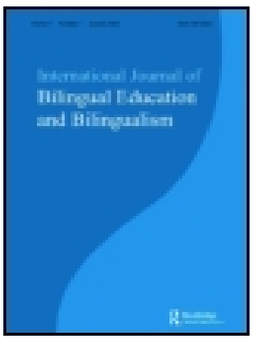

International Journal of Bilingual Education and Bilingualism

ISSN: 1367-0050 (Print) 1747-7522 (Online) Journal homepage: http://www.tandfonline.com/loi/rbeb20

Are background variables good predictors of need for L2 assistance in school? Effects of age, L1, amount, and timing of exposure on Icelandic language and nonword repetition scores

Elin Thordardottir

To cite this article: Elin Thordardottir (2017): Are background variables good predictors of need for L2 assistance in school? Effects of age, L1, amount, and timing of exposure on Icelandic language and nonword repetition scores, International Journal of Bilingual Education and Bilingualism, DOI: 10.1080/13670050.2017.1358695

To link to this article: http://dx.doi.org/10.1080/13670050.2017.1358695

Published online: 21 Aug 2017.

INTERNATIONAL JOURNAL OF BILINGUAL EDUCATION AND BILINGUALISM, 2017
https://doi.org/10.1080/13670050.2017.1358695

Are background variables good predictors of need for L2 assistance in school? Effects of age, L1, amount, and timing of exposure on Icelandic language and nonword repetition scores

Elin Thordardottir[a,b]

[a]School of Communication Sciences and Disorders, McGill University, Montreal, Canada; [b]ReykjavíkurAkademían, Reykjavik, Iceland

ABSTRACT

In response to the recent sharp increase of L2 students in Reykjavik schools, allocation criteria for special L2 services were adopted that were based on length of residence and on whether children's home language was tonal or not tonal. This study set out to evaluate the appropriateness of these criteria, and to replicate previous findings of a smaller scale study of the Icelandic and nonword repetition performance of L2 learners of Icelandic.

Participants: Included L2 learners and native speakers of Icelandic (n = 266) at three grade levels (grades 1–3, 5–6 and 8–9 (n = 266); the L2 learners included children from tonal and non tonal home languages.

Method: All the children were administered a new test of Icelandic vocabulary and grammar developed expressly for Icelandic, a test of Icelandic nonword repetition, and a background information questionnaire.

Results: L2 speakers in each age group performed significantly lower than L1 speakers in Icelandic vocabulary and grammar; less than a third of the L2 speakers performed within the normal L1 range, and over half performed more than 2 *SD* below this range. Low performers were particularly numerous in the oldest age group. NWR performance was related to age and Icelandic exposure, but scores were nevertheless uniformly high. No differences were found between children from tonal and non tonal home languages. The relationship between input and performance was complex, making fair allocation criteria based on background variables hard to formulate. Input variables (amount and timing of Icelandic exposure) were strongly related to input for the L2 group as a whole, and for the two older groups. However, the relationship was not significant for the youngest group. The pattern suggested that fast progress in Icelandic is related to higher age and recency of onset of L2 exposure.

ARTICLE HISTORY
Received 11 November 2016
Accepted 17 July 2017

KEYWORDS
Second language; L2; bilingual; vocabulary; grammar; nonword repetition; language processing; age of acquisition (AoA); length of residence (LoR); Icelandic

Introduction

Iceland is a country that has a centuries' long history of being linguistically homogeneous and predominantly monolingual in the sense that the overwhelming majority of its inhabitants had Icelandic as their native language and relatively few grew up with bilingualism in their homes. Traditionally, most Icelanders have also had a level of proficiency in foreign languages: both Danish and English are part of the compulsory curriculum and many students take additional language courses in upper grades, such as French, German or Spanish. In contrast, until recently

CONTACT Elin Thordardottir elin.thordardottir@mcgill.ca

very few people learned Icelandic as a second language (L2). Only recently, this language environment underwent a sudden and drastic change with a sharp rise in immigration within the last two decades. According to the 2011 census, 10.5% of the Icelandic population was born outside of Iceland (Statistics Iceland 2016) – with the addition of children of immigrants born in Iceland, as many as 15% of the nation may currently speak Icelandic as an L2. The Reykjavik School Board, by far the largest in the country, responded to the sudden and dramatic increase in students with a home language other than Icelandic, by providing special support services to these students. Allocation criteria for eligibility for these services needed to be formulated in the absence of specific evidence for this particular context. The initial criteria were based on two factors: length of residence (LoR) in Iceland and home language background: depending on whether children's home language was not tonal or tonal, they were allocated two or three years of services. The first criterion was based on the assumption that weak Icelandic skills were to be expected initially during an initial period of adaptation to a new language and environment, with subsequent improvement such that children could eventually be expected to function in the Icelandic school environment. The second criterion was based purely on teacher observation that children from tonal language homes encountered more difficulty in the school environment. One purpose of this study was to evaluate the appropriateness of these allocation criteria. A broader goal was to better understand the relationship of background factors such as amount of L2 exposure and the age at which this exposure starts on L2 proficiency.

Few studies have been published to date on L2 speakers of Icelandic. In a pioneering longitudinal study on the Icelandic L2 proficiency of school-age children (Elin Thordardottir and Juliusdottir 2013) 39 school-age children ranging in age from 5;10 to 17;7 (years; months) at the beginning of the study, were followed over a 3 year period. At the first test time, 89% of the L2 speakers scored more than 1 *SD* below and 73.7% more than 2 *SD* below the mean of native speaker norms in vocabulary and grammar. On subsequent testings over 3 years, few of the L2 speakers moved closer to the native mean in spite of making progress both in vocabulary and grammar. In contrast to these low language scores, the L2 speakers scored close to ceiling on Icelandic tests of nonword repetition at all three test times. These findings are further supported by a study by Ólafsdóttir et al. (2016), who reported that school-age L2 speakers of Icelandic lagged significantly behind native peers in vocabulary and reading, placing them at risk for academic difficulty.

The results of the longitudinal study by Elin Thordardottir and Juliusdottir (2013) offer important insights; however, the study has several limitations: The number of participants was fairly low and the participants varied widely in age. Further, the study employed tests of Icelandic that were originally constructed in English. The present study addressed these limitations by including a larger sample of participants targeting three age groups, the inclusion of a control group of native speakers, and the use of a language test constructed expressly for Icelandic.

Expected levels of L2 performance given length of exposure

Basing allocation criteria for special language services on factors such as length of residence rests on the assumption that input factors strongly predict L2 language performance and that at some point identifiable from background characteristics, children typically no longer need such services. Early reports of L2 learning in a school context estimated that 3–7 years were required to achieve L2 mastery, with less time required for conversational skills than academic language (Collier 1989; Genesee et al. 2005). These estimates come from studies that generally dealt with children learning English as their L2, and they varied in whether language skill was assessed directly or indirectly through measures of school performance. A number of large scale studies continue to confirm that L2 learners, as a group, lag behind native speakers in L2 proficiency for a long time, and some of them possibly permanently. However, it is not easy to determine the size of the difference in performance between L1 and L2 speakers from these studies or for exactly how long it should be expected to last – the type of information required for allocation criteria based on such factors.

Findings of lower performance in the language of school have been reported for L2 learners of English in the US of Hispanic background (Hammer, Lawrence, and Miccio 2008; Jackson, Schatschneider, and Leacox 2014; Oller, Pearson, and Cobo-Lewis 2007) and of Asian language backgrounds (Hammer, Jia, and Uchikoshi 2011; Pham and Kohnert 2014), and for English learners in Canada (Bialystok et al. 2010; Smithson, Paradis, and Nicoladis 2014). A number of recent studies focusing on more varied language groups have added importantly to this literature including Albanian L1 speakers learning Greek (Simos et al. 2014), Turkish L1 speakers learning Norwegian (Rydland, Grøver, and Lawrence 2014) and Dutch (Unsworth 2016), children of various backgrounds learning Icelandic, (Elin Thordardottir and Juliusdottir 2013; Ólafsdóttir et al. 2016), and French (Elin Thordardottir forthcoming). These studies surveying different language combinations learned in different countries are remarkably consistent in finding significant and persistent gaps between L2 and native speakers; however, the studies do not typically report the size of the difference directly. By inspecting the results, a difference of around 1 *SD* seems to be a common finding, with some variability. One study involving L2 learners from migrant families that were disadvantaged in terms of certain SES and living conditions reported gaps as large as 2 *SD* (Jackson, Schatschneider, and Leacox 2014). In contrast, a few studies have reported better performance (Bialystok et al. 2010; Smithson, Paradis, and Nicoladis 2014; Unsworth 2016). However, even when higher performance is reported relative to norms, a significant difference may still be reported relative to the control group (e.g. Bialystok et al. 2010). To conclude, studies on the language performance of L2 speakers consistently report them to perform significantly less well than native peers in that language, and they consistently report that these differences are maintained over time. These findings do not, therefore, suggest that most school age L2 speakers outgrow the need for special language services. The findings also reveal considerable variability that could be related to factors beyond amount of input, such as SES and other social and environmental factors.

The effect of age of acquisition

Children who are schooled in an L2 vary in whether their first exposure to the L2 occurs at the onset of school, earlier or later. Relatively few studies have examined the effects of the age of first exposure (AoE) in child L2 learners and available findings are contradictory. Early studies suggested that children who started L2 in middle childhood had better long term outcomes than children starting L2 exposure later (Collier 1989). However, other studies have also documented better L2 performance by adolescent L2 learners, at least in an initial period of L2 learning (Muircheartaigh and Hickey 2008). Older children may approach the task in a different manner and have a stronger base in their L1 on which to build. The effect of AoE has proven difficult to separate from the effect of amount of exposure, as a match on one of these variables often results in a mismatch on the other. For example, in a large study of Spanish L1 children acquiring English, Bedore et al. (2016) showed that an earlier AoE to English was associated with better English performance, however, possibly at a cost to Spanish performance. In contrast, Blom and Bosma (2016) reported that *later* AoE was associated with better L2 vocabulary performance and better grammatical performance, although to a lesser degree. The findings of these two studies are contradictory – importantly, both report as a limitation that the effect of amount of exposure was not controlled. Only a few studies have controlled both AoE and amount of exposure. Elin Thordardottir (2011a) showed that French-English simultaneous bilingual children's vocabulary in both languages at age 5 was unaffected by whether their bilingual exposure started before 6 months or age or after 21 month of age, when the groups were equated on cumulated exposure to the two languages. Unsworth (2016) found no differences on vocabulary or grammatical measures of school-age children whose bilingual exposure had started at age 1–3 years versus 4–7 years. Elin Thordardottir (forthcoming) found comparable performance in the French vocabulary performance of simultaneous and sequential learners of French enrolled in grades 1 and 3 in French schools. To summarize, many studies of AoE have not effectively controlled for the effects of amount of exposure; the few studies that have

controlled both factors indicate that amount of exposure is the more important determiner of L2 performance, for both preschool and young to middle school-age children.

L2 performance across domains of language

Among preschool children, the amount of bilingual exposure has been shown to exert a strong influence not only on vocabulary but also on grammatical development (Elin Thordardottir 2015). Studies on L2 acquisition in the school years have focused primarily on vocabulary size, therefore, clear findings on grammatical development by school-age L2 learners are relatively scarce. In a large study of Welsh and English learners spanning ages 3 years to older adults, significant differences in grammatical attainment between groups of people with different language exposure patterns were found in the primary school and teen age years for Welsh, such that children from Welsh only homes performed better than children from English-only homes and than children from Welsh-English homes. English grammar performance, in contrast, was equivalent across home language groups (Mueller Gathercole, Kennedy, and Thomas 2016). The previous longitudinal study of Icelandic (Elin Thordardottir and Juliusdottir 2013) reported significant gains in raw scores of grammatical subtests across test times. Younger and older participants were tested using different tests. Among the younger school-age children, gains were seen on subtests of grammatical morphology and syntactic knowledge, but not in grammatical comprehension. Among the older school-age children, gains were seen on subtests of syntax and detection of malapropisms. A study of the spontaneous language of the children in that study over two consecutive test times revealed overall high scores in the production of grammatical morphology, which is quite complex in Icelandic. Accuracy of verb morphology increased between the two test times from 85 to 90% among this group of children varying widely in age, whereas accuracy of case marking remained stable at 85% (Elin Thordardottir and Eiríksdottir 2012). A more recent study of another group of L2 learners of Icelandic teenagers similarly reported very high accuracy scores in the production of complex sentences and inflectional morphology in spontaneous language (Nowenstein and Elin Thordardottir forthcoming). In spite of high scores, significant differences were found between L2 and native speakers in the accuracy of grammatical morphology, but not in syntactic complexity. In contrast, Paradis, Tulpan, and Arppe (2016) reported a plateau in the productive verb morphological accurary of Chinese L2 speakers of English occurring at age 6 years.

Tonal versus non-tonal L1 background

The performance of children coming from tonal versus non-tonal home language backgrounds was included in this study for the sole reason that different service allocation criteria had been adopted for these groups by the Reykjavik school board based on teacher observation. Information is not available on the prevalence of this observation among teachers or on the specific types of difficulties they reported among children from tonal language backgrounds. In tonal languages, pitch patterns (such as pitch height, pitch direction, amplitude and duration) are used to distinguish meaning (lexical or grammatical) among words that would otherwise be homophonous. Tonal languages are reported to make up 60–70% of the world's languages (Maddiesen 2013). Icelandic is not a tonal language. A literature search yielded a number of articles that have examined the acquisition of tone by L1 speakers of non-tonal languages as well as by speakers of other tonal languages (Hallé, Chang, and Best 2004; Hao 2012; Li, Shao, and Bao 2017; Wang and Saffran 2014). These studies vary in whether they found speakers of non-tonal languages to have greater difficulty perceiving and producing the target tonal patterns (Hallé, Chang, and Best 2004; Li, Shao, and Bao 2017) or not (Hao 2012; Steien and Dommelen 2016). Differences in performance perceiving or producing the tones of a particular language were also seen among speakers of other tonal languages (Li, Shao, and Bao 2017). Indeed, tonal patterns, their categorization, and the cues that are used to distinguish them differ across tonal languages (Brunelle 2009; Li, Shao, and Bao 2017; Maddiesen 2013). Our

literature search did not yield any studies on the reverse situation, comparing the performance of speakers of tonal and non-tonal languages on L2 performance in a non-tonal language. This situation might be reasoned to be easier, because it does not involve having to learn contrasts that have no parallel in the native language. However, it could be speculated that if speakers of tonal languages encounter difficulty going in this direction, it might be because they incorrectly associate pitch or prosodic variations with meaning. Icelandic has a stress pattern of alternating stressed and unstressed syllables, with the first syllable of each word always being stressed. Inflectional morphemes are always unstressed. In Icelandic, however, stress does not distinguish meaning. Importantly, in the studies that were found on the acquisition of tone, the participants were adults, and their task was to distinguish or produce tonal contrasts. The studies did not examine whether their ability to do so influenced the acquisition of vocabulary or grammar in the tonal language. Those studies on the acquisition of English as L2 that have focused on native speakers of Asian languages, including Chinese and Vietnamese, do not stand out as reporting lower English performance than other studies (Hammer, Jia, and Uchikoshi 2011; Pham and Kohnert 2014). Interestingly, in the cross-linguistic study of nonword repetition, children with language impairment are uniformly found to perform significantly more poorly than peers with typical development across a wide range of languages (see next section), with the notable exception of Cantonese (Stokes et al. 2006). It is not clear whether this is related to Cantonese being a tonal language. Other possibilities include that multisyllabic words used in norword repetition tasks do not occur in Cantonese.

Nonword repetition performance

The interest in nonword repetition (NWR) in this study stems from its potential use to help rule out language impairment in bilingual children, due to its sensitivity to language impairment (LI) but relative insensitivity to varying abilities of bilingual children in the language of the nonwords (see Elin Thordardottir and Brandeker 2013). Children with and without LI have been shown to vary significantly in NWR performance across a number of languages, including Icelandic (Bortolini et al. 2006; Conti-Ramsden, Botting, and Faragher 2001; Girbau and Schwartz 2008; Elin Thordardottir 2008; Elin Thordardottir et al. 2011). In a landmark study of bilingual children with and without LI, Elin Thordardottir and Brandeker (2013) showed that in children with typical development age 5 years, NWR was unaffected by varying levels of bilingual exposure, in contrast to vocabulary scores. Further, it was shown that NWR correctly separated groups of children with and without LI regardless of whether they were bilingual. Subsequently, a number of studies have reported high NWR scores by bilingual children with typical development using NWR in various languages, including Dutch, Icelandic and Korean (Boerma et al. 2015; Brandeker and Elin Thordardottir 2015; Lee, Kim, and Yim 2013; Elin Thordardottir and Juliusdottir 2013). At the same time, a number of studies, in particular targeting Hispanic children in the US have reported less clear cut results for NWR in bilingual children (see e.g. Gibson et al. 2015). Differences between studies may depend on a number of factors, including age and the construction of the nonwords. For example Elin Thordardottir and Brandeker found no relationship between previous amount of exposure and NWR for their French NWR test, but a significant, albeit weak relationship for their English nonwords, which were of more complex construction (see also Duncan and Paradis 2016; Evans and Coady 2008). Boerma et al. reported higher diagnostic accuracy for the detection of LI in bilingual children using a NWR designed to be quasi-universal, following Chiat and Polisenska (2016), than for a Dutch test. The quasi-universal test included a limited number of vowels and consonants that have a high likelihood of occurring across many languages. The French test used earlier by Elin Thordardottir & Brandeker, while not designed to be universal, had a similar item construction as that used by Boerma et al., involving only CV syllables with no clusters. However, the NWR test yielding high scores by L2 speakers of Icelandic (Elin Thordardottir and Juliusdottir 2013) included more complex items including both word-like and non word-like words, as well as words with consonant clusters. Thus, high scores by bilingual children have been reported both using NWR tests specifically simplified to fit

bilingual children (Boerma et al. 2015) and using tests that have not been specifically adapted for bilingual children (Elin Thordardottir and Brandeker 2013; Elin Thordardottir and Juliusdottir 2013; Lee, Kim, and Yim 2013). The ability of NWR tests to rule out language impairment (specificity) hinges on the ability of children without LI to repeat the words correctly – in the present context, in spite of low proficiency in the language of the nonwords. The present study sought to confirm the previous finding of high scores on the Icelandic NWR test by L2 speakers with varying Icelandic abilities in a larger sample of children using the same NWR test, previously documented by Elin Thordardottir and Juliusdottir (2013).

A test constructed for Icelandic

Previous findings on the L2 learning of Icelandic have relied on tests that were adapted from English, such as the PPVT (Dunn and Dunn 1994) and the TOLD (Hamill and Newcomer 1997; Símonardóttir and Guðmundsson 1996; Símonardóttir et al. 1995). However, given that these were originally constructed in English, their contents are based on an English sequence of acquisition and on language structures found to be vulnerable in LI in English. In Elin Thordardottir and Juliusdottir (2013), although raw score gains were generally observed between successive test times, some subtests evidenced ups and downs over time. One possible reason for these irregularities over time could be that the tests do not adequately capture the development of the Icelandic language. As an example, the PPVT has been used clinically in an unnormed Icelandic translation for many years in Icelandic clinics serving young children. However, the translation of the PPVT into Icelandic reveals itself as particularly problematic for the more advanced items required for older school-age children and adolescents, where English frequently employs words of Latin (French origin), which are associated with more academic language. Icelandic does not have such a dichotomy of translation equivalents from different origins and use in more casual versus academic contexts, and, therefore, cannot maintain the relative level of difficulty of items of the PPVT. As an example, the English item 'homunculus' has no Icelandic translation other than 'little man'- and is thus an advanced vocabulary item in English but not in Icelandic. Another reason for scores not uniformly increasing with age in Icelandic could be that L2 learning does not proceed in the same manner (involving sequence) as L1 learning. The present study employed a new test constructed originally in Icelandic and that specifically targets language used in Icelandic schools. Vocabulary items, and grammatical structures included were based on the language of Icelandic school curricula at various grade levels. Such materials reflect the language requirements of students in Icelandic schools more closely than tests originally constructed in other languages.

Purpose of study

This study follows up on a previous longitudinal investigation of L2 learning of Icelandic by school-age children (Elin Thordardottir and Juliusdottir 2013). Although that study provided important insights, limitations included a fairly limited sample size of children varying largely in age and language background, and the use of language tests adapted from English. The present study set out to document the Icelandic proficiency of a larger group of school-age L2 learners at three different grade levels, with a systematic examination of the effect of the children's L1. Specific questions included the following: 1) How does the typical Icelandic performance of L2 speakers of Icelandic compare with that of L1 speakers at three grade levels, 2) Is there a siginficant difference in the L2 Icelandic performance of children from tonal and non tonal home language backgrounds?, 3) How does L2 Icelandic performance relate to long term and recent exposure to Icelandic as well as the age at which Icelandic exposure started? Based on previous results, it was predicted that L2 speakers would score lower than L1 counterparts as a group, and that individual variability would be related to previous exposure to Icelandic. Further, it was predicted that L2 speakers would perform similarly to

L1 speakers on Icelandic NWR. No firm predictions were made in terms of whether children from tonal and non-tonal home language backgrounds would differ in their Icelandic proficiency.

Methods

Participants

Participants were recruited through 14 public primary and middle schools (grunnskóli) in Reykjavik, with the collaboration of the Reykjavik school board (Skóla-og frístundasvið Reykjavíkurborgar). Parents of children who participated signed an informed consent form. The study was overseen by the Institutional Review Board of the Faculty of Medicine of McGill University and by the Data Protection Authority (Persónuvernd) of Iceland. Participants were 266 children, including children enrolled in grades 1–3, grades 5–6, and grades 8–9. Each grade level intially comprised three groups: native speakers of Icelandic (L1IS), L2 speakers of Icelandic whose L1 is a tonal language (L2TL), and L2 speakers of Icelandic whose L1 is not a tonal language (L2NTL). Beyond an initial comparison not finding significant differences in Icelandic performance between L2 speakers from tonal and non tonal L1s (see results section), the two groups of L2 speakers were collapsed together in subsequent analyses into one group of L2 speakers. Table 1 presents background characteristics for L1 and L2 speakers (including those from tonal and non tonal L1s), including the number of children in each group, age and socio-economic status (SES), indexed by total years of maternal education. Within each grade level, the groups of L1 and L2 speakers of Icelandic did not differ significantly in chronological age, but did differ significantly in SES, with lower maternal education in the L2 than L1 groups at all grade levels. Information on SES was available for only 100 children, equally distributed over the L1 and L2 groups. Many parents of both L1 and L2 speakers left this information out even though they filled out other parts of the background information form. The native (L1IS) speakers were required not to have lived in other countries; however, with the exception of the youngest children, they cannot be described as fully monolingual as foreign language instruction starts with English in grade 3 and Danish in grade 7.

Table 2 presents background information for the L2 speakers differentiated into L2TL (tonal L1) and L2NTL (non tonal L1) speakers. This information is relevant to the comparison made between the Icelandic language and NWR of these two groups. Tonal home languages in this study included Chinese, Vietnamese and Thai. Non-tonal home languages included Polish, Russian, Ukrainian, Slovenian, Tagalog and Cebuano. Detailed information on the language exposure history of the L2 children was collected by a parent questionnaire described in detail in the Procedures section. Table 2 reports the age of arrival (AoA), the age of first significant exposure to Icelandic (AoE), the length of residence in Iceland (LoR) and the length of significant exposure to Icelandic (LoE) of the L2 speakers. The distinction between AoA and AoE, and LoR and LoE, respectively, is made because some of the children were reported to have stayed at home without receiving significant Icelandic exposure for some months or years before attending Icelandic daycare, preschool or school. Table 2 also reports the children's % exposure to Icelandic since birth and over the last 4 years preceding the study. This exposure

Table 1. Background characteristic of participants, including native Icelandic speakers (L1IS) and L2 speakers of Icelandic.

	1–3 grade		5–6 grade		8–9 grade	
	L1IS	L2	L1IS	L2	L1IS	L2
n	42	62	37	54	25	46
Age in mos.	92.7	91.2	132.8	130.2	167.4	169.3
SD	(9.4)	(9.4)	(9.2)	(9.8)	(9.9)	(7.6)
p	.998		.980		.980	
Mat. Ed.	16.9	13.3	17.8	13.4	16.4	12.4
SD	(3.3)	(2.2)	(4.4)	(3.3)	(3.4)	(2.7)
t	4.665		3.416		2.299	
p	<.000		<.000		<.040	

Table 2. Language exposure background characteristics of groups of L2 speakers of Icelandic, by grade level and by home language background (tonal first language, L2TL and non-tonal first language, L2NTL). All measures are expressed in the mean and (SD) number of months.

	1–3 grade		5–6 grade		8–9 grade	
	L2TL	L2NTL	L2TL	L2NTL	L2TL	L2NTL
n	18	44	21	33	18	28
AoA	4.0 (8.7)	26.4 (21.1)*	13.8 (34.2)	60.6 (34.0)*	61.8 (57.5)*	100.1 (48.4)*
AoE	23.2 (16.01)*	36.0 (21.3)*	34.8 (28.0)*	64.1 (27.8)*	73.8 (47.3)*	101.1 (45.3)*
LOR	89.2 (17.9)	66.3 (19.9)*	116.5 (32.8)	76.9 (39.3)*	102.4 (61.8)*	74.1 (48.7)*
LOE	62.5 (5.6)*	56.5 (19.4)*	94.6 (26.1)*	72.5 (34.6)*	87.8 (45.5)*	73.1 (46.3)*
%Ice	33.3 (11.9)*	25.1 (10.0)*	30.8 (11.4)*	22.1 (11.8)*	21.6 (13.5)*	13.5 (10.1)*
%Ice4	42.5 (13.3)*	36.3 (10.3)*	40.3 (11.2)*	38.1 (11.9)*	35.8 (15.0)*	31.8 (14.1)*

*group scored differed significantly from same grade level group of native speakers of Icelandic (not displayed in the Table).
Underlined groups: L2TL and L2NTL groups at same grade level differed significantly from each other.
%Ice: percent of waking hours in Icelandic-speaking environments since birth.
%Ice4: percent of waking hours in Icelandic-speaking environments in last 4 years.

refers to the % of waking hours spent in Icelandic-speaking environments. Some of the children had trilingual exposure because of two languages being spoken in the home, particularly the speakers of Cebuano, many of whom also had some English exposure at home. Further, some L2 speakers of Icelandic had some level of Icelandic exposure in the home. If so, this was taken into account in the computation of their exposure to Icelandic. Data on the native speakers of Icelandic are not displayed in the Table, but were used in statistical comparisons between the groups. For the native speakers, AoA and AoE were set at 0 months, LOR and LOE are the same, and are both equivalent to chronological age; for this group of children, and the % exposure to Icelandic was set at 100%.

Table 2 indicates whether the language groups (L1IS, L2TL and L2NTL) within each grade level differed significantly from each other on the background variables. These comparisons were made by comparing all groups (language and grade level) in a one-way ANOVA for each variable, with post hoc Tukey tests used for pair-wise comparisons of groups.

Procedure

Children were tested individually at their school in a session lasting about an hour. Two tests were administered:

1) A newly developed test of Icelandic proficiency, Milli mála (Elin Thordardottir 2011b). This test has a total of 128 items divided into 4 subsections: receptive language (42 items), productive vocabulary (25 items), definitions (23 items) and grammaticality judgement (38 items). Milli mála was developed as a test of Icelandic proficiency primarily in L2 speakers in grades 1–10 (ages 6–16 years). Vocabulary and sentence structure items of varying difficulty levels were selected based on the language content of textbooks used at various grade levels in Icelandic schools in various subjects, including mathematics, geography, history and Icelandic studies. Several vocabulary items are also included that are considered unlikely to be learned in the school context, as a gauge of contact with Icelandic outside of school. In the first subsection of the the test (receptive language), children select among four pictures the one that best corresponds to a word, sentence or short paragraph read by the examiner. This subsection has an internal consistency (Cronbach's alpha) of .871 based on the data from the present study. In the second subsection, children are asked to name pictured items. Correct answers include object labels as well as more complex concepts such as providing a date and telling time in Icelandic. The internal consistency for this section was .903. In the third subsection, children are asked to define items of increasing complexity, including words referring to abstract concepts, phrases and idioms. In addition to being asked to provide their own definition (which many children are unable to do), children select one of three definitions that are provided visually and are read to the child as well (only the multiple choice response was used in this study).

This section has an internal consistency of .661. The last section is a grammaticality judgment in which the examiner reads sentences and the child is asked to indicate whether they are correct or not. The types of errors that are incorporated into this grammaticality judgment were based on language samples from previous studies on L2 speakers of Icelandic with and without language impairment (Elin Thordardottir 2008, 2016) and school-age L2 speakers of Icelandic (Elin Thordardottir and Eiríksdottir 2012) as well as on error types encountered in Icelandic newspapers and newscasts, which represent morphological variations that prove difficult even for native speakers. This section of the test has an internal consistency of .784. The Milli Mála test was used in this study in a preliminary version – since this study, it has been normed on a larger sample of L2 speakers (Elin Thordardottir forthcoming).

2) Two lists of nonword repetition (NWR), developed in a previous study (Elin Thordardottir 2008), including a set of word-like nonwords and a set of non word-like nonwords. Each set includes 25 items, ranging in length from 1 to 5 syllables (5 at each length). Both sets of nonwords respect Icelandic phonotactic rules. The word-like words have characteristic Icelandic intonation and real nominative word endings. Longer nonwords in that list are formed like real Icelandic compound words, with the first part having a genitive inflection and the second part a nominative ending. The non word-like items have uncharacteristic intonation patterns and word endings. Longer words are not formed as compound words. The NWR lists were scored on line using the method presented in Dollaghan and Campbell (1998), leading to a score of percent phonemes correctly repeated. Omissions and substitutions are scored as incorrect; additions are ignored.

Testing was conducted by trained examiners who were employees of the Reykjavik School Board. Parents were sent a background information form surveying the child's general development, parental education level (number of years of education) as a proxy for socio-economic status (SES), and the child's language exposure history. The language exposure history form was one used in a number of previous studies of bilingual development (Elin Thordardottir 2011a; Elin Thordardottir et al. 2006). Parents are asked factual questions detailing the language(s) spoken in the home, dates of daycare, preschool and school attendance, and language of use in these settings, other relevant settings in which the child has spent time on a regular basis. From this form, a single number is calculated representing the proportion of the child's waking hours since birth spent in Icelandic environments. The calculation includes contexts of potential interaction – thus counting contexts involving human contact, but not counting TV watching. Calculations can be made of Icelandic exposure over the child's entire life, or over smaller segments. In this study, we computed the percentage of waking hours spent in Icelandic environments over the child's life, over the last 4 years preceding testing and over the last 8 years preceding testing. These different measures were meant to explore the usefulness of exposure measures emphasizing the child's entire versus most recent experiences in predicting L2 performance. Basic information on schooling and language use in the home was also asked of the children when they were tested. For those children for whom the parent questionnaire was not returned, information obtained from the child was used to document exposure.

Results

Icelandic performance of children from tonal and non-tonal backgrounds

Independent sample *t*-tests were conducted at each grade level separately comparing the total Milli mála performance of the L2TL and L2NTL groups. No signficant differences between these groups were found: Grades 1–3 ($p = .789$), grades 5–6 ($p = .133$), and grades 8–9 ($p = .141$). Subsequently, the L2 speakers are combined in further analyses of performance on the Milli mála test in one single L2 group.

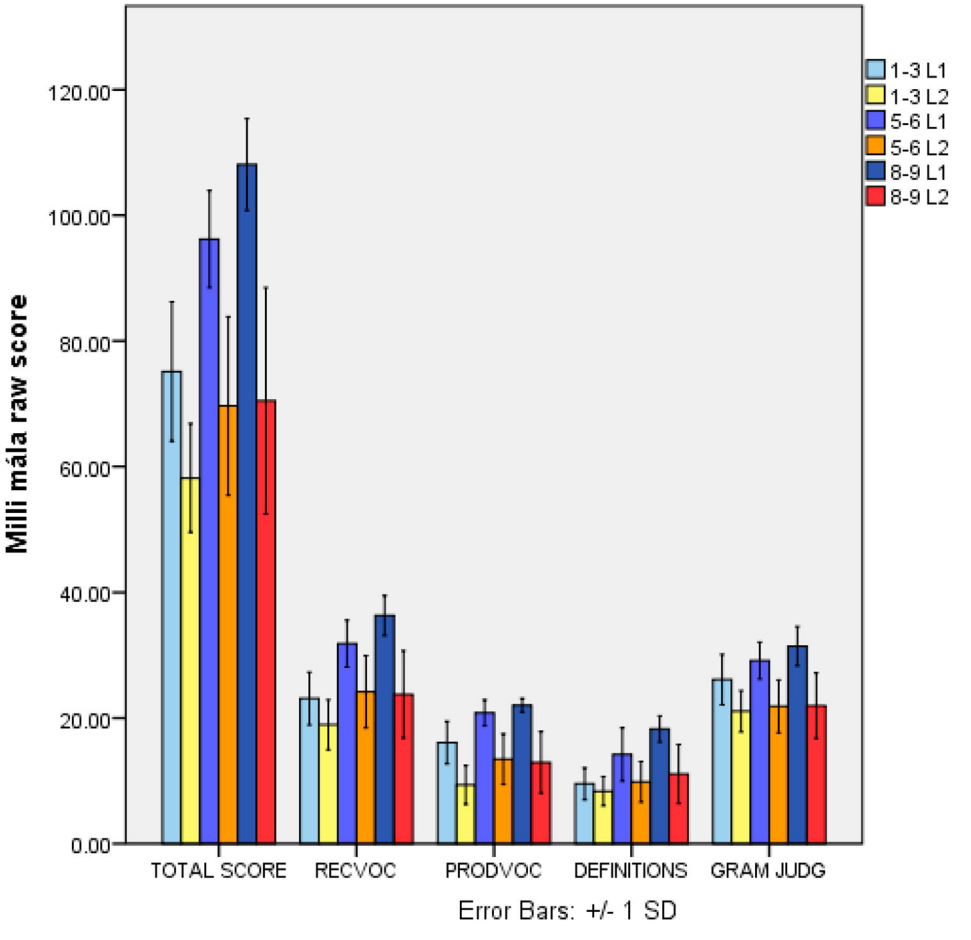

Figure 1. Milli Mála total score and subsection scores (Receptive Language, Productive Vocabulary, Definitions and Grammaticality Judgment) by language group (L1 and L2 speakers of Icelandic) by grade level (grades 1–3, grade 5–6, and grades 8–9).

Developmental sensitivity of the Milli mála test for L1 and L2 speakers

For the Milli mála test to be used to gauge advances in language level and how closely L2 speakers resemble L1 speakers of the same grade level, it is important to ascertain that the test is sensitive to the development of Icelandic in the school years and that it produces significant differences between age groups. Total scores on the Milli Mála Icelandic test as well as scores on each subsection are presented in Figure 1, displaying the results of the L1 and L2 language groups at each age (grade) level. To test for the effect of age on the total Milli mála score, a two-factor ANOVA was run, with Language group (L1, L2) and Age (youngest, middle and oldest) as the two factors. Results revealed a significant effect of Age ($F(2,247) = 74.916$, $p < .000$); a significant effect of Language ($F(1,247) = 297.569$, $p < .000$) and a significant Age x Language interaction ($F(2,247) = 13.992$, $p < .000$). This result was followed up by One-Way ANOVAs on the effect of Age in each Language group (L1 speakers: $F(2,96) = 109.31$, $p < .000$; L2 speakers: $F(2,151)$, $p < 000$) and Fischer LSD post hoc tests. The post hoc tests revealed that the source of the Age x Language interaction is that, among L1 speakers, all age groups differed from both other age groups ($p < .000$), whereas among L2 speakers, the youngest group differed from each of the older groups ($p < .000$), whereas the two older groups did not differ from each other ($p = .450$).

Performance across subsections of the Milli Mála test (different language domains)

ANOVA analyses were conducted to examine age effects for each section of the Milli Mála test for the L1 and L2 groups separately, revealing a similar pattern as on the total score of the test. For the L1 speakers, a significant effect of age group was found for each subsection of the test (receptive vocabulary ($F(2, 100) = 107.212$, $p < 000$), productive vocabulary ($F(2, 103) = 53.928$, $p < .000$), definitions ($F(2,101) = 60.860$, $p < .000$), and grammaticality judgment ($F(2,101) = 19.693$, $p < .000$). Post hoc Tukey HSD tests revealed that all age groups differed significantly from both other groups on all sections, except in productive vocabulary where the middle and oldest groups did not differ significantly from each other ($p = .142$).

For the L2 speakers, a significant effect of age group was found for the first three sections of the test: receptive vocabulary ($F(2,157) = 12.413$, $p < .000$, productive vocabulary $F(2,156) = 15.985$, $p < .000$, and definitions $F(2,156) = 8.379$, $p < .000$, but not for grammaticality judgment ($p = .627$). Post hoc Tukey HSD tests revealed, for receptive and productive vocabulary, a significant increase from the youngest to the middle group, but not from middle to oldest ($p = .979$). For definitions, only the youngest and oldest groups differed significantly from each other.

Percentage of L2 speakers scoring within normal limits compared to L1 peers

A main question in this study is the extent to which, and the time frame during which L2 speakers require special assistance to achieve Icelandic skills that will allow them to benefit from school work

Table 3. Number and percentage of L2 children in each age group scoring WNL in Icelandic proficiency relative to L1 peers, and background variables for each performance group. AoA and EoA are in months; hours of Icelandic is in thousands of hours.

	WNL (within − 1 SD of mean)	−1 to −2 SD of mean	−2SD from mean
Grades 1–3			
Number	17	20	12
Percentage of group	34.7	40.8	24.5
%Icelandic since birth	28.3 (10.2)	27.1 (12.9)	23.3 (12.0)
%Icelandic last 8 years	28.6 (10.1)	27.1 (12.9)	24.1 (11.3)
%Icelandic last 4 years	40.7 (11.5)	36.5 (12.0)	35.4 (13.1)
AoE	32.5 (17.8)	36.5 (12.0)	35.4 (13.1
LoE	62.8 (15.4)	57.5 (20.6)	52.9 (27.9)
Hours of Icelandic	8,006 (2,821)	7,508 (3,584)	6,318 (2,858)
Maternal ed.	13.8 (1.3)	11.6 (2.3)	14.0 (1.6)
Grades 5–6			
Number	8	5	34
Percentage of group	17.0	10.6	72.3
%Icelandic since birth	36.1 (14.8)	21.2 (3.8)	25.4 (11.3)
%Icelandic last 8 years	41.8 (17.3)	26.4 (1.6)	31.0 (13.0)
%Icelandic last 4 years	43.1 (14.6)	38.4 (7.02)	38.8 (11.8)
AoE	33.0 (28.5)	60.0 (28.1)	54.9 (31.2)
LoE	101.1 (23.5)	65.4 (20.4)	78.2 (34.4)
Hours of Icelandic	15,176 (6,583)	7,784 (918)	10,231 (5,162)
Maternal ed.	(12.5 (3.5))	(18.0 (0))	12.8 (2.56)
Grades 8–9			
Number	3	3	28
Percentage of group	8.8	8.8	82.3
%Icelandic since birth	38.0 (14.5)	30.0 (14.2)	14.4 (9.0)
%Icelandic last 8 years	51.7 (16.7)	43.7 (22.8)	21.2 (10.4)
%Icelandic last 4 years	51.7 (16.7)	44.3 (22.3)	31.1 (11.6)
AoE	40.0 (34.6)	56.3 (21.0)	94.0 (48.7)
LoE	118.5 (4.9)	102.5 (31.8)	74.7 (47.3)
Hours of Icelandic	20,701 (8,309)	15,332 (8,383)	7,137 (4,276)
Maternal ed.	–	(15.0 (0))	11.8 (2.6)

conducted in Icelandic. The L2 speakers' language scores relative to native speakers of the same grade level offer an indication of their standing in this respect. For this analysis, children scoring at or above −1 *SD* relative to L1 grade peers were considered to score Within Normal Limits (WNL) for their grade level. A middle group included children scoring between −1 and −2 *SD*, and the lowest scoring group included children scoring more than 2 *SD* below the mean. Table 3 displays the number of L2 children at each grade level falling into each of these proficiency groups as well as the language exposure variables of these children. Descriptive information is reported for each of the groups for several different measures of previous exposure to Icelandic, reflecting our interest in exploring which type of measure is most highly related to Icelandic performance given that previous exposure was being used as a criterion of service allocation. The measures include the percentage of waking hours spent in Icelandic-speaking environments since birth, over the last eight years and last four years, AoE in months, LoE in months, maternal and paternal education in months. A number of participants who filled out background questionnaires did not include information on their education, therefore data on this variable are missing for a number of children. In the table, the mean numbers in parentheses are based on very few data points. Given that the grade level groups differ in age, and that each age group covers a 2–3 year age range, the percentage of time spent in Icelandic does not correspond to the same amount of time across age groups. As a means of providing a measure that can be compared more directly across age levels, an estimate of the absolute number of hours spent in Icelandic environments was computed by multiplying the percentage of time since birth spent in Icelandic environment x the age in months x the number of waking hours per month spent with other people with the possibility of linguistic interchange (set at 300). Table 3 reveals that within each age group, only a minority of children scored WNL. The percentage of children doing so diminished with each higher age group. The percentage of children scoring more than 2 *SD* below the mean increased with age. From studies of preschool simultaneous bilinguals, it has been proposed that a %exposure of 40% or greater since birth allows bilingual children to score WNL in each language separately (Elin Thordardottir 2011a). A similar finding was reported for bilingual school-age children in grades 1 and 3 (Elin Thordardottir forthcoming). In Table 3, children in the two oldest groups who scored WNL did have close to or more than 40% exposure to Icelandic since birth, whereas children scoring lower in Icelandic did not reach this amount of exposure.

Relationship of Icelandic proficiency and background variables

Correlation between Icelandic performance and language exposure measures

Table 4 presents correlational analyses relating the total Milli Mála score to different language exposure variables to explore whether greater or lesser relationships with performance were obtained using measures of overall, lifetime exposure, or more recent exposure. Milli Mála scores were significantly correlated with age for both L1 and L2 speakers; however much more strongly so for the L1 speakers. Maternal education was not significantly correlated with Icelandic performance for either group. For the L2 group as a whole, significant correlations were found between

Table 4. Correlations between Icelandic performance (Milli mála total score) and background Icelandic input variables for L2 and L1 children by age group.

	L2 gr 1–3	L2 gr 5–6	L2 gr 8–9	All L2	All L1
% Ice birth	.115	.497**	.767**	.340**	
% Ice8 years	.125	.495**	.750**	.477**	
% Ice 4 years	.219	.239	.537**	.274**	
AoE	−.019	−.475**	−.613**	−.197*	
Hours Ice	.184	.516**	.764**	.586**	
Age	.329*	.180	.144	.391**	.705**
Maternal ed.	.042	.077	–	−.061	−.031

**correlation significant at the 0.01 level.
*correlation significant at the 0.05 level.

Icelandic proficiency and all the linguistic background variables: exposure from birth, exposure over 8 years, exposure over 4 years, AoE, number of hours of exposure and chronological age. Of the background variables, the one most strongly correlated with Icelandic performance was the number of hours spent in Icelandic environments since birth.

Language exposure variables and Milli mála performance across age groups

Figure 2 is a scatterplot for the entire group of participants of individual children's Milli mála scores plotted as a function of age (left panel) and as a function of the total number of hours of Icelandic exposure over the child's lifetime (right panel). In Figure 3, scatterplots are displayed of each age group of L2 speakers separately, showing Milli Mála scores as a function of the number of hours of Icelandic exposure. The number of hours was selected here based on the results of the correlational analysis. Descriptively, Figure 2 illustrates well the marked difference in distribution of scores between the L1 and L2 groups. A steep increase in scores with age is seen for the L1 group with fairly low variability. L2 learners evidence a less steep increase in scores as a function of age and a much wider scatter of scores, particularly among the oldest children. None of the L2 children score better than L1 age mates; L2 scores range from the mean of the L1 speakers to a low of 40/128 items correct throughout the age range and almost no overlap in seen between the scores of the L1 and L2 groups when matched on age (left panel). The right panel of Figure 2 reveals almost no overlap between the L1 and L2 groups in the absolute amount of previous Icelandic exposure. Yet, many of the L2 speakers obtain scores comparable to those of L1 speakers. This indicates a different relationship between amount of exposure and speed of learning in the two groups.

Multiple simultaneous regressions were run for the group as a whole, and for each grade level group, to predict Milli Mála score from chronological age, number of hours of Icelandic exposure, and AoE to Icelandic. Standardized Beta values were used to assess the relative contribution of each of the three variables. The regression results are presented in Table 5. For the group as a whole, and for grades 8–9 and grades 5–6, the three factors together significantly predicted Icelandic performance, but did not for children in grades 1–3. Whereas the three factors each contributed significantly for the group as a whole, analysis of the individual age groups showed that the hours of exposure contributed significantly only for the oldest group, age at time of testing contributed significantly only for the two youngest groups, and AoE contributed significantly only for the middle group.

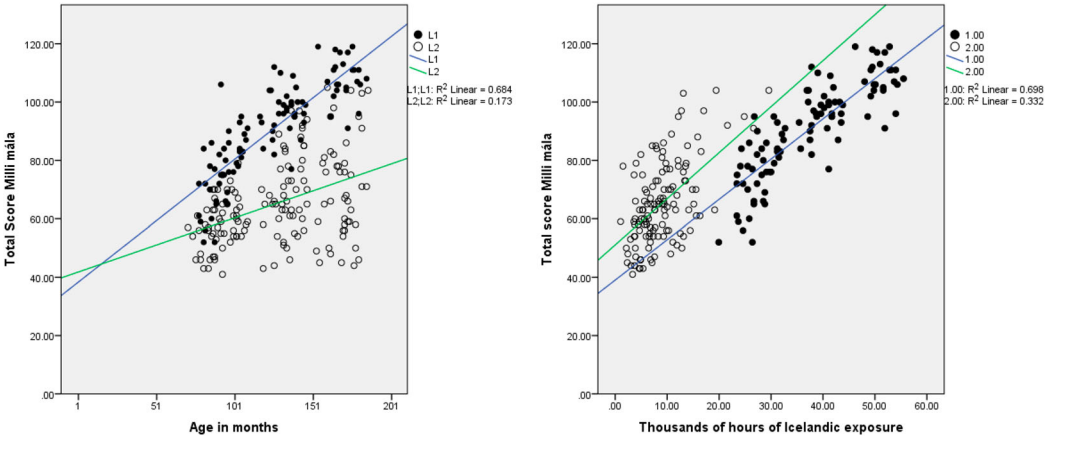

Figure 2. Icelandic proficiency (Total Milli mála score) for L1 speakers and L2 speakers as a function of chronological age (left panel) and as a function of the absolute number of hours of previous Icelandic exposure since birth (right panel).

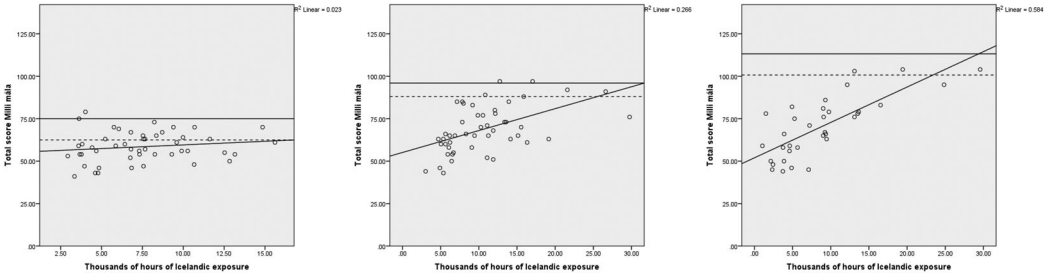

Figure 3. Icelandic proficiency (Total Milli mála score) as a function of number of hours spent in Icelandic-speaking environments since birth for children in grades 1–3 (left panel), grade 5–6 (middle panel) and grades 8–9 (right panel). The solid straight line in each graph represents the mean score of L1 speakers in the age group – the dotted line represents the −1 *SD* mark relative to the L1 mean.

These results suggest explanations for why L2 speakers in the two oldest groups vary in their Icelandic performance, but provide no explanation of the variability seen in the youngest group. In a further search for a language exposure variable that could contribute to the distribution of the youngest children into ability groups, the L2 speakers in grades 1–3 were divided into simultaneous and sequential learners of Icelandic based on an AoE before or after 36 months. This comparison is not meant to advocate for such a dichotomous classification of bilingual children, but is conducted because such a classification is traditionally used in clinical work and it is, therefore, of relevance to test its usefulness. This yielded 28 simultaneous and 20 sequential speakers of Icelandic. Thirty nine percent of the simultaneous speakers scored WNL, 43% between −1 and −2 *SD*, and 14% below −2 *SD*. Of the sequential speakers, 30% scored WNL, 40% between −1 and −2 *SD*, and 30% below −2 *SD*. A *t*-test comparing the simultaneous and sequential learners revealed no significant difference in Icelandic performance (t, 46 = .710, p = .481). Examination of the home languages also revealed no notable differences; the majority of children in both the highest and lowest ability groups were of Polish L1. Finally, examination of which school the children came from gave no indication that particular schools were associated with better or worse performance.

Nonword repetition (NWR)

Mean performance on each of the two NWR tests (word like and nonword like) by group (age and language) ranged from 90.9 (*SD* 5.3) to 97.4 (*SD* 2.3). Age and language group effects were examined by a factorial ANOVA with arcsine transformed scores, with Age Group (3 levels) and Language Group (3 levels) as the factors. For the word-like list, a significant main effect of Age was found (F(2,255) = 16.146, p < .000). There was no significant main effect of Language (p = .537) and no significant Age Group x Language Group interaction (p = .099). Post hoc Fisher LSD tests on the age effect revealed that the youngest children (grades 1–3) differed from each of the two older age groups (grades 5–6 and grades 8–9). For the non word-like list, the main effects of Age Group and Language Group were significant, whereas the interaction between these factors was not (p = .476): Age group: F(2, 256) =

Table 5. Regression analysis results predicting Milli Mála score from number of hours of exposure to Icelandic, chronological age, AoE to Icelandic for the entire group of L2 speakers and for the subgroups of L2 speakers in Grades 1–3, 5–6 and 8–9.

			Standardized Beta					
Grade Level	(*df*)*F*	*p*	Hours	*p*	Age	*p*	AoE	*p*
All	(3,226) 186.322	<.000	.438	<.000	.431	<.000	.347	<.000
1–3	(3,47) 2.584	.065	.043	.853	.396	.026	−.111	.654
5–6	(3.48) 8.882	<.000	.110	.570	.331	.025	−.395	.027
8–9	(3,33) 14.842	<.000	.658	<.001	−.011	.932	−.158	.348

12.239, $p < .000$; Language group: $F(92,256) = 5.398$, $p = .005$. Post hoc tests on the Language group factor revealed a significant difference between L1IS speakers and L2TL speakers, with no other significant difference. Post hoc tests on the Age factor were as for the other NWR list: the youngest group differed from each of the older groups.

Figure 4 shows the performance of L1 speakers and L2 speakers in each of three age groups on the two NWR tasks as a function of number of hours of Icelandic exposure (word-like items in left panel and non word-like items in right panel). Only the non word-like item list correlated significantly with percent exposure to Icelandic (nonwordlike: $r = .296$, $p < .01$, word-like $r = .137$, ns.). Both lists correlated significantly with number of hours of Icelandic at the 0.01 level (non word-like: $r = .296$, $p < .01$, word-like $r = .250$, $p < .01$), and both lists correlated with age at the 0.05 level (non word-like: $r = .171$, $p < .05$, word-like: $r = .186$, $p < .05$). For the L1 speakers, both lists correlated significantly with age at the 0.01 level (non word-like: $r = .501$, $p < .01$, word-like: $r = .307$, $p < .01$). As Figure 4 shows, however, in spite of these significant relationships between exposure variables, and significant age group differences, performance on both NWR lists was quite high and similar in range across levels of previous exposure. The horizontal line in the graphs shows the mean performance of native Icelandic speaking 9 year olds with LI on each of these two lists (Elin Thordardottir 2008). On the word-like list (left panel of Figure 4), 12/266 children (4.5%) of children scored below this line; on the non word-like list (right panel), 8/266 children (3%) scored below it. Children scoring below the line did not include more L2 than L1 speakers.

Discussion

The results of this study confirm previous findings indicating that school-age L2 learners of Icelandic, as a group, lag significantly and importantly behind native speakers in their mastery of the Icelandic language. The findings show that a large proportion of the L2 speakers, in particular in the oldest age groups, score more than 2 *SD*s below native peers. The findings indicate that large numbers of L2 speakers in Icelandic schools are in need of services to help them master Icelandic sufficiently well for their educational needs. A main goal of this study was to evaluate the appropriateness of service allocation criteria for L2 students used in Reykjavik schools, based on length of residence

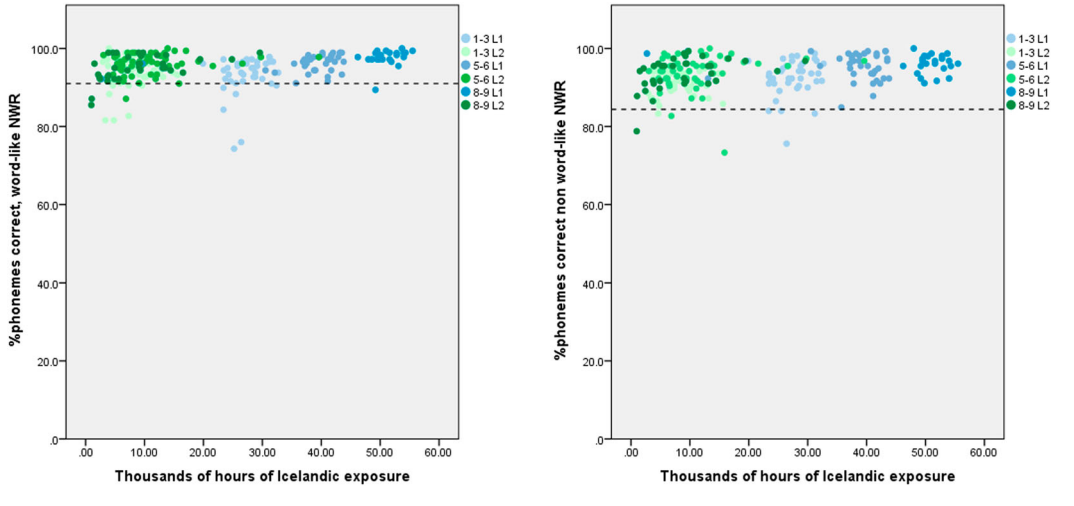

Figure 4. Percent correct scores on NWR for word-like nonwords (left panel) and non word-like nonwords (rigth panel) for L1 speakers and L2 speakers. Age groups are represented in progressively darker shades of blue for L1 speakers and green for L2 speakers. The horizontal lines indicate the mean NWR scores for each list obtained by native Icelandic 9-year olds (grade 3) with LI (Elin Thordardottir 2008).

in Iceland and home language background as tonal or non tonal. The findings do not support the use of these criteria. No difference was found in Icelandic performance between speakers from tonal versus non tonal home language background. In terms of length of residence, no specific length of time in terms of number of years was identified that was associated with WNL performance in Icelandic, although, for the two oldest groups, a percent exposure level of 40% or more was associated with WNL status, consistent with previous studies (Elin Thordardottir 2011a, forthcoming). Among a number of exposure measures explored, the one most highly correlated with Icelandic performance was the absolute amount of Icelandic exposure. However, the relationship between the amount and timing of exposure and Icelandic performance was found to be complex, and importantly, different in the different age groups. In contrast to their highly variable Icelandic language scores, the L2 speakers scored uniformly high on two Icelandic tests of NWR, indicating that the NWR lists have the potential of being used to rule out language impairment in this group. Each of these issues is discussed in more detail in the following sections.

Milli mála performance of native and L2 speakers of Icelandic

In documenting low oral language Icelandic performance of school-age L2 speakers of Icelandic, the study confirms those of a previous study (Elin Thordardottir and Juliusdottir 2013) with a larger sample of participants, and using a different test of Icelandic performance developed expressly for Icelandic. Vocabulary and grammatical performance on the Milli Mála test of the L1 speakers increased systematically with age with fairly low variability and with significant differences in performance between all three age groups. This, along with the generally high internal consistency of the test, indicates that the Milli Mála test is sensitive to Icelandic language development in the age range of 6–16 years. Neither the L1 or L2 group showed evidence of a floor or ceiling effect on the test as a whole (the oldest L1 speakers obtained maximum scores on the vocabulary sections, but not Definitions or Grammaticality Judgement). The four subtests of the Milli mála test address different domains of language. Native speakers evidenced a significant increase with age in each of those. L2 speakers evidenced lesser age effects on the whole, but evidenced a similar degree of age effect in vocabulary (receptive and expressive) and in definitions. However, their lack of age effect in grammaticality judgment differs from their performance on the other subtests and from the native speaker pattern. This might indicate that the L2 speakers encounter particular difficulty in the mastery of Icelandic inflectional morphology – the focus of this subtest – or that they experience a plateau effect similar to that reported by Paradis et al. (2016) for Chinese learners of English. The low performance in grammatical morphology is, however, in disagreement with the high performance of L2 speakers of Icelandic in spontaneous language (Elin Thordardottir and Eiríksdottir 2012; Nowenstein and Elin Thordardottir forthcoming). A possible reason for this discrepancy may be that the grammaticality judgment subtest contains a wide range of difficulty of inflections, including distinctions that prove difficult even to native speakers. This was done to avoid a ceiling effect for the L1 speakers. In spontaneous language, in contrast, speakers can deliberately avoid the use of inflections they do not master. This is an important topic that should be investigated further. The L2 speakers' apparent arrested development in grammaticality judgment could be due to a critical period effect of some kind. However, the fact that performance in grammatical morphology increased over time in the longitudinal study speaks against such an explanation (Elin Thordardottir and Eiríksdottir 2012). Alternatively, it might be possible to remedy this weakness by modifying educational practices. Importantly, the mastery of Icelandic for professional purposes does require advanced mastery of the inflectional system and should be given careful consideration in order to prepare students adequately for higher education and professional opportunities. Icelandic proficiency was not related to SES (years of maternal education) for either L1 or L2 speakers. This result must be interpreted with caution, however, because SES information was available for less than half the children.

Proportion of L2 speakers scoring WNL

The extent to which the performance of L2 speakers differs from that of L1 peers provides a way to gauge how well equipped the L2 speakers are to participate in school work in Icelandic and to use Icelandic for social purposes. Overall, two thirds of the L2 speakers scored one *SD* or more below the L1 mean for their grade level and over half scored 2 or more *SD* below the L1 mean. If the performance of the L2 speakers were distributed similarly to that of the L1 speakers, only about 16% of the children would be expected to score below −1 *SD*, and only 2% of the children would be expected to score below −2*SD*. Clearly, the two distributions are very different. Low L2 performers in relation to L1 speakers were particularly numerous in the older age groups, where the great majority of children scored more than 2 *SD* below L1 peers. However, there was also a much larger scatter of scores among L2 speakers such that the highest scorers in absolute terms among the L2 children were also in the two oldest groups (see Figures 2 and 3). Overall, studies of L2 speakers in various countries consistently show them to score significantly lower in the L2 than do L1 peers (see review in introduction). The size of the difference is often difficult to determine, however, and varies across studies. In the planning of educational services for L2 students, it is important to understand how rapidly and how well they can be expected to master the language used in school. The present findings show that individual children vary greatly in this respect. One way in which this study differs from many previous studies of L2 children is that it includes children of a wide age range and who also vary greatly in when they arrived in Iceland (as opposed to studies who have followed children who were schooled in the L2 from the beginning). This is a reality of the current school-age immigrant population in Iceland. It illustrates the difficulty in describing the needs of this population as a whole. Further, this illustrates the difficulty of conceiving of a bilingual norm for a population of such great heterogeneity.

L2 speakers from tonal versus non-tonal home languages

Service allocation criteria in Reykjavik schools distinguished between children from tonal and non tonal home languages based on teacher observation that these groups performed differently. This hypothesis was not borne out in this study. It is not impossible that the teacher observations may still be valid. It may be that differences between these language groups exist in some aspect of language that was not measured in this study. This could involve, for example, reading and spelling. Chinese uses a writing system of characters corresponding to whole words, whereas Thai and Vietnamese use alphabetic spelling systems. Other areas of potential difficulty not assessed here include pronunciation or pragmatic skills. Further study would be required to examine such possibilities. A significant difference did emerge in NWR between L1 speakers and L2TL speakers only, and only on non word-like items. This difference, if replicated, may potentially relate to greater phonological differences between Icelandic and tonal languages than Icelandic and non-tonal languages. It can be speculated that it could also be related to differences in the writing systems of tonal languages, that is, whether they involve phonological spelling as it has been suggested that children's strategy in completing NWR tests may be influenced by whether or not they have learned to read (see Elin Thordardottir and Juliusdottir 2013).

NWR performance

NWR has been suggested as a particularly important tool in identifying the presence or absence of language impairment in bilingual children (Elin Thordardottir and Brandeker 2013). Low performance on NWR is a strong indication of LI; however, several studies have shown NWR to be essentially unaffected by bilingualism (Boerma et al. 2015; Lee, Kim, and Yim 2013; Elin Thordardottir and Brandeker 2013; Elin Thordardottir and Juliusdottir 2013). However, there have also been contradictory findings (Gibson et al. 2015). Another issue is whether NWR tests differ in how much they are affected by

bilingual status, with the possibility that simpler items are necessary to avoid such effects (Chiat and Polisenska 2016). This study used two NWR lists that varied in wordlikeness, but were not specifically designed to include phonologically simple items. Results confirm the previous finding by Elin Thordardottir and Anna Gudrun Juliusdottir, showing very high scores by the L2 speakers, and also add to the previous findings by showing systematic age effects which could not be documented in the previous study because of a much smaller sample size. The effect of age was similar among both L1 and L2 speakers, as evidenced by no language group by age group interaction. This is an interesting finding in light of the fact that age effects on the test of Icelandic proficiency were not similar across L1 and L2 speakers, and thus speaks further to a dissociation between proficiency in Icelandic and the ability to repeat Icelandic non-words. NWR scores were significantly related to previous exposure to Icelandic. Nevertheless, the range of NWR scores was much smaller than the range of Milli Mála scores. Further, the NWR scores of L1 and L2 children overlapped greatly whereas Milli Mála scores hardly overlapped at all. Importantly in terms of the clinicial use of NWR tests for the purpose of ruling in or ruling out language impairment, all but a handful of the children scored in the range characteristic of native speakers with TD rather than with LI suggesting that the NWR can be of help in ruling out LI in this population of Icelandic speakers. NWR scores were uniformly above 90% for the word-like list and above 85% for the non word-like list, the levels of performance documented previously for native Icelandic speaking 9-year-olds with LI (Elin Thordardottir 2008). Only 12/266 (4.5%) children scored below this point on the word-like lists and 8/266 (3%) on the non word-like list. These results indicate that the NWR lists used in this study are promising as a clinical tool to help rule out LI in Icelandic bilingual school-age children. In a more general sense, the results indicate that bilingual children can score high, or significantly above the range of scores associated with LI even on NWR tests that use items that are fairly complex phonologically. Further study is required to examine the diagnostic precision of these tests with bilingual children with and without language impairment.

Effect of input on L2 vocabulary and grammar

The findings of this study offer several novel insights into the relationship between L2 exposure and L2 proficiency. Previous literature in general shows a relationship betwen exposure and proficiency (see detailed review in the introduction). Across a number of studies, children schooled in an L2 make progress over time, even if they are generally found not to close the gap with L1 age or grade mates. In terms of the relative effects of amount of L2 exposure and the age at which this exposure starts, findings have been contradictory (see Bedore et al. 2016; Blom and Bosma 2016). However, studies of school-age children that have carefully controlled both amount of exposure and AoE have found that amount of exposure plays a more important role (Elin Thordardottir forthcoming; Unsworth 2016). The present study offers new insights into the role of amount and timing of input in that it shows the relationship between these variables and L2 performance to be markedly different across age groups of school age children. Whereas amount of input exerted a strong influence on Icelandic performance in L2 speakers in grades 8–9, AoE was a more important influence among children in grades 5–6. Unexpectedly in light of studies generally showing exposure variables to impact L2 performance, no significant relationship was found between exposure and L2 performance among the youngest children, in grades 1–3.

In searching for explanations for this pattern of findings, it is pertinent to consider the fact that the present study sample differs from that of many previous studies of L2 performance in school-age in that the children span a large age range, from 6 to 16 years, and in that the sample includes children who vary greatly in AoE. Whereas school-age L2 studies typically include children with AoEs occurring at school entry or before (e.g. Elin Thordardottir forthcoming; Unsworth 2016), the present study includes a considerable number of children with AoEs throughout the school-age range. This is true in particular of children in the upper grades, as is evident from Table 3, and is reflective of

the current school-age immigrant population in Iceland. The different relationship between amount of exposure and Icelandic performance across the three age groups is depicted in Figure 3. Examination of background variables (Table 3) reveals that only very few of the L2 speakers in the oldest group were born in Iceland or were early arrivers; the mean AoA for this group is 94 months (7 years, 10 months). Figure 3 shows a steep increase in Icelandic performance as a function of number of hours of Icelandic exposure in grades 8–9, a somewhat less steep increase in grades 5–6, and a relative lack of increase in scores with increased hours in grades 1–3. Although this study is cross-sectional, these patterns do suggest that the oldest group is adding to their Icelandic proficiency with added exposure the fastest and the youngest group the slowest. The explanation for these different patterns can only be speculated on at this point. Older children may learn an L2 more rapidly than younger children for a number of reasons, including enhanced cognitive skills, greater use of metalinguistic abilities and deliberate strategies (see Genesee et al. 2005), and possibly, greater personal motivation. More efficient L2 learning by older children has been found in several previous studies (e.g. Muircheartaigh and Hickey 2008). Another characteristic of the oldest age group in this study (see Table 4) is a relatively higher AoE. This suggests that the rapid Icelandic learning in relation to hours of exposure may be linked to the recency of first Icelandic exposure. Previous studies have reported enhanced rates of learning at the beginning of L2 exposure in young school-age children (Elin Thordardottir forthcoming; Hammer, Lawrence, and Miccio 2008; submitted). The present study suggests that recency of exposure may produce a period of enhanced learning also in children with higher AoEs.

When input factors do not explain performance differences

Perhaps the most puzzling finding of this study is that none of the background exposure variables were correlated with Icelandic performance for L2 children in grades 1–3. A first possible explanation might be that the children in this age group are too similar in their backgrounds or in their Icelandic performance scores for any effect to appear. This may seem plausible in light of the fact that most of these children had attended Icelandic preschools, which are run by the city and are highly standardized in their quality and teaching approach. However, there is considerable variability in the children's backgrounds. They range in AoE from 0 to 72 months (given that they did not all start preschool at the earliest possible age), and in number of hours of exposure from 2,937 to 15,606. On the Milli Mála test, they range in scores from 41 to 79 points out of a maximum of 128. Table 3 reveals that the highest and lowest performers (WNL and – 2*SD*) differ only slightly on the background variables: for AoE, the group means differ by merely 4 months (from 32 to 35 months), the % of waking hours spent in Icelandic speaking environments differs only by 5%; the LoR a bit more, or by 10 months. Importantly, both groups range fairly widely in each of these background variables.

In keeping with the hypothesis that recency of first exposure plays a role in setting the rate of L2 learning, could it be that the pattern in this youngest group indicates that these children have reached a stable state where new L2 learning has slowed down significantly and is therefore no longer related to amount of input as strongly as before? Such slowing could similarly explain the finding of Ólafsdóttir et al. (2016) that children with early AoEs learn Icelandic more slowly than children with higher AoEs. It can be speculated that the youngest children in the present study may have had an earlier period in which they learned more quickly. But if so, why the slowing down? It is likely relatively easier to acquire basic language skills than more advanced language (Genesee et al. 2005). An initial boost in L2 learning could also be related to a strong need to be able to function in the new language. Therefore, the slowing might not be related only to the absolute complexity level achieved, but also to whether a level has been achieved that permits basic functioning and thus does not create as urgent a need for progress. The younger children require a lower level in L2 than the older children to get by. In a study of vocabulary and grammatical development of simultaneous bilingual preschool children, relative amount of exposure received over the child's lifetime in each language exerted a

strong influence on the level of proficiency in that language, in both vocabulary and grammatical development (Elin Thordardottir 2011, 2015). However, more exposure produced rapid gains mainly at low exposure levels up to roughly 50%. This was also the exposure level that permitted the bilingual children to score within the normal range relative to monolinguals. Alternative explanations can be speculated on as well. One factor that should be carefully considered in future resarch is the children's level of attainment and maintenance of the L1. Late arrivers have the advantage of having likely developed their L1 to a high degree of proficiency, whereas this may be more questionable for the L2 speakers of Icelandic with lower AoEs who acquire their L1 as a heritage language.

The large number of children in the study as a whole with low Icelandic peformance scores is certainly cause for worry. However, this result may be most surprising for the youngest group and raises compelling questions regarding why language learning sometimes does not happen at the rate expected based on the language environment that is provided to the children (see also Hu and Schuele 2015). Further study is required to understand what factors other than exposure influence L2 learning. These results show early exposure to an L2 certainly does not guarantee good L2 outcomes, and may in some cases be counterproductive (see also Blom and Bosma 2016; Elin Thordardottir forthcoming; Ólafsdóttir et al. 2016). Other variables that were documented in this study also failed to provide an explanation for the range of scores obtained by the youngest children, including SES, whether the children were simultaneous or sequential learners, what home language group they belonged to or which school they were attending. The literature suggests that additional important factors may include the quality of input, including the type of language activities occurring the children's homes (Scheele, Leseman, and Mayo 2010), the primary caregivers' vocabulary knowledge (Buac, Gross, and Kaushanskaya 2014), as well as patterns of strengths and weaknesses in language related skills that may characterize over- and underachievers (Hu and Schuele 2015).

Conclusions

The findings of this study indicate that a very significant proportion of school-age learners of Icelandic as L2 achieve lower levels of mastery of Icelandic than are required for academic and social purposes. These results underscore the fact that learning L2 in the school context does take considerable time and is not associated with automatic success. The poor performance of a large proportion of the L2 speakers relative to L1 peers does not qualify as language impairment, given that it reflects only part of the overall language knowledge – also, the uniformly high NWR scores suggest typical language learning abilities. This low level of mastery of the language of school and the surrounding community is, nevertheless, bound to present these children with considerable difficulty both in academic and social contexts and is likely to impact their future access to educational and employment opportunities, with significant impacts on their quality of life.

The results of this study provide novel insights into the complex relationship between exposure patterns and L2 performance. Importantly, background exposure patterns exerted a different level of impact in different age groups, and had no relationship to performance for the youngest children As a result, allocation criteria for special services based on length of residence or home language background are hard to formulate for the group as a whole. Based on the findings of this study, the Reykjavik school board changed its allocation criteria, which now include individual assessment of Icelandic proficiency. Further research to refine those criteria is in progress.

Acknowledgments

The author wishes to acknowledge the collaboration of the School Board of the City of Reykjavik (Skóla-og frístundasvið, Reykjavíkurborg) in the recruitment and data collection for this study. Thanks are also extended to school personnel who assisted with recruitment, and last but not least, to the participating students and their parents.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Elin Thordardottir is professor in the School of Communication Sciences and Disorders of McGill University, Montreal, and a researcher at ReykjavikurAkademian in Iceland. She holds a Ph.D. in Communicative Disorders and clinical certification in Speech-Language Pathology and Audiology. Her current research focus is on typical language development and language disorders (assessment and intervention) in monolingual and bilingual children. Her work has appeared in various journals such as *The Journal of Speech, Language and Hearing Research, Applied Psycholinguistics, The International Journal of Bilingualism, The Journal of Communication Disorders, The International Journal of Speech Language Pathology*, and *The Canadian Journal of Speech Language Pathology and Audiology*.

References

Bedore, L., E. Peña, Z. Griffin, and G. Hixon. 2016. "Effects of Age of English Exposure, Current Input/Output and Grade on Bilingual Language Performance." *Journal of Child Language* 43: 687–706.

Bialystok, E., G. Luk, K. Peets, and S. Yang. 2010. "Receptive Vocabulary Differences in Monolingual and Bilingual Children." *Bilingualism: Language and Cognition* 13: 525–531.

Blom, E., and E. Bosma. 2016. "The Sooner the Better? An Investigation Into the Role of Age of Onset and Its Relation With Transfer and Exposure in Bilingual Frisian–Dutch Children." *Journal of Child Language* 43: 581–607.

Boerma, T., S. Chiat, P. Leseman, M. Timmermeister, F. Wijnen, and E. Blom. 2015. "A Quasi-Universal Nonword Repetition Task as a Diagnostic Tool for Bilingual Children Learning Dutch as a Second Language." *Journal of Speech Language and Hearing Research* 58: 1747–1760.

Bortolini, Y., B. Arfé, C. Caselli, L. Degasperi, P. Deevy, and L. Leonard. 2006. "Clinical Markers for Specific Language Impairment in Italian: The Contribution of Clitics and Nonword Repetition." *International Journal of Language and Communication Disorders* 41: 695–712.

Brandeker, M., and Elin Thordardottir. 2015. "Language Exposure in Bilingual Toddlers: Performance on Nonword Repetition and Lexical Tasks." *American Journal of Speech-Language Pathology* 24: 126–138.

Brunelle, M. 2009. "Tone Perception in Northern and Southern Vietnamese." *Journal of Phonetics* 37: 79–96.

Buac, M., M. Gross, and M. Kaushanskaya. 2014. "The Role of Primary Caregiver Vocabulary Knowledge in the Development of Bilingual Children's Vocabulary Skills." *Journal of Speech Language and Hearing Research* 57: 1804–1816.

Chiat, S., and K. Polisenska. 2016. "A Framework for Cross-Linguistic Nonword Repetition Tests: Effects of Bilingualism and Socioeconomic Status on Children's Performance." *Journal of Speech Language and Hearing Research* 59: 1179–1189.

Coady, J., and J. Evans. 2008. "Uses and Interpretations of Non-word Repetition Tests in Children with and without Specfic Language Impairment." *International Journal of Language and Communication Disorders* 43: 1–40.

Collier, V. 1989. "How Long? A Synthesis of Research on Academic Achievement in a Second Language." *TESOL Quarterly* 23: 509–531.

Conti-Ramsden, G., N. Botting, and B. Faragher. 2001. "Psycholinguistic Markers for Specific Language Impairment (SLI)." *Journal of Child Psychology and Psychiatry* 42: 741–748.

Dollaghan, C., and T. Campbell. 1998. "Nonword Repetition and Child Language Impairment." *Journal of Speech Language and Hearing Research* 41: 1136–1146.

Duncan, T., and J. Paradis. 2016. "English Language Learners' Nonword Repetition Performance: The Influence of Age, L2 Vocabulary Size, Length of L2 Exposure and L1 Phonology." *Journal of Speech Language and Hearing Research* 59: 39–48.

Dunn, L., and L. Dunn. 1994. *Peabody Picture Vocabulary Test*. 3rd ed. Circle Pines, MN: American Guidance Service.

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. 2011a. "The Relationship Between Bilingual Exposure and Vocabulary Development." *International Journal of Bilingualism* 15 (5): 426–445.

Elin Thordardottir. 2011b. *Milli mála: mat á íslensku sem öðru máli* [Milli mála: Assessment Test for Icelandic as a Second Language]. Reykjavik: Framsaga.

Elin Thordardottir. 2015. "The Relationship Between Bilingual Exposure and Morphosyntactic Development." *International Journal of Speech-Language Pathology* 17: 97–114.

Elin Thordardottir. 2016. "Grammatical Morphology is Not a Sensitive Marker of Language Impairment in Icelandic in Children Aged 4–14 Years." *Journal of Communication Disorders* 62: 82–100.

Elin Thordardottir, and M. Brandeker. 2013. "The Effect of Bilingual Exposure Versus Language Impairment on Nonword Repetition and Sentence Imitation Scores." *Journal of Communication Disorders* 46: 1–16.

Elin Thordardottir, and Gudbjörg Kristín Eiríksdottir. 2012. "*Use of Icelandic Morphology by School-Age L2 Speakers*." Poster presented at the bilingual and multilingual interaction conference, ESRC Research Centre on Bilingualism, Bangor, Wales, March–April.

Elin Thordardottir. forthcoming. "Amount Trumps Timing in Bilingual Vocabulary Acquisition: Effects of Input in Simultaneous and Sequential School-Age Bilinguals." *International Journal of Bilingualism*.

Elin Thordardottir, and Anna Gudrun Juliusdottir. 2013. "Icelandic as a Second Language: A Longitudinal Study of Language Knowledge and Processing by School-Age Children." *International Journal of Bilingual Education and Bilingualism* 16: 411–435.

Elin Thordardottir, E. Kehayia, B. Mazer, N. Lessard, A. Majnemer, A. Sutton, N. Trudeau, and G. Chilingarian. 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, A. Rothenberg, M.-E. Rivard, and R. Naves. 2006. "Bilingual Assessment: Can Overall Proficiency be Estimated from Separate Measurement of Two Languages?" *Journal of Multilingual Communication Disorders* 4: 1–21.

Genesee, F., K. Lindholm-Leary, W. Saunders, and D. Christian. 2005. "English Language Learners in U.S. Schools: An Overview of Research Findings." *Journal of Education for Students Placed at Risk* 10: 363–385.

Gibson, T., C. Summer, E. Peña, L. Bedore, R. Gillam, and T. Bohman. 2015. "The Role of Phonological Structure and Experience in Bilingual Children's Nonword Repetition Performance." *Bilingualism: Language and Cognition* 18: 551–560.

Girbau, D., and R. G. Schwartz. 2008. "Phonological Working Memory in Spanish–English Bilingual Children With and Without Specific Language Impairment." *Journal of Communication Disorders* 41: 124–145.

Hallé, P., Y.-C. Chang, and C. Best. 2004. "Identification and Discrimination of Mandarin Chinese Tones by Mandarin Chinese vs. French Listeners." *Journal of Phonetics* 32: 395–421.

Hamill, D., and P. Newcomer. 1997. *Test of Language Development (TOLD)*. Austin, TX: Pro-Ed.

Hammer, C., G. Jia, and Y. Uchikoshi. 2011. "Language and Literacy Development of Dual Language Learners Growing Up in the United States: A Call for Research." *Child Development Perspectives* 5: 4–9.

Hammer, C., F. Lawrence, and A. Miccio. 2008. "Exposure to English Before and After Entry Into Head Start: Bilingual Children's Receptive Language Growth in Spanish and English." *International Journal of Bilingual Education and Bilingualism* 11: 30–56.

Hao, Y.-C. 2012. "Second Language Acquisition of Mandarin Chinese Tones by Tonal and Non-Tonal Language Speakers." *Journal of Phonetics* 40: 269–279.

Hu, C.-F., and M. Schuele. 2015. "When Language Experience Fails to Explain Word Reading Development: Early Cognitive and Linguistic Profiles of Young Foreign Language Learners." *The Modern Language Journal* 99: 754–770.

Jackson, C., C. Schatschneider, and L. Leacox. 2014. "Longitudinal Analysis of Receptive Vocabulary Growth in Young Spanish English–Speaking Children from Migrant Families." *Language Speech and Hearing Services in Schools* 45: 40–51.

Ólafsdóttir, Sigríður, Freyja Birgisdóttir, Hrafnhildur Ragnarsdóttir, and Sigurgrímur Skúlason. 2016. "*Orðaforði og lesskilningur hjá börnum með íslensku sem annað mál: Áhrif aldurs við komuna til Íslands*" [Icelandic Vocabulary and Reading Comprehension among Icelandic Second Language Learners: Influence of Age on Arrival in Iceland]. *Netla* Special Issue: 1–25.

Lee, H. J., Y. T. Kim, and D. Yim. 2013. "Nonword Repetition Performance in Korean-English Bilingual Children." *International Journal of Speech-Language Pathology* 15: 375–382.

Li, B., J. Shao, and M. Bao. 2017. "Effects of Phonetic Similarity in the Identification of Mandarin Tones." *Journal of Psycholinguistic Research* 46: 107–124.

Maddiesen, I. 2013. "Tone." In *The World Atlas of Language Structures Online*, edited by M. Dryer and M. Haspelmath. Leipzig: Max Planck Institute on Evolutionary Anthropology. http://wals/info/chapter13.

Mueller Gathercole, V., I. Kennedy, and E. Thomas. 2016. "Socio-Economic Level and Bilinguals' Performance on Language and Cognitive Measures." *Bilingualism: Language and Cognition* 19: 1057–1078.

Muircheartaigh, J., and T. Hickey. 2008. "Academic Outcome, Anxiety and Attitudes in Early and Late Immersion in Ireland." *International Journal of Bilingual Education and Bilingualism* 11: 558–576.

Nowenstein, Iris Edda, and Elin Thordardottir. Forthcoming. "Setningager og samræmi í íslensku sem öru máli: Dulinn munur í frásagnarmálsnum táninga [Syntactic Structure and Agreement in Icelandic as a Second Language: A Hidden Difference in Adolescent's Narrative Samples]." *Talfræingurinn*.

Oller, K., B. Pearson, and A. Cobo-Lewis. 2007. "Profile Effects in Early Bilingual Language and Literacy." *Applied Psycholinguistics* 28: 191–230.

Paradis, J., Y. Tulpan, and A. Arppe. 2016. "Chinese L1 Children's English L2 Verb Morphology Over Time: Individual Variation in Long-Term Outcomes." *Journal of Child Language* 43: 553–580.

Pham, G., and K. Kohnert. 2014. "A Longitudinal Study of Lexical Development in Children Learning Vietnamese and English." *Child Development* 85: 767–782.

Rydland, V., V. Grøver, and J. Lawrence. 2014. "The Second Language Vocabulary Trajectories of Turkish Immigrant Children in Norway from Ages Five to Ten: The Role of Preschool Talk Exposure, Maternal Education and Co-Ethnic Concentration in the Neighborhood." *Journal of Child Language* 41: 352–381.

Scheele, A., P. Leseman, and A. Mayo. 2010. "The Home Language Environment of Monolingual and Bilingual Children and Their Language Proficiency." *Applied Psycholinguistics* 31: 117–140.

Simos, P., G. Sideridis, A. Mouzak, and A. Chatzidaki. 2014. "Vocabulary Growth in Second Language among Immigrant School-Aged Children in Greece." *Applied Psycholinguistics* 35: 621–647.

Smithson, L., J. Paradis, and E. Nicoladis. 2014. "Bilingualism and Receptive Vocabulary Achievement: Could Socio-Cultural Context Make a Difference?" *Bilingualism: Language and Cognition* 17: 810–821.

Símonardóttir, Ingibjörg, and Einar Guðmundsson. 1996. *TOLD-I málþroskaprófíslensk staðfærsla* [Language Assessment Test, Icelandic Adaptation]. Reykjavik: Rannsóknarstofnun uppeldis- og menntamála.

Símonardóttir, Ingibjörg, Einar Guðmundsson, Steingrímur Skúlason, and Sigríður Pétursdóttir. 1995. *TOLD-P málþroskaprófíslensk staðfærsla* [Language Assessment Test, Icelandic Adaptation]. Reykjavik: Rannsóknarstofnun uppeldis- og menntamála.

Statistics Iceland [Hagstofa Íslands]. 2016. http://www.hagstofa.is.

Steien, G., and W. Dommelen. 2016. "The Production of Norwegian Tones by Multilingual Non-native Speakers." *International Journal of Bilingualism*. doi:10.1177/1367006916673218.

Stokes, S., A. Wong, P. Fletcher, and L. Leonard. 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.

Unsworth, S. 2016. "Early Child L2 Acquisition: Age or Input Effects? Neither or Both?" *Journal of Child Language* 43: 608–634.

Wang, T., and J. Saffran. 2014. "Statistical Learning in a Tonal Language: The Influence of Bilingualism and Previous Language Experience." *Frontiers Psychology* 953: 1–9.