bilinguals in France. The results showed that children in immersion patterned with the other bilinguals for NWR, but had much weaker SR performance. Within-subjects analyses revealed that, for SR, the children in immersion had stronger scores on *wh*-questions and relative clauses, which suggests that these structures may be less sensitive to language exposure.

Keywords: child L2 acquisition; early total French immersion; language exposure; non-word repetition; sentence repetition

One major challenge in assessing language development in bilingual children involves setting appropriate expectations for language performance based on age, as well as patterns of language exposure.[1] Recent work on sequential and simultaneous bilinguals acquiring French in contexts in which it is spoken as a majority (first language; L1) language have reported that sentence repetition (SR) varied by some exposure-related measure, while non-word repetition (NWR) did not (de Almeida et al., 2017; Thordardottir & Brandeker, 2013). The current paper aims to shed light on the impact of language exposure on NWR and SR by extending the study of these tasks to a group of children acquiring French in a minority (second language; L2) context. While the bulk of the literature on distinguishing between typical and atypical language development in bilinguals has focused on sequential bilinguals in L1 minority/L2 majority contexts (Marinis, Armon-Lotem, & Pontikas, 2017), a focus

on children in minority L2 immersion allows for additional insight into how bilingual children with limited L2 exposure can be expected to perform compared to other bilinguals and how different language properties can affect this performance. Additional insight into which properties are mastered quickly regardless of the acquisition context can lead to the design of less biased repetition tools for bilinguals. Furthermore, understanding which language properties lead to weaker performance in typically developing (TD) bilinguals with limited exposure can allow for the exclusion of these language features from assessment tools or for a more informed interpretation of difficulty with such properties in these learners.

Background

SR and NWR in French majority language contexts

SR and NWR tasks have been shown to be effective at identifying atypical language development in monolingual children (Conti-Ramsden, Botting, & Faragher, 2001), but it is unclear how performance should be expected to vary as a function of language exposure in bilinguals. Thordardottir and Brandeker (2013) reported on the NWR and SR performance of simultaneous bilingual children in Montreal, Canada. Their study included 84 5-year-olds, 49 of whom were acquiring French and English. The bilingual children varied in the relative quantity of exposure to each language and represented a continuum of more exposure to English to more exposure to French. With respect to SR performance, the authors reported strong correlations with exposure in both languages. With respect to NWR performance, the authors found a significant correlation between English NWR performance and exposure to English, but they argued that this correlation was weaker than the one between SR and exposure. For French NWR, the authors found no significant correlation between this task and amount of exposure to French.

Thordardottir and Brandeker (2013) explained that the makeup of the English NWR, which included more phonologically complex properties and stimuli that closely resembled English words, was likely behind the stronger link between NWR and exposure to English compared to the correlation observed between French NWR and exposure to French. While much research on NWR tasks does not discuss in detail the phonological properties of the stimuli, Thordardottir and Brandeker's results suggest that such information is important in understanding the link between exposure and NWR performance in bilinguals (see Gallon, Harris, & van der Lely 2007, who raise similar issues for monolingual assessment). In terms of their SR results, there were presumably specific morphosyntactic properties that presented greater obstacles to correct repetition than others; however, the sentences in these SR tasks were designed to increase in length, vocabulary difficulty, and syntactic complexity as the test progresses, which does not (easily) allow for the impact of exposure on these individual language properties to be teased apart.

Thordardottir and Brandeker's results highlight the need for greater insight into which language properties lead to particular difficulty in repetition tasks in TD bilinguals, especially in those with limited language exposure. With such issues in mind, the LITMUS (Language Impairment Testing in Multilingual Settings)

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

series of tasks was designed to be appropriate for bilingual language assessment. This was done as part of a collaborative effort within the COST Action IS0804, and the tasks were adapted into different languages (Armon-Lotem, de Jong, & Meir, 2015). The tasks used in the present study include the French non-word repetition task, LITMUS-NWR-French, and the French sentence repetition task, the LITMUS-SR-French. The LITMUS-NWR-French task aims to distinguish between typical and atypical development in bilingual children by targeting phonological properties that vary in crosslinguistic frequency and complexity (see dos Santos & Ferré, 2018, for additional details). The stimuli are divided into two groups: language-dependent and language-independent items. Non-words falling into the language-dependent category involve arguably complex phonological properties that are found in French, but are less common crosslinguistically (sC clusters and [l] in coda position). Non-words falling into the language-independent category involve onset clusters (consonant–consonant–vowel; CCV), which are quite common crosslinguistically (Maddieson, Flavier, Marsico, & Pellegrino, 2011). Cluster reduction has been shown to occur in children with language impairment (Gallon et al., 2007), so such non-words may present difficulty to affected children, but perhaps not to TD bilingual children, as these properties will likely be present in the L1. Moreover, stimuli were kept relatively short (maximum three syllables) so as to specifically target phonological abilities. The French SR task from the LITMUS series, the LITMUS-SR-French, was designed to be lexically simple (i.e., it uses early acquired and frequent words) while targeting morphosyntactic properties that have been shown to be difficult for children with language impairment (i.e., tense marking, *wh*-movement, and clausal embedding, for more details, see Fleckstein, Prévost, Tuller, Sizaret, & Zebib, 2018).

Using the LITMUS-NWR-French task, dos Santos and Ferré (2018) examined NWR performance with 30 TD bilinguals whose mean age was 6 years, 3 months (6;03; range: 5;04–8;02) and 13 bilingual children with language impairment whose mean age was 6;06 (range: 5;04–8;01), in addition to 10 monolingual French-speaking children with language impairment (mean age: 7;05, range: 6;02–8;05) and 14 TD monolingual controls (mean age: 5;08, range: 5;04–6;03). The bilingual children had either L1 English or L1 Arabic. All bilingual children were living in France, and most were simultaneous bilinguals. Contrary to their predictions, the authors found that all groups (monolingual children with and without language impairment and bilingual children with and without language impairment) had lower performance on the language-independent items compared to the language-dependent items. It is important to note that the TD bilingual children had significantly higher performance than the bilinguals with language impairment, and overall NWR performance was not related to language exposure in the bilingual groups. In terms of the specific effects of the phonological properties of the non-words, the TD bilingual group outperformed the bilingual group with language impairment with respect to length of stimuli (i.e., number of syllables) and clusters, suggesting that these are interesting properties to focus on in bilingual development. Moreover, in a similar study, Ferré, dos Santos, and de Almeida (2015) showed that the stimuli with [l] in coda position appeared to be sensitive to language impairment, but not bilingualism.

De Almeida et al. (2017) report on data from both the LITMUS-NWR-French and the LITMUS-SR-French with 61 TD bilinguals and 21 bilinguals with language impairment, some of whom were also included in dos Santos and Ferré (2018) and Fleckstein et al. (2018). For NWR, their results showed that this task had good sensitivity (81%) and specificity (79%) and that there was no link between NWR and dominance or exposure factors. However, the SR task resulted in 76% sensitivity and 72% specificity; furthermore, there were significant correlations between SR and exposure factors (use of French at home, use of French during activities and with friends, and a Language Dominance Index, which included measures of language exposure in the past and daily language use). This NWR/SR asymmetry confirms the pattern reported by Thordardottir and Brandeker (2013) for bilinguals in Montreal. De Almeida et al. (2017) concluded that phonological development proceeds faster than morphosyntactic development and that more needs to be done to better understand which language properties underlie the variation in SR performance. Taken together, the findings reported by Thordardottir and Brandeker (2013) and de Almeida et al. (2017) suggest that more research is required to (a) confirm whether the stronger performance in NWR compared to SR extends to French minority contexts (i.e., in contexts where there are fewer possibilities for language use) and (b) better understand which language properties contribute to variation in SR performance.

Children in French immersion as models for children with limited language exposure

The bilinguals in Thordardottir and Brandeker (2013) and de Almeida et al. (2017) were all acquiring French in contexts in which the language enjoys majority status and learners have ample access to native speakers. In such contexts, it can be difficult to recruit enough participants with limited exposure to French. For example, de Almeida et al. (2017) noted the weak SR results (an identical repetition score of 20%) of a child with exposure to French of less than 12 months; however, only four children from their study had less than 24 months of exposure, so conclusions about performance from children with particularly limited exposure could not be made. However, the immersion children in the present study all have equally limited exposure to French, making for interesting models from which to study the impact of limited language exposure. Thus, the goal of examining such learners is to better understand within-group variability (i.e., differences in SR and NWR performance in immersion children compared to bilingual children with more exposure) and within-subject variability as a function of language property (i.e., where immersion learners' strengths and weaknesses lie in these tasks).

Early total French immersion in English-speaking Canada

In English-speaking Canada, children can elect to enroll in early total French immersion programs in which all academic instruction is given in French. These learners are acquiring L2 French in a classroom setting with their teacher as their primary model speaker, but they are nevertheless exposed to natural language data via instruction delivery. The French immersion classroom teachers are often L2

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

speakers themselves, but it is unclear what exact impact this has on L2 oral language abilities. These minority L2 learners are generally thought to be in a quite different learning context compared to children acquiring a majority L2 (Paradis, Genesee, & Crago, 2011). For English-speaking children in French immersion, they elect to receive instruction in French, which is not the case for sequential bilingual children from family immigration who speak a minority L1 at home and must acquire the majority L2 at school. Furthermore, for children with L1 English in English-speaking parts of Canada, it is highly likely that most, if not all, classmates will share the same L1, which is often not the case for majority L2 learners.

Research on older children who were educated in early French immersion programs in Canada have shown that academic and L1 language outcomes are on a par with those of children who have been educated in the L1 (see Genesee, 2004, for a review). However, in terms of oral language production abilities in French, prior evidence suggests that these learners do not attain native-like or near-native-like skills (Genesee, 2004). While the bulk of this research has been carried out on older students, studies on oral language abilities in the early grades suggest that these learners have particular difficulty with gender marking, verb morphology, and object clitics (Adiv, 1980; Grüter, 2005; Harley, 1979, 1992). These same properties have been shown to cause difficulty in sequential bilinguals in L2 French majority contexts as well (see Prévost, 2009, for a review), but it is often assumed that the majority L2 children will eventually achieve production skills that closely resemble those of native speakers, even if this takes time. Overall, the work on oral French language abilities of children in early total French immersion has focused on properties that are prone to error; therefore, much less is known about L2 properties that might already be mastered during the early school years. Understanding which L2 properties are acquired early could allow for better identification of (a)typical language performance in the early stages of bilingual development. This in turn could allow for the development of language tasks that are less biased against bilingual children with less L2 exposure.

The present study

The overall aim of this study is to better understand how limited language exposure impacts the performance of bilingual children on a NWR task and a SR task, two assessment tools that are commonly used in clinical settings. Quantitative analyses will involve the comparison of the performance of children acquiring L2 French in an early total French immersion program in English-speaking Canada to that of French-speaking bilinguals in France. This comparison will be done using cutoffs that have been proposed in previous work with bilinguals in France (i.e., de Almeida et al., 2017). The cutoff data will not be used for diagnostic purposes, but as a way of comparing performance between these two bilingual groups. Within-subjects analyses of the NWR and SR data from the children in immersion will be carried out in order to better understand how specific language properties impact repetition performance and to look for relative strengths and weaknesses in these learners' early L2 abilities.

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

The following research questions will be addressed:

1. How do the children in early total French immersion (i.e., who have limited daily exposure to French) compare to bilingual children in France on French NWR and SR tasks?
2. Of the language properties targeted by the NWR and SR, which stand out as being particularly difficult or easy for children in early total French immersion?

Given the limited daily exposure to French in this context, stronger performance with a certain language property will be argued to mean that the property develops quite independently of language exposure. In contrast, weaker performance with a certain language property will be argued to mean that the property is quite sensitive to variation in language exposure.

Method

Recruitment procedure and exclusionary criteria

Recruitment for this study was done in collaboration with teachers and staff at an elementary school with an early immersion program located in St. John's, Newfoundland. Consent forms and parental questionnaires were sent home with each first grader, and those whose parents returned a completed consent form participated in the study. The parental questionnaire was used to find out if there were any concerns about the child's language development and to collect information about language use outside of school and socioeconomic status (parental education). Table 1 presents the general characteristics of the immersion group.

As this study focuses on typical L2 development by L1 English-speaking children, the following exclusionary criteria were used: (a) English language performance below age criterion per the Clinical Evaluation of Language Fundamentals—5th edition screening tool (Semel, Wiig, & Secord, 2013); (b) hearing or other impairment that may affect language or the ability to do the tasks as reported by questionnaire; (c) L1 language other than English; and (d) a score below the 9th percentile on Raven's Coloured Progressive Matrices (Raven, Raven, & Court, 1998). A total of 42 children participated in the study; however, participants were excluded from the present analyses for the following reasons: not meeting the age criterion for English ($n = 3$), hearing impairment ($n = 2$), and native language other than English ($n = 1$). In addition, other children's data were excluded for the following reasons: technical problems during data collection ($n = 1$) and refusal to complete both repetition tasks ($n = 2$). The final data analysis was therefore carried out on 33 children.

Taking mother's education as a measure of socioeconomic status, the overall level for these children was mid to high, without much variability: out of 33 mothers, 26 had completed a 4-year university degree, 13 of whom had completed a master's degree or higher. All had completed high school.

Exposure to French

French speakers are a clear minority in St. John's, Newfoundland, with 0% of the population being monolingual French-speaking and just 9.2% of the population

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 1. General characteristics of the children in French immersion

	Mean	*SD*	Range
Age	6;10	0;3	6;4–7;5
Age of onset	5;4	0;3	4;10–5;8
Length of exposure[a]	1;7	0;1	1;5–1;9
Cumulative exposure[b]	0;3	0;0.11	0;2–0;4

[a]Length of exposure is the difference between age at time of testing (age) and age of onset.
[b]Cumulative exposure considers the total number of hours of language exposure of the life of the child (see Unsworth, 2013).

speaking both French and English (Statistics Canada, 2017). The information from the parental questionnaires confirmed that all of the children had acquired English as an L1 and that none of the participants were exposed to French outside of school on any regular basis. Parents also reported that regular exposure to French began upon entering half-day kindergarten in the same early immersion program. Thus, the children's classroom teacher, who is a non-native speaker of French, was their primary model French speaker.

Data collection took place during the second half of first grade. As can be seen in Table 1, this means that the overall length of exposure for these participants was between 17 and 23 months, which represents the time elapsed from the beginning of kindergarten until the time of testing. However, if cumulative exposure is taken into account per Unsworth (2013), exposure is about 2–3 months.[2]

Standardized test scores

Standardized receptive measures of vocabulary and morphosyntax from the Nouvelles Épreuves pour l'Examen de Langage (Chevrie-Muller & Plaza, 2001) battery were obtained so as to have independent measures of the French abilities of these children. This battery was normed in Europe, but is commonly used by clinicians and researchers in Canada. Expressive measures were obtained initially, but these tasks proved to be quite difficult for these children, so collection of expressive measures was discontinued. The mean raw scores and *z* scores for the receptive vocab and morphosyntax tasks are reported in Table 2. These results suggest performance that is well below those of the TD bilinguals in France for whom Fleckstein et al. (2018) reported a mean *z* score of –0.6 ($SD = 1.3$, Range: –4.8 – 1.3) on the same receptive morphosyntax task.

Materials

LITMUS-NWR-French

The LITMUS-NWR-French task includes a total of 50 non-words, which fall into two different categories: language-dependent and language-independent items (see Table 3). Dos Santos and Ferré (2018) and de Almeida et al. (2017) reported on a longer version of the LITMUS-NWR-French (71 items), while the current

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 2. French standardized test scores of receptive vocabulary and morphosyntax from the Nouvelles Épreuves pour l'Examen de Langage (Chevrie-Muller & Plaza, 2001)

Task (out of total possible raw score)	N^a	Raw scores Mean (*SD*); *range*	*Z* scores Mean (*SD*); *range*
Receptive vocabulary (out of 36 points)	33	19.52 (3.39); *12 – 26*	–13.12 (2.96); *–19.7 – –7.8*
Receptive morphosyntax (out of 8 points)	33	3.45 (1.37); *1 – 7*	–3.10 (1.35); *–5.5 – 0.16*

[a]Number of participants.

Table 3. Description of LITMUS-NWR-French stimuli (50 items)

	Segments	Syllable types	Examples
Language independent ($n = 21$)	[p, k, f, l, a, i, u]	CV, CCV, CVC#	*pilu*, *fli*, *pukif*, *flaplu*, *plaklu*
Language dependent ($n = 29$)	same as language independent, plus [s]	same as language independent, plus #sC, #sCCV, sC#, and internal coda	*spu*, *skla*, *fips*, *plal*, *filpu*, *kufalpi*, *pliks*

study reports on data from a shorter version of the same task (50 items), which includes a subset of the items from the longer version. Recall that language-dependent items ($n = 29$) involve complex phonological properties that are found in French, but are less frequent crosslinguistically. Language-independent items ($n = 21$) are composed of segments and syllable types that are more frequent crosslinguistically. Thus, by definition, the language-dependent and language-independent items were designed to differ with respect to syllable structures involving sC (*spu*), Cs (*fips*), as well as [l] in coda position (*filpu*; see Table 3 for more details), but the two categories also differed in other ways. Specifically, the language-independent items had more multisyllabic stimuli than the language-dependent items. Out of the total 20 stimuli with only one syllable, 17 were in the language-dependent category while only three were language-independent items. Thus, the language-dependent items contain a greater proportion of monosyllabic stimuli (17/29, 59%) than the language-independent items (3/21, 14%). Another difference concerned the number of CCV clusters per word. There were only two stimuli with two clusters (*flaplu* and *plaklu*), and both of these were language-independent items. Because of these proportions, the independent/dependent distinction will be analyzed separately from the other phonological factors of syllable length, clusters, and [l] in coda position. In terms of scoring the NWR, the only method used here involves identical repetition of the target stimuli.

LITMUS-SR-French

The LITMUS-SR-French (Prévost, Tuller, & Zebib, 2012) includes 30 test sentences, which are divided into five different properties (see Table 4). Within each of the five

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 4. Details of the LITMUS-SR-French items (30 total test items)

	Property	Subproperty	Number	Length in syllables
Monoclausal	Present	3SG	3	6.7
		3PL	3	
	Past	3SG	3	8.7
		3PL	3	
	Wh-question	*Who*	3	7
		Which + noun	3	
Biclausal	Complement clause	Nonfinite	3	11.8
		Finite	3	
	Relative clause	Subject	3	11.3
		Object	3	

properties, two subproperties are distinguished. While there are some differences in length across the five main properties (compare, e.g., *wh*-questions and relative clauses), subproperties within one property type did not differ in length (e.g., 3SG past vs. 3PL past). Test stimuli are presented in Appendix A.

Two different coding schemas will be considered for the analysis of the SR data: identical repetition and target structure score. For identical repetition, a score of 1 was assigned when the child repeated the test sentence verbatim. In the event of any sort of modification (substitution, omission, addition, or word order change), an identical repetition score of zero was given. However, phonological errors, such as *grissé* [grise] for *griffé* [grife] "scratched" in (1), which did not lead to a change in meaning, were counted as identical repetition.

(1) J'ai vu le chat qui a grissé [= griffé] la vache.
I have seen the cat who has scratched the cow.

For target structure scoring, a score of 1 was given if a targeted subproperty (Table 4) was correctly repeated, even if there were other deviations from the test sentence, such as gender or lexical errors (as in [2]). In other words, the target structure score was given a 1 if certain structure (e.g., the object relative clause in [2]) was repeated as such. In another example, the stimuli in (3) is from a monoclausal sentence with past tense, but the child omitted the auxiliary. Hence, this item receives a target score of zero. The reduction of a multiclausal sentence to monoclausal one would also lead to a target score of zero (as in [4]).

(2) Je vois le garçon que le [=la] fille a poussé.
I see the boy that the.MASC [= the.FEM] girl has pushed.

(3) Les parents ∅ [=ont] rangé les jouets.
The parents ∅ [=have] put.away the toys.

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

(4) Le papa a [=sait] très bien conduire la voiture.
The dad has [=knows] very well drive.INF the car.
"The dad knows how to drive the car very well."

Procedure

Data collection

The test items for both tasks were prerecorded by a female speaker of European French, and the audio files were incorporated into a PowerPoint presentation. Given that European French is frequently spoken by teachers in French immersion programs in St. John's (many teachers attend French programs in France or St. Pierre et Miquelon), the use of these European French recordings was deemed appropriate. Furthermore, standard European and Canadian French varieties do not differ for the morphosyntactic properties targeted in the task.

Two sessions were organized for each child: one in French and the other in English. Testing took place in a quiet room at school. Trained research assistants with the appropriate language background (including the author) administered the tasks. The French and English sessions for the same child were never carried out by the same experimenter. During the French session, two receptive subtests from the standardized battery Nouvelles Épreuves pour l'Examen de Langage, as well as the LITMUS-SR-French and LITMUS-NWR-French were administered. During the English session, the Clinical Evaluation of Language Fundamentals—5th edition screener and Raven's Coloured Progressive Matrices (Raven et al., 1998) were administered. Each session took about 20–25 min. The French session was generally administered first.

Data analysis

The data were transcribed and coded using Phon (Hedlund & Rose, 2019) by trained research assistants with the appropriate language background (some of whom had been involved in the data collection). Transcriptions were checked in their entirety by a second transcriber. Transcriptions that were initially performed by non-native-speaking research assistants were then verified by a native French-speaking assistant. Any initial disagreements were resolved after discussion, and in difficult cases, benefit of the doubt was given to the child.

Statistical tests were run on raw scores using software package *R* (R Core Team, 2017). For NWR, paired *t* tests were used to determine any within-group differences between language-dependent and language-independent items. Following the method described in Crawley (2007), a generalized linear model with a binomial distribution and logit error family was used in order to see which phonological properties contributed significantly to the variance in performance. For LITMUS-SR-French, a robust one-way repeated-measures analysis of variance (Wilcox, 2005, function *rmanova*) was used to examine the effect of property on target structure scores. This was done using the *WRS2* package for *R* (Mair, Schönbrodt, & Wilcox, 2017) and was followed by robust post hoc tests (Wilcox, 2005, function *rmmcp*), which were used to evaluate differences across property types (Table 4). Robust tests were run using a default of 20% trimmed

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

means. Family-wise error was controlled for using the calculation described in Wilcox (2005) and the critical (corrected) p value will be used to determine statistical significance for each comparison.

Results

LITMUS-NWR and LITMUS-SR performance in immersion children compared to other bilinguals

Figure 1 presents the NWR and SR identical repetition data of the children in immersion, which shows that their NWR performance was quite high ($M = 88\%$), with little variability ($SD = 0.07$). When compared to the LITMUS-NWR-French cutoff that was proposed in previous research (i.e., 80%), the immersion children patterned with the TD bilinguals from France. A total of 87% of the children in immersion (29/33) and 79% of TD bilinguals in France scored above the cutoff of 80% (see Table 5, which summarizes the cutoffs and sensitivity rates from de Almeida et al., 2017). This pattern of results suggests that the LITMUS-NWR-French is neutral to variation in language exposure.

Turning to the SR data, the immersion children had a mean identical repetition rate of 30.9% ($SD = 14.9\%$), which is much lower than the NWR score. Furthermore, none of the children in immersion performed above 60% (raw score of 18/30), which was the cutoff proposed by de Almeida et al. (2017) for TD bilinguals living in France.[3] Thus, the children in immersion overlapped considerably with the TD bilinguals in France for NWR, but not for SR.

Phonology: Strengths and weaknesses

Recall that the NWR stimuli were divided into two main groups: (a) language-dependent items and (b) language-independent items. The results of a paired t test revealed that children had significantly lower scores on language-independent items versus language-dependent items, $t\ (32) = 3.38$, $p = .002$, 95% CI [0.02, 0.09]. The language-dependent and language-independent scores are displayed in Figure 2.

One possible reason that the children had higher scores on language-dependent items concerns length and number of clusters. As mentioned below, the language-dependent items are on average shorter than the language-independent items. Furthermore, only two non-words contained two clusters (*plaklu* and *flaplu*), and these were both language-independent items. These two language-independent items were repeated correctly by the fewest number of children (17/33 and 21/33, respectively). Given the potential impact of length (monosyllabic vs. multisyllabic stimuli) and clusters (no cluster vs. one or two), regression was performed to better understand their role on the NWR task. A third factor, the presence or absence of [l] in coda position (word-medial or final), was also included in the analysis, given that Ferré, dos Santos, and de Almeida (2015) had argued that this created difficulty for children with language impairment.

The mean identical repetition scores for items containing these three phonological factors are presented in Table 6. The regression results revealed that each factor was a significant predictor of NWR performance (see Table 7).[4] Thus, although overall variability in the current data set was low ($SD = 0.07$), it appears that items

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 5. Cutoff points and specificity rates for LITMUS-NWR-French and LITMUS-SR-French as reported by de Almeida et al. (2017)

	Cutoff	Specificity	Sensitivity
LITMUS-NWR-French	80%	79%	81%
LITMUS-SR-French	60%	72%	76%

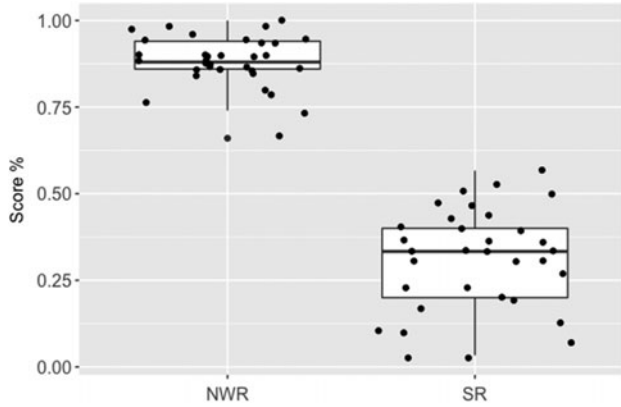

Figure 1. Proportion of identical repetition in the non-word repetition (NWR) and sentence repetition (SR) tasks.

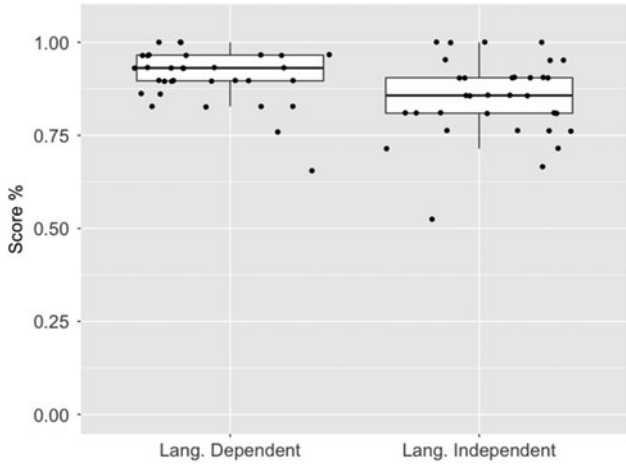

Figure 2. Comparing scores on language-dependent and language-independent NWR items.

with two and three syllables, clusters, and [l] in coda position were associated with somewhat lower performance in this task. The role played by length and clusters most likely explains why lower scores have been observed for language-independent items.

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 6. Non-word repetition performance (% correct) by phonological property

	Length in syllables		Clusters		[l] coda	
	1	2–3	None	1–2	No	Yes
Mean	.94	.84	.90	.85	.89	.81
SD	.05	.10	.08	.11	.07	.19
Range	.75–1.0	.60–1.0	.64–1.0	.64–1.0	.68–1.0	.33–1.0

Table 7. Results of generalized linear model including length (monosyllabic vs. multisyllabic), CCV clusters (no clusters vs. 1 or 2 clusters), and [l] in word-internal or word-final coda position

	B (SE)	Odds ratio	2.5% CI	97.5% CI
Constant	3.30 (0.21)*			
Syllable length (>1)	−1.17 (0.19)*	0.31	0.21	0.45
Clusters (≥1)	−0.72 (0.16)*	0.49	0.35	0.67
[l] coda	−0.82 (0.21)*	0.44	0.29	0.66

Note: $R^2 = .198$ (Hosmer and Lemeshow), $\chi^2\ (3) = 66.25$, $p < .001$. *$p < .001$.

Morphosyntax: Strengths and weaknesses

While the immersion children patterned with TD bilinguals acquiring French in France for NWR, their SR performance was much weaker. Applying the 60% identical repetition cutoff (see Table 5) would mean that all 33 of the immersion children, who are TD children according to L1 scores, would score in the atypical range. Clearly, identical repetition scores from the SR task are much more sensitive to variation in language exposure than those of the NWR task (see Figure 1).

In order to better understand the relative strengths and weaknesses in SR performance in the immersion children, a more qualitative analysis of the SR data was carried out. This revealed that many errors had little to do with the syntactic structure of the test items. For example, errors involving grammatical gender, which is often said to be notoriously difficult for L2 learners, occurred frequently: on average, each child made 9 gender errors (a total of 304 gender errors were made across all children). In an identical repetition score analysis, these errors obscure any relative difficulty or facility with other language properties. For example, it could be that most of the errors in the biclausal sentences are related to gender and not to embedding. Therefore, the remaining analyses will involve the target structure score.

The boxplots in Figure 3 show target structure scores by property and reveal particularly weak scores on sentences targeting monoclausal sentences with present or past tense marking and complement clauses. A robust repeated-measures analysis of variance revealed a significant effect of property, $F\ (3.29,\ 65.73) = 18.70$, $p < .001$, and post hoc tests with corrected alpha for multiple comparisons revealed that the target structure scores of complement clauses and past and present tense were significantly weaker than those of *wh*-questions and relative clauses (see Figure 3 and Table 8 for statistical results). Furthermore, the difference between *wh*-questions and relative clauses was not significant, suggesting that many of these

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Table 8. Post hoc statistics for comparing target structure scores across five structures of the LITMUS-SR-French

Structure	Ψ	2.5% CI	97.5% CI	*p*	Critical *p*	Significance
Wh-question vs. Relative	0.05	−0.66	0.75	.833	.05	*ns*
Complement vs. Relative	−1.90	−3.02	−0.79	.000	.01	*
Complement vs. *Wh*-question	−1.86	−2.96	−0.75	.000	.01	*
Past vs. Relative	−1.67	−2.81	−0.52	.000	.01	*
Past vs. *Wh*-question	−1.62	−2.76	−0.47	.000	.01	*
Present vs. Relative	−0.76	−1.52	0.00	.005	.01	*
Present vs. *Wh*-question	−0.90	−1.63	−0.18	.001	.01	*
Complement vs. Present	−0.95	−1.73	−0.18	.001	.01	*
Complement vs. Past	−0.33	−1.38	0.71	.327	.03	*ns*
Past vs. Present	−0.76	−1.85	0.33	.039	.02	*ns*

Note: CI, confidence interval. *P* value (*p*) is the value that is uncorrected for multiple comparisons. The critical *p* is the value to which uncorrected *p* is compared and which determines statistical significance.

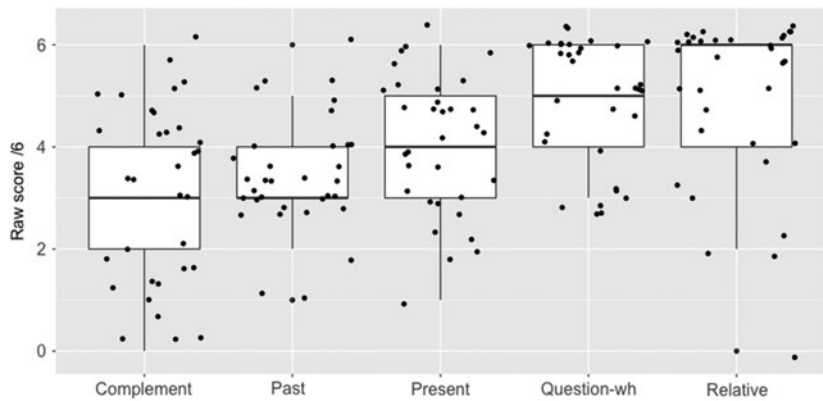

Figure 3. Target structure performance on the LITMUS-SR-French by property.

children have already mastered structures involving *wh*-movement in their L2. Taken together, these results suggest that these immersion children (i.e., with limited L2 exposure) have difficulty with complement clauses and verb morphology, but perform well on relative clauses and *wh*-questions.

Figure 4 presents histograms for the 10 subproperties. The relatively weaker target structure scores for present and past monoclausal items reflected problems repeating the 3PL marker on the main verb (5) or omission of the 3PL auxiliary (6). Only 1 child was able to repeat the 3PL auxiliary all three times, and fewer than 10 children were able to repeat all three 3PL present tense items with the correct verb morphology. Within the complement clause items, children performed particularly

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

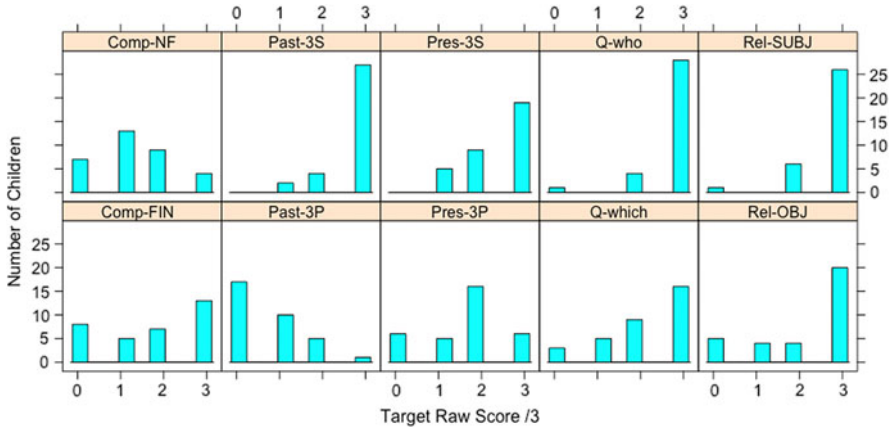

Figure 4. Frequency of target structure scores by subproperty in the LITMUS-SR-French. Comp-NF, non-finite complement clause. Comp-FIN, finite complement clause. Q-who, object question with who. Q-which, object question with which + noun. Rel-SUBJ, subject relative clause. Rel-OBJ, object relative clause.

poorly on non-finite complement clauses. A very frequent error observed in this case involved the substitution of a past tense auxiliary for the matrix verb, thereby reducing the sentence to a monoclausal one (7)–(8). In terms of the performance patterns with *wh*-movement, while scores on object relative clauses were somewhat lower than subject relative clauses and while *which* object question scores were lower than *who* object questions, even these structures, which have been argued in the literature to be particularly complex (e.g., Friedmann, Belletti, & Rizzi, 2009; Friedmann & Novogrodsky, 2011), did not present as much difficulty as the items targeting 3PL morphology or complement clauses.

(5) Les enfants prend [=prennent] un bain.
The children take.3SG [=take.3PL] a bath.

(6) Les enfants ø [=ont] fermé la porte.
The children ø [=have] closed the door.

(7) Le lapin a [=veut] mangé la salade maintenant.
The rabbit has [=wants] eaten the lettuce now.

(8) La maman a [=sait] très bien dessiné le lapin.
The mom has [=knows] very well drawn the bunny.

Discussion

The overall goal of this study was to better understand the role of (limited) language exposure on NWR and SR task performance. The guiding research questions were as follows: (a) how do children in immersion perform on repetition tasks relative to previously used cutoffs for distinguishing between typical and atypical performance; and (b) what are the relative strengths and weaknesses in their performance?

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

NWR and SR performance in immersion children compared to other bilinguals

The first research question was addressed by extending the cutoffs proposed for SR and NWR in previous work with bilinguals in French majority contexts to children attending an early total French immersion program in English-speaking Canada, thus allowing for increased insight into how language exposure impacts repetition task performance. The results from the immersion children revealed a similar, but more extreme, NWR/SR asymmetry compared to previous work (de Almeida et al., 2017; Thordardottir & Brandeker, 2013). The immersion children, who are exposed to French only at school, were able to score within the normal range on the LITMUS-NWR-French at about the same rate as bilinguals in France.[5] However, none of them were able to score in the normal range on LITMUS-SR-French, according to the 60% correct identical repetition cutoff used with bilinguals in France. These results confirm that the LITMUS-NWR-French is quite neutral to language exposure, whereas the LITMUS-SR-French is quite sensitive to variation in language exposure when identical repetition scores are used. The low identical repetition results in SR from the children in immersion could mean that a lower cutoff or a different scoring schema would need to be used for children with limited L2 exposure, such as those children in minority L2 acquisition contexts. It could also be the case that similar results would be observed in children in majority L2 contexts who do not frequently use the L2 outside of school. De Almeida et al. (2017)'s significant correlation between identical repetition SR scores and their Language Dominance Index (which includes measures of daily language use) suggests that weaker SR performance can be expected with majority L2 children who do not use French frequently on a daily basis. While more work needs to be done on SR with different bilingual populations and with different language combinations, it is clear that an identical repetition analysis of the LITMUS-SR-French results presents bias against children with limited exposure to French, even if the task may be effective in bilinguals who have considerable exposure quantity.

Phonological strengths and weaknesses as measured by the LITMUS-NWR-French

The second research question was addressed using a more qualitative analysis of the phonological and morphosyntactic properties targeted by each repetition task. The qualitative NWR analyses revealed the same phonological patterns as those reported for bilinguals in France (i.e., lower scores on language-independent items, most likely due to the number of syllables and clusters in these items; de Almeida et al., 2017; Ferré et al., 2015; dos Santos & Ferré, 2018). Although these prior studies focused more on intergroup differences (TD bilinguals vs. bilinguals with language impairment), their overall findings indicated that some variability in the TD data was due to number of syllables, clusters, and [l] in coda position. In other words, bilinguals with more exposure to French appear to have the same phonological performance as the children in immersion. Thus, some variability with respect to these phonological properties should be expected in TD children in this age range and with these language combinations, regardless of quantity of exposure to French.

Assuming that these repetition tasks are sensitive to language knowledge and do not simply reflect memory capacity (Gallon et al., 2007; Klem et al., 2015; Polišenská,

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

Chiat, & Roy, 2015), the NWR/SR asymmetry lends support to de Almeida et al.'s (2017) conclusion that phonology develops faster than morphosyntax. However, the stronger NWR performance could be due to the fact that the properties targeted by the LITMUS-NWR-French task (i.e., CCV clusters, sC clusters, and [l] in coda position) are also properties that are found in English. It could instead be the case that English-speaking children in French immersion do well on this task because of this overlap and not because phonology develops more quickly than morphosyntax. Yet, de Almeida et al. (2017) included children whose first languages were Arabic, European Portuguese, Turkish, and English, and they did not find any significant effect of L1 on NWR scores. Turkish arguably does not allow [l] in coda position (Ferré et al., 2015), and a large percentage of the Turkish–French bilinguals in de Almeida et al. (2017) were dominant in Turkish, so one might predict L1 influence on these children's NWR performance, but that was not what was observed. Thus, the available evidence to date suggests that this NWR task is neutral to exposure to French, no matter how typologically similar or distant the L1 phonology is from that of French. However, more research needs to be done on children who speak languages that do not allow the same cluster types, such as Spanish, which does not allow word initial sC clusters, or Japanese, which lacks tautosyllabic CC clusters.

Morphosyntactic strengths and weaknesses as measured by the LITMUS-SR-French

While the identical repetition scoring schema for SR presented a general bias against bilinguals with limited L2 exposure, due in part to difficulties with gender agreement, the target structure scoring schema revealed differences in morphosyntactic performance by property type. Specifically, the children in immersion had difficulty repeating monoclausal structures that targeted verb morphology, especially the 3PL forms, in addition to difficulty with the structure of complement clauses, an issue that will be addressed further below. The difficulty with these properties suggests that they are particularly sensitive to variation in language exposure. Regarding tense morphology, Paradis et al. (2011) suggested that TD bilinguals in majority L2 English contexts take about 3–5 years of exposure to English in order to become accurate with past tense *–ed* and 3SG present tense marking. However, it is unclear how long it might take children in a minority L2 immersion environment to master such forms. Regarding the 3PL forms, in a study on English-speaking children enrolled in early total French immersion programs in Ontario, Harley (1992) showed that even children in Grade 10 immersion produced 3PL markers in French in only 30% of obligatory contexts in a structured interview setting, suggesting that learners in early total immersion may not become accurate with 3PL verb morphology in French, even after continuing with French immersion into the high school years. Furthermore, in a longitudinal study on sequential bilingual children acquiring L2 French (with L1 English) in francophone primary schools in the Montreal area, Paradis, Le Corre, and Genesee (1998) reported that 3PL forms emerged later than past tense forms. Thus, it appears that French 3PL forms are acquired later by bilingual learners in different types of acquisition contexts. The difficulty with 3PL verb morphology is not surprising given that forms that are overtly marked for 3PL agreement are much less frequent in the input.

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

This difference in relative frequencies is used to explain the later acquisition of the 3PL forms in both monolinguals and bilinguals (see Prévost, 2009, for a review). Moreover, recall that sentences that differ with respect to the 10 subproperties targeted by the task (e.g., 3SG past and 3PL past) are very similar in length (see [9]–[10] and Appendix A), so rote memorization abilities are unlikely to explain the difficulty with the 3PL forms. While the exact mechanisms that underlie SR performance are not well known, there is increasing consensus that the ability to correctly repeat a sentence depends on language knowledge and not on rote memorization skills (Klem et al., 2015; Polišenská et al., 2015).

(9) Le lapin a mangé la carotte.
The rabbit has eaten the carrot.

(10) Les parents ont rangé les jouets.
The parents have put.away the toys.

In sharp contrast to difficulties with gender marking and verb morphology, children in immersion had significantly stronger performance on the target structure of relative clauses and *wh*-questions. This suggests that a target structure score of SR items involving *wh*-movement, even of those structures that are particularly complex (i.e., object relative clauses and *which* object questions), poses much less difficulty to children with limited language exposure, and could be a promising solution for reducing bias in SR tasks. Additional evidence for such a solution is found in prior research showing that structures with *wh*-movement pose problems for monolingual children with language impairment (Contemori & Garraffa, 2010; Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2004, 2011; Jakubowicz, 2011; Stavrakaki, 2001; van der Lely & Battell, 2003), so it is likely that focusing on these properties would allow for more accurate identification of language impairment in bilingual children. More research on using *wh*-movement to distinguish between bilinguals with and without language impairment is therefore warranted.

The relatively stronger results with *wh*-movement compared to gender marking and verb morphology in the immersion data appear to support previous work in L2 acquisition showing that syntactic properties are more easily acquired than morphological ones (e.g., Lardiere, 1998, 2008; Slabakova, 2013). However, the children in immersion had difficulty repeating the biclausal structure of the complement clause items, which relies on knowledge of syntactic embedding. As prior research on spontaneous language samples has shown that biclausal structures in general, and complement clauses in particular, emerge early and are produced frequently by TD sequential bilinguals (Paradis, Rusk, Duncan, & Govindarajan, 2017; Scheidnes & Tuller, 2018), the fact that children in immersion struggled to repeat this structure is surprising. This result is furthermore unexpected from a frequency point of view, but also from a complexity point of view, as well as a L1 typology point of view. In terms of the influence of relative frequency in the input, in a study based on corpus analyses of English child-directed speech, Diessel (2004) showed that complement clauses are more frequent than relative clauses, which he used to explain the emergence of complement clauses before that of relative clauses in L1 English. Moreover, Paradis et al. (2017) made a similar frequency-based argument

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

for the earlier mastery of complement clauses compared to relative clauses in sequential bilinguals (L2 English with various L1s). Assuming that the same relative frequency holds in a school context and for L2 French, then the fact that complement clauses led to lower target structure scores than relative clauses is surprising.

In terms of complexity, relative clauses, especially object relative clauses, are thought to involve greater computational complexity than complement clauses (e.g., Tuller, Henry, Sizaret, & Barthez, 2012). Prior work suggests that sequential bilingual children will avoid complexity in the L2 when possible, even when the equivalent structure in the L1 requires the same operation (Prévost, Strik, & Tuller, 2014). Therefore, even though word order in subject and object relative clauses is the same in French and English, we may expect children in immersion, who have low proficiency in French, and thus presumably weaker processing capacity, to display difficulty with such a structure. For example, we may expect these children to display avoidance strategies, such as producing a subject instead of an object relative clause. However, this is not what was observed.

In terms of L1 typology, we might expect acquisition to be facilitated by similarity between the L1 and L2 for some property. French and English share similar word orders for complement and relative clauses. Thus, if L1 and L2 similarity does reduce the exposure quantity required for the acquisition of a certain property, then one would expect similar performance for both complement and relative clauses, which is not what was observed.

One possible explanation is that complement clauses are acquired later by these children because more time is needed to acquire the subcategorization properties of verbs in the L2 (i.e., which verbs can take complement clauses). Thus, complement clauses may be more sensitive to language exposure because they involve more lexical learning than relative clauses. Another possible explanation for the weaker complement clause scores concerns non-native input. While the exact influence of non-native speaker input is unclear, prior work has shown that exposure to native input can positively impact L2 development (e.g., Jia & Aaronson, 2003; Paradis, 2011). However, it is unclear how exposure to only non-native input (from an L1 English speaker) would lead to the problems with complement clauses that were observed here (e.g., replacing the matrix verb with an auxiliary).

Finally, what can explain the relative ease with which children in immersion repeated relative clauses? Here, sheer age may be the crucial factor. Research on interveners in A-bar dependencies suggests that the ability to parse object relative clauses improves with age in typical (monolingual) development (Friedmann et al., 2009). If this is the case, then we could expect to see TD bilingual children perform at age-appropriate levels for object relative clauses, even if language exposure varies. If this is the case, age at time of testing should be more important for lexically restricted object relative clauses and object *which*-questions than exposure to the form in the input. If age counts more than exposure for the repetition of these structures, then this type of task and structure could be very promising for correctly identifying bilinguals with language impairment. It could be that processing certain A-bar dependencies in bilingual children is more closely related to mature performance systems than it is to crosslinguistic influence. Thus, more clinical research into this question is warranted, but it should include children with typologically different language combinations in order to better understand the link

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

between exposure and L1 typology in the acquisition of object relative clauses and *which*-questions.

Conclusion and limitations

In conclusion, children in immersion, who have limited daily exposure to French, have NWR performance that overlaps with that of bilinguals in France, thus confirming that a carefully designed NWR task can be neutral to language exposure, though this may depend on the phonological properties of the child's two languages. However, there was almost no overlap in SR performance between children in immersion and bilinguals in France, despite some typological similarities between French and English morphosyntax, thus confirming that SR is particularly sensitive to variation in language exposure. Yet, the relatively stronger scores on relative clauses and *wh*-questions suggest that focusing on (a target analysis) of these structures may lead to less biased SR tasks.

One limitation of the present study is that quality and quantity of input are not analyzed independently. The immersion children receive limited exposure in terms of quantity due to the fact that French is not spoken to them outside of school, but their exposure is also limited in quality as there is no regular contact with native French speakers. It could be that the weak SR performance of the children in French immersion, especially with respect to complement clauses, depends more on the lack of native speakers and not on the sheer quantity of exposure to French. However, it is clear that the children in immersion had stronger performance on NWR and *wh*-questions and relative clauses, in spite of the lack of native speakers. Thus, the results of the current study suggest that limited exposure in terms of both quantity and quality does not lead to lower NWR scores or to problems with *wh*-movement. Future work could endeavor to better examine these two factors independently, to the extent to which this is possible.

Acknowledgments. I would like to sincerely thank the parents, teachers, and children for their participation. I would also like to thank the researchers at the Université de Tours for sharing their tasks with me and for their helpful discussions: Alice Fleckstein, Philippe Prévost, Laurie Tuller, Sandrine Ferré, Christophe dos Santos, Rasha Zebib, Eléonore Morin, and Laetitia de Almeida. A big thanks to Yvan Rose and Greg Hedlund for their help with Phon and to Sharon Unsworth for sharing her tools for calculating cumulative language exposure. I am grateful to Alice Brun-Newhook, Camille Hellec, Laura Barr, Rebecca Hobbs, Kaylyn Turner, Sarah Fradsham, and Trevor Bessette for their help with data collection. This research was supported by the Social Sciences and Humanities Research Council of Canada (Grant #430-2016-00035) and by Memorial University of Newfoundland.

Notes

1. NWR = non-word repetition, SR = sentence repetition, PL = plural, SG = singular, Comp-NF = non-finite complement clause, Comp-FIN = finite complement clause, Q-who = object question with *who*, Q-which = object question with *which* + noun, Rel-SUBJ = subject relative clause, and Rel-OBJ = object relative clause.

2. Cumulative exposure was calculated using the Utrecht Bilingual Language Exposure Calculator (Unsworth, 2011). This calculation is a proportion of the total number of hours of daily language exposure out of the total number of hours that the child is estimated to be awake (i.e., 5300, or 14.5 hr of wakefulness per day). The school year in this district includes 190 days or 38 weeks. The children in this study participated in half-day kindergarten, which involved 2.5 hr of school per day for 5 days per week for 38 weeks,

which is the equivalent of 475 hr ([2.5 × 5] × 38) and represents .09 of a year (475/5300 = .09). For first grade, the children are at school for 30 hours per week ([30 × 38] = 1140 / 5300 hr = .22), so after completing first grade, a child in this context (i.e., with no exposure to French outside of school) would have a cumulative exposure of about a third of a year (.09 + .22 = .31 = 3.72 months). However, the children in this study were tested before the end of the school year, so only the part of first grade that the child had completed was counted in the calculation. It is possible that this calculation overestimates their exposure quantity, because periods for recess and lunch were not taken into account, and these are times during which the children most likely only speak English to each other. However, the point still remains that these children have overall limited exposure to French.

3. A very recent study on the LITMUS-SR-French proposed a 53% identical repetition cutoff for typical performance in bilinguals (Tuller et al., 2018). Even with this lower cutoff, only 2/33 (6%) children in immersion would score in the normal range.

4. Following dos Santos and Ferré (2018), these statistical analyses were run only on CCV clusters, but similar results were obtained when sC clusters were included.

5. Preliminary comparative analyses of the immersion children and children from France on the longer version of the *LITMUS-NWR-FR* revealed a similar overlap (Morry & Scheidnes, in press).

References

Adiv, E. (1980). *An analysis of second language performance in two types of immersion programs* (Unpublished doctoral dissertation, McGill University).

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

Chevrie-Muller, C., & Plaza, M. (2001). *Nouvelles epreuves pour l'examen de langage*. Paris: Les Editions du Centre de Psychologie Appliquée.

Contemori, C., & Garraffa, M. (2010). Comparison of modalities in SLI syntax: A study on the comprehension and production of non-canonical sentences. *Lingua*, **120**, 1940–1955. doi: 10.1016/j.lingua.2010.02.011

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for Specific Language Impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, **42**, 741–748. doi: 10.1017/S0021963001007600

Crawley, M. J. (2007). *The R book*. Chichester, UK: Wiley.

de Almeida, L., Ferré, S., Morin, E., Prévost, P., dos Santos, C., Tuller, L., … Barthez, M.-A. (2017). Identification of bilingual children with Specific Language Impairment in France. *Linguistic Approaches to Bilingualism*, **7**, 331–358. doi: 10.1075/lab.15019.alm

Deevy, P., & Leonard, L. B. (2004). The comprehension of *wh*-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **47**, 802–815.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge, NY: Cambridge University Press.

dos Santos, C., & Ferré, S. (2018). A nonword repetition task to assess bilingual children's phonology. *Language Acquisition*, **25**, 58–71. doi: 10.1080/10489223.2016.1243692

Ferré, S., dos Santos, C., & de Almeida, L. (2015). Potential phonological markers for SLI in bilingual children. In E. Grillo & K. Jepson (Eds.), *Proceedings of the Annual Boston University Conference on Language Development* (Vol. 1, pp. 152–164). Somerville, MA: Cascadilla Press.

Fleckstein, A., Prévost, P., Tuller, L., Sizaret, E., & Zebib, R. (2018). How to identify SLI in bilingual children: A study on sentence repetition in French. *Language Acquisition*, **25**, 85–101. doi: 10.1080/10489223.2016.1192635

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, **119**, 67–88.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, **31**, 661–681. doi: 10.1017/S0305000904006269

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of *wh* questions in three subtypes of SLI. *Lingua*, **121**, 367–382. doi: 10.1016/j.lingua.2010.10.004

Gallon, N., Harris, J., & van der Lely, H. (2007). Non-word repetition: An investigation of phonological complexity in children with grammatical SLI. *Clinical Linguistics & Phonetics*, **21**, 435–455. doi: 10.1080/02699200701299982

Genesee, F. (2004). What do we know about bilingual education for majority-language students? In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 547–576). Malden, MA: Blackwell.

Grüter, T. (2005). Comprehension and production of French object clitics by child second language learners and children with specific language impairment. *Applied Psycholinguistics*, **26**, 363–391.

Harley, B. (1979). French gender "rules" in the speech of English-dominant, French-dominant and monolingual French-speaking children. *Working Papers on Bilingualism, 19*. Retrieved from =http://eric.ed.gov/?id=ED184336

Harley, B. (1992). Patterns of second language development in French immersion. *Journal of French Language Studies*, **2**, 159–183. doi: 10.1017/S0959269500001289

Hedlund, G., & Rose, Y. (2019). *Phon 3.0* [Computer program]. Retrieved from https://www.phon.ca

Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, **121**, 339–351. doi: 10.1016/j.lingua.2010.10.006

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, **24**, 131–161. doi: 10.1017/S0142716403000079

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S.-A. H., Gustafsson, J.-E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, **18**, 146–154. doi: 10.1111/desc.12202

Lardiere, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research*, **14**, 359–375. doi: 10.1191/026765898672500216

Lardiere, D. (2008). Feature-assembly in second language acquisition. In J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). London: Erlbaum.

Maddieson, I., Flavier, S., Marsico, E., & Pellegrino, F. (2011). LAPSyD: Lyon–Albuquerque Phonological Systems Databases (Version 1.0). Retrieved from http://www.lapsyd.ddl.ish-lyon.cnrs.fr/

Mair, P., Schönbrodt, F., & Wilcox, R. (2017). WRS2: Wilcox robust estimation and testing. Retrieved from https://cran.r-project.org/web/packages/WRS2/index.html

Marinis, T., Armon-Lotem, S., & Pontikas, G. (2017). Language impairment in bilingual children. *Linguistic Approaches to Bilingualism*, **7**, 265–276.

Morry, J., & Scheidnes, M. (in press). The effect of linguistic experience on a non-word repetition task: Testing children with little exposure to their second language. In C. dos Santos & L. de Almeida (Eds.), *Bilingualism and specific language impairment: Selected proceedings of Bi-SLI 2015*. Amsterdam: Benjamins.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, **1**, 213–237. doi: 10.1075/lab.1.3.01par

Paradis, J., Le Corre, M. L., & Genesee, F. (1998). The emergence of tense and agreement in child L2 French. *Second Language Research*, **14**, 227–256. doi: 10.1191/026765898675500501

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

Paradis, J., Rusk, B., Duncan, T. S., & Govindarajan, K. (2017). Children's second language acquisition of English complex syntax: The role of age, input, and cognitive factors. *Annual Review of Applied Linguistics*, **37**, 148–167. doi: 10.1017/S0267190517000022

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language & Communication Disorders*, **50**, 106–118. doi: 10.1111/1460-6984.12126

Prévost, P. (2009). *The acquisition of French: The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition*. Amsterdam: Benjamins.

Prévost, P., Strik, N., & Tuller, L. (2014). *Wh*-questions in child L2 French: Derivational complexity and its interactions with L1 properties, length of exposure, age of exposure, and the input. *Second Language Research*, **30**, 225–250. doi: 10.1177/0267658313519814

Prévost, P., Tuller, L., & Zebib, R. (2012). *LITMUS-SR-French*. Tours, France: François Rabelais University.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven's Coloured Matrices*. London: H. K. Lewis.

R Core Team. (2017). R: A Language and Environment for Statistical Computing. Retrieved from http://www.R-project.org

Scheidnes, M., & Tuller, L. (2018). Using clausal embedding to identify language impairment in sequential bilinguals. *Bilingualism: Language and Cognition*. Advance online publication. doi: 10.1017/S1366728918000949

Semel, E., Wiig, E., & Secord, W. (2013). *Clinical Evaluation of Language Fundamentals (CELF-5)* (5th ed.). Bloomington, MN: NCS Pearson.

Slabakova, R. (2013). What is easy and what is hard to acquire in a second language: A generative perspective. In M. del P. García Mayo, M. J. Gutierrez Mangado, & M. Martínez Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 5–28). Amsterdam: Benjamins.

Statistics Canada. (2017). *St. John's, CY [Census subdivision], Newfoundland and Labrador and Newfoundland and Labrador [Province] (table). Census Profile*. (2016 Census. Statistics Canada Catalogue no. 98-316-X2016001). Retrieved from https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=1001519&Geo2=PR&Code2=10&SearchText=St.+John%27s&SearchType=Begins&SearchPR=01&B1=Language&TABID=1&type=1

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language*, **77**, 419–431. doi: 10.1006/brln.2000.2412

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

Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prevost, P., ... **Zebib, R.** (2018). Identifying language impairment in bilingual children in France and in Germany. *International Journal of Language & Communication Disorders*, **53**, 888–904. doi: 10.1111/1460-6984.12397

Tuller, L., Henry, C., Sizaret, E., & Barthez, M.-A. (2012). Specific language impairment at adolescence: Avoiding complexity. *Applied Psycholinguistics*, **33**, 161–184. doi: 10.1017/S0142716411000312

Unsworth, S. (2011). *Utrecht Bilingual Language Exposure Calculator (UBiLEC)*. Unpublished manuscript, Utrecht University, Utrecht.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, **16**, 86–110. doi: 10.1017/S1366728912000284

van der Lely, H. K. J., & Battell, J. (2003). Wh-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language*, **79**, 153–181. doi: 10.1353/lan.2003.0089

Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing* (2nd ed.). Burlington, MA: Elsevier.

Appendix A

Test sentences, property, subproperty, and length in syllables from the LITMUS-SR-French in order of presentation

Order	Sentence	Property/ Subproperty	Length in Syllables
1	Le garçon prend un bain.	Present-3SG	6
	The boy takes a bath.		
2	Le lapin a mangé la carotte.	Past-3SG	9
	The rabbit has eaten the carrot.		
3	Le papa sait très bien conduire la voiture.	Complement-non-finite	11
	The dad knows very well drive.INF the car.		
	"The dad knows how to drive the car very well."		

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

4	Qui le monsieur regarde?	*Wh*-question-*who*	6
	who the man watches		
	"Who is the man watching?"		
5	J'ai vu le chat qui a griffé la vache.	Relative-subject	10
	I have seen the cat who has scratched the cow.		
6	Les parents ont rangé les jouets.	Past-3PL	8
	The parents have put.away the toys.		
7	Je vois le garçon que la fille a poussé.	Relative-object	11
	I see the boy who the girl has pushed.		
8	Quel enfant la maîtresse punit?	*Wh*-question-*which*	8
	which child the teacher punishes		
	"Which child is the teacher punishing?"		
9	Les chats boivent du lait.	Present-3PL	5
	The cats drink.3PL some milk.		
10	Le garçon dit que la maman a lu un livre.	Complement-finite	12
	The boy says that the mom has read a book.		
11	Le singe a pris la banane.	Past-3SG	9
	The monkey has taken the banana.		
12	Tu vois le garçon qui a dessiné la mamie.	Relative-subject	13
	You see the boy who has drawn the grandma.		
13	La maîtresse punit les enfants.	Present-3SG	8
	The teacher punishes the children.		
14	La dame dit que le garçon a pris le ballon.	Complement-finite	12
	The lady says that the boy has taken the ball.		
15	Quel garçon le papy connaît?	*Wh*-question-*which*	8
	which boy the grandpa know		
	"Which boy does the grandpa know?"		
16	Tu as vu le cheval que le chien a mordu.	Relative-object	12
	You have seen the horse that the dog has bitten.		
17	Les tortues ont mangé la salade.	Past-3PL	9
	The turtles have eaten the lettuce.		
18	La maman sait très bien dessiner des lapins.	Complement-non-finite	12
	the mom knows very well draw.INF some rabbits		
	"The mom knows how to draw rabbits very well."		
19	Les enfants prennent un bain.	Present-3PL	6
	The children take. 3PL a bath.		

https://doi.org/10.1017/S0142716419000420 Published online by Cambridge University Press

20	Qui la maîtresse punit ?	*Wh*-question-*who*	6
	who the teacher punishes		
	"Who is the teacher punishing?"		
21	Le lapin veut manger la salade maintenant.	Complement-non-finite	12
	the bunny wants eat.INF the lettuce now		
	"The bunny wants to eat the lettuce now."		
22	La maman a fermé la fenêtre.	Past-3SG	9
	The mom has closed the window.		
23	Tu as vu la vache que le chat a griffée.	Relative-object	11
	You have seen the cow that the cat has scratched.		
24	Qui la mamie connaît?	*Wh*-question-*who*	6
	who the grandma knows		
	"Who does the grandma know?"		
25	Les parents punissent les enfants.	Present-3PL	8
	The parents punish.3PL the children.		
26	J'ai vu le chien qui a mordu le cheval.	Relative-subject	11
	I have seen the dog who has bitten the horse.		
27	Les enfants ont fermé la porte.	Past-3PL	8
	The children have closed the door.		
28	La maman lit une histoire.	Present-3SG	7
	The mom reads a story.		
29	Quel garcon le monsieur dessine ?	Wh-question-*which*	8
	which boy the man draw		
	"Which boy is the man drawing?"		
30	La fille croit que le papi a cassé un verre.	Complement-non-finite	12
	The girl believes that the grandpa has broken a glass.		

Cite this article: Scheidnes, M. (2020). Sentence repetition and non-word repetition in early total French immersion. *Applied Psycholinguistics* **41**, 107–131. https://doi.org/10.1017/S0142716419000420