CrossMark

RESEARCH PAPER

Literate and preliterate children show different learning patterns in an artificial language learning task

Naomi Havron (ORCID) · **Limor Raviv** · **Inbal Arnon**

Abstract Literacy affects many aspects of cognitive and linguistic processing. Among them, it increases the salience of words as units of linguistic processing. Here, we explored the impact of literacy acquisition on children's learning of an artifical language. Recent accounts of L1–L2 differences relate adults' greater difficulty with language learning to their smaller reliance on multiword units. In particular, multiword units are claimed to be beneficial for learning opaque grammatical relations like grammatical gender. Since literacy impacts the reliance on words as units of processing, we ask if and how acquiring literacy may change children's language-learning results. We looked at children's success in learning novel noun labels relative to their success in learning article-noun gender agreement, before and after learning to read. We found that preliterate first graders were better at learning agreement (larger units) than at learning nouns (smaller units), and that the difference between the two trial types significantly decreased after these children acquired literacy. In contrast, literate third graders were as good in both trial types. These findings suggest that literacy affects not only language processing, but also leads to important differences in language learning. They support the idea that some of children's advantage in language learning comes from their previous knowledge and experience with language—and specifically, their lack of experience with written texts.

Keywords Language learning · Literacy · Artificial language · Communication · Linguistic units

N. Havron (✉)
Laboratoire de Sciences Cognitives et Psycholinguistique, Dept d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS, Paris, France
e-mail: naomi.havron@mail.huji.ac.il

L. Raviv
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

I. Arnon
Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel

Introduction

Contrary to the classical view that writing is simply a representation of spoken language, learning to read and write has far-reaching cognitive and linguistic consequences. In the past few decades, literacy has been shown to affect the functional organization of the brain (Castro-Caldas et al. 1998; Dehaene et al. 2010; Dehaene and Cohen 2011), as well as multiple aspects of cognitive and linguistic processing (Cheung and Chen 2004; Dellatolas et al. 2003; Huettig and Mishra 2014; Morais et al. 1986; Ventura et al. 2013). For example, literate adults display quantitatively better short term memory abilities (Kosmidis et al. 2011),

and also differ qualitatively from illiterate adults in their performance on short term memory and long term retrieval tasks – namely, literates rely more on semantic features and less on phonological features in retrieval (Reis and Castro-Caldas 1997). Moreover, literacy promotes the development of metalinguistic abilities such as phonological and morphological awareness, such that literate children are more able to recognize and manipulate the morphemes and phonemes that make up words (Bassetti 2005; Brunswick et al. 2012; Kurvers et al. 2006; Levin et al. 1999).

Importantly, literacy is associated with enhanced awareness of word boundaries (Havron and Arnon 2017a, b; Holden and MacGinitie, 1972; Kurvers et al. 2007; Veldhuis and Kurvers 2012). In spoken language, words are usually not separated by pauses, and it is not always clear where one word ends and another begins. Yet in many writing systems, words are separated by spaces, encouraging readers to process language more along word boundaries, and attend more to word units. Indeed, in such writing systems, literacy was found to influence the size of the linguistic units that children and adults segment speech into: compared to preliterate children and illiterate adults, literate children and adults are more likely to segment speech into single word units, and less likely to segment them into multiword units (Havron and Arnon 2017a, b; Veldhuis and Kurvers 2012). In writing systems which do not mark word boundaries (e.g., Chinese), such an effect is not found (Bassetti 2005). Put differently, literacy in writing systems that mark word boundaries boosts the salience of words as the relevant units of linguistic processing, and biases against the use of larger, multiword phrases as units of processing.

The current study asks whether, in addition to changing the way people perceive and process language, literacy also impacts the way they learn a new language. Since literacy increases the salience of words as units of processing, it may shift learners' attention from multiword phrases to single words, with consequences for language learning outcomes. Specifically, such differences in the size of processing units might impact certain aspects of language learning. A recent theoretical proposal relates adults' difficulty in learning a second language to their smaller reliance (compared to children) on multiword units during learning (Arnon

2010; Arnon and Christiansen 2017). Multiword units are predicted to be beneficial for learning adjacent grammatical relations, such as gender agreement, by increasing the association between the linguistic elements in question. In support of this prediction, having adults learn from larger units can facilitate their mastery of such relations (Arnon and Christiansen 2017; Arnon and Ramscar 2012; Paul and Grüter 2016; Siegelman and Arnon 2015). Literacy may similarly affect how a new language is learned by shifting learners' attention from multiword phrases to single words. Interestingly, while this shift may aid single word learning (i.e. vocabulary), it might actually hinder learning of multiword relations (i.e., grammatical agreements). To test whether literacy indeed contributes to vocabulary learning, but hinders the learning of grammatical relations, we examine children's learning of a grammatical-gender-like system in an artificial language twice, before and after learning to read. Given that literacy affects unit size of processing, we predict that learning to read will lead to differences in the way children acquire certain aspects of the artificial language. In particular, we predict that preliterate children will be better at learning article-noun agreement (larger units) than at learning nouns (smaller units), while literate children will not, and might even show the opposite pattern.

This prediction is motivated by findings from the artificial language learning and second language acquisition literatures. First, adults show better learning of gender agreement in an artificial language when they are exposed first to full utterances and only later to nouns in isolation (compared to first hearing isolated nouns and only then full utterances; Arnon and Ramscar 2012), and show better memory for article-noun mappings when they are exposed to unsegmented utterances (without pauses between words) compared to segmented utterances (with pauses inserted between words; Siegelman and Arnon 2015; Havron and Arnon, under revision). The idea is that learning from larger units leads to better learning of semantically opaque grammatical relations such as gender agreement by increasing the association between the grammatical element and the word. Learning words first, in contrast, may block (or reduce) these associations (see Arnon and Ramscar 2012, for more details and a formal model). Indeed, in

Siegelman and Arnon (2015), participants who treated the article and noun as one word rather than two separate words (as indicated by their typed responses) showed better learning of the agreement between them. A similar pattern was found in natural language learning, looking at the learning of classifier-noun associations in Chinese (Paul and Grüter 2016). Together, these studies show that unit size affects language learning, with multiword units facilitating the learning of non-transparent grammatical relations. Interestingly, in all these studies, multiword units did not facilitate learning the noun labels themselves (Arnon and Ramscar 2012; Havron and Arnon, under revision; Paul and Grüter 2016; Siegelman and Arnon 2015), suggesting that while larger units promote learning of opaque grammatical relations, they do not necessarily aid vocabulary learning.

The literature on second language learning makes a similar distinction between learning vocabulary and learning grammatical relations. Studies on second language acquisition repeatedly find that those immigrating at a younger age eventually reach higher proficiency in the new language (e.g., DeKeyser et al. 2010; Flege et al. 1999; Hakuta et al. 2003; Johnson and Newport 1989). However, while domains such as learning collocations, formulaic language, and grammatical gender agreement are highly affected by age of arrival (Laufer and Waldman 2011; Sugiura 2002; Wray 1999, 2000; Yorio 1989; Dewaele and Véronique 2001; Lew-Williams and Fernald 2010), vocabulary acquisition seems less affected by age (MacWhinney 2005). This pattern is consistent with the idea that some of children's advantage in language learning (and in particular, in learning grammatical relations) is related to younger children's tendency to learn from larger, multiword units (Arnon and Christiansen 2014, 2017; Arnon 2010): If adults rely on larger units to a lesser extent compared to children, this will negatively impact their learning of certain grammatical relations—but not of vocabulary items. This difference in unit size may be related to literacy skills. Adults may struggle less with vocabulary than with opaque grammatical relations because of their superior reading skills, which lead them to rely less on multiword units and more on single word units (Havron and Arnon 2017a, b; Veldhuis and Kurvers 2012). If so, then learning to read should benefit some aspects of language learning (i.e., vocabulary learning) but not others (i.e., gender agreement).

Here, we explore this prediction by looking at the effect of literacy on children's learning of grammatical gender in an artificial language. We suggest that acquiring literacy will lead children to attend to smaller sized units, which will, in turn, lead them to show more adult-like learning patterns. Specifically, we predict that literacy acquisition will lead to better performance on vocabulary relative to grammatical agreement. To test this prediction, we use an artificial-language learning paradigm similar to the ones described above. We examine the effect of literacy on children's learning outcomes by looking at first-graders twice six-month apart, before and after they learn to read. We also test a control group of third graders at the same time intervals. Since the third graders will not have acquired literacy in the time between testing sessions (they could already read at the first session), their performance serves as a control for any general improvements having to do with being tested twice on the same paradigm. If the pattern of results changes in the second session for first graders, but not third graders, this would suggest that literacy acquisition contributed to this change. Of course, the ability to read is not an all-or-nothing skill, and even third graders are still improving their reading abilities. However, the additional gains over those six months, are not the same as the initial reading acquisition experienced by the first graders. It is the actual discovery of word boundaries, rather than the continued practice with them, that seems to impact unit size in language processing (Havron and Arnon 2017a). Therefore, we expect the first graders, but not the third graders, to show a significant change in their learning patterns.

All children are taught an artificial language with gender-like agreement between articles and nouns, similar to the one used in previous studies (Siegelman and Arnon 2015; Havron and Arnon, under revision). During the learning phase, children see different objects and hear their matching descriptions in the new language (i.e., an utterance with a carrier phrase, then one of two articles followed by a noun). Importantly, unlike previous work we do not manipulate how segmented the input is, but instead expose all the children to unsegmented input in order to see how literacy will impact their processing of it. After exposure, children are tested on their knowledge of vocabulary (which noun represents which object) and grammatical agreement (which article appears with

which noun). If acquiring literacy affects unit size and leads children to attend to (and learn from) smaller linguistic units, then preliterate children should show better learning of article-noun mappings compared to vocabulary learning in the first session. This agreement advantage should be non-existent or reversed in the control group of third grade literate children. When comparing the two sessions, learning to read should improve vocabulary learning in first graders (who learned to read in the meantime), but possibly hinder their learning of gender agreement—thus eliminating the agreement advantage. In contrast, the already-literate third graders should not differ between sessions. Literacy was operationalized as the ability to fluently read a short paragraph in Hebrew without diacritics. We focus on first graders who are preliterate in the first session and reached sufficient levels of literacy in the second session for two reasons. First, children who were still preliterate at the second session, while (still) not diagnosed with any learning disability, may constitute a separate population, and the factors which led to their inability to read may also affect their artificial language learning results. Second, children who could already read at the beginning of the school year may also constitute a separate group. Early readers have been found to differ from non-early readers in their general language development and verbal intelligence (Huba and Ramisetty-Mikler 1995), memory and visuospatial ability (Brunswick et al. 2012), and inhibitory control (Blair and Peters Razza 2007). Thus, factors confounding with early or late literacy may make it harder for us to isolate the effect of literacy from that of cognitive abilities. In order to disentangle the effect of literacy from other confounding factors such as age and cognitive abilities, we also tested children's productive vocabulary in their native language (with picture naming) and short-term memory (with digit span), and controlled for these in our analyses.

In sum, the present study tests the hypothesis that literacy leads children to learn differently from the same input. We predict that preliterate first graders will be better at learning article-noun agreement than at learning nouns (contrary to what was repeatedly found for adults), while third graders who can already read will not. We also predict that following literacy acquisition, first graders will show a reduction in the article-noun agreement advantage, while third graders will show no change.

Methods

Participants

86 first and third graders participated in the study. All children were from middle-to-high socio-economic status attending schools in the center of Israel. The study was approved by the ethics committee of the department of psychology at the Hebrew University of Jerusalem, and by the Israeli Ministry of Education. Children were tested at their school after parents signed consent forms. All children were native speakers of Hebrew. Of the 86 students who participated in the study, five were removed from further analyses: Four were special education students learning in a regular class, and one requested to stop after the first task. This left us with a final sample of 31 first graders (mean age 6.45, 18 girls) and 27 third graders (mean age 8.89, 14 girls). Additionally, two other groups of children were excluded from analysis (see full explanation above): first graders who were already able to read the paragraph at the first session (n = 12); and first graders that were still not able to read the paragraph in the second session (n = 11).

Materials

Cognitive assessments

Short-term memory Evaluated using the digit span task (only forward section; Kaufman 1994). We only used the digits one to five to prevent arithmetic ability from affecting performance.

Vocabulary measure Evaluated using the Gitit-Kave (Kavé 2006) Hebrew naming test for productive vocabulary, a validated tool suitable for use with children at this age, and adults.

Literacy assessment

Performed using a rated paragraph reading task. The paragraph was at a level suitable for readers at the end of the first grade (the same paragraph was used with the third graders). The responses were recorded and later rated by two native speakers who have not met the participants and were blind to children's grade and testing session. Reading was rated for fluency (0–3) and confidence (0–3), and aggregated to create a single score (with a minimum score of 0 and a maximum

score of 6). This scale was developed by Tarone et al. (2007), and used previously in Hebrew with literate and illiterate adults and children (Havron and Arnon 2017a, b). We considered children who scored below four on the literacy scale as preliterate. Inter-rater reliability was high (weighted Kappa = 0.88, $z = 11.2$, $p < 0.001$). See appendix 1 for instructions for raters.

Artificial language

The artificial language was similar to that taught in Siegelman and Arnon (2015), but with several modifications to make it more child friendly. There were fewer words to learn (eight vs. twelve in Siegelman and Arnon's study), the carrier phrase used before the article-noun pair (equivalent to "this is", see below) was shorter, and the syllable structure of words was simpler (only CV syllables, except for the carrier phrase which ended with a CVC). Children were also given shorter exposure (one block of exposure as opposed to two in Siegelman and Arnon's Study), and were asked to repeat the utterances out loud to facilitate learning.

The artificial language contained eight unique bisyllabic novel labels describing concrete items (e.g., "keba" for clock, "nadi" for chair), and two articles ("do" and "ga", see appendix for a full list of stimuli). All objects had high-frequency, early-acquired Hebrew labels. The novel labels were divided into two noun 'classes' which were only discriminated based on the article, so that each noun appeared with only one of the articles. There were no semantic or phonological cues to class membership—the only cue was distributional (which article the label appeared with). To ensure that learning was not affected by the grammatical gender of labels in children's first language, the two noun 'classes' were matched on the gender of the nouns in Hebrew: each class had an equal number of masculine and feminine nouns, so that classifying nouns in the artificial language based on grammatical gender in Hebrew would not result in successful learning.

The artificial language had a fixed word order: Utterances always started with a carrier phrase ("Kamek"), followed by an article and a noun. A 250 ms silence separated the carrier phrase from the article and the nouns (see example 1).

(1) **Kamek** [250ms] **do.tibo**

 carrier phrase article.noun

The same recorded token of each noun, article and carrier phrase was used throughout the experiment. The duration of the two articles was identical to ensure they were equally prominent. For the second session, we used another artificial language, which was identical to the first language in training and testing parameters, but had a different carrier phrase, different articles, different nouns and different objects (See appendix 2 for the two languages). Thus, the two versions of the artificial language had the same structure, but differed in the lexical form of the items. We replaced the forms so that children will not rely on memory for the particular items or articles.

Procedure

Children were tested individually in a quiet room in their school. The first session was conducted in October (Israeli children have many holidays in September and do not learn much in the first month of school). The second session was conducted six months later in April (before they went on spring break), after the teachers reported most children could already read. The order of the tasks in a given session was always the same. In the first session, children first completed a productive-vocabulary task, followed by the artificial-language learning task, and ending with the literacy assessment. In the second session, they first completed the artificial-language task (with the new carrier phrase, noun labels and articles), followed by a working memory task (the digit span), and a second literacy assessment last.

In the exposure phase of the artificial-language learning task, children were told that they were about to learn how different objects are called in an alien language. The exposure consisted of four repetitions of each of the eight utterances with items presented in a random order (32 trials in total, lasting approximately 4 min). At every trial, children saw a cartoon alien pointing at objects and heard the utterance describing it (see Fig. 1 for a sample exposure trial). Children were asked to repeat the alien's utterance in order to enhance their learning. Following the
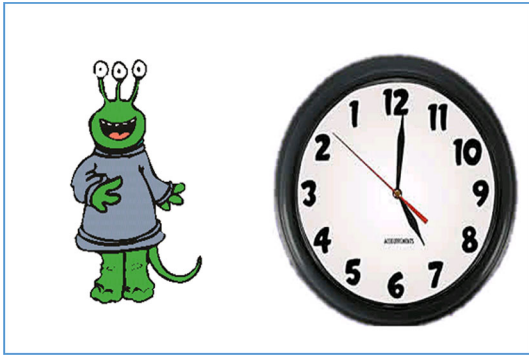
Fig. 1 A sample exposure trial from the experiment

exposure phase, children completed a test phase (lasting about 3 min) where the alien reappeared on the screen, and children were told that it would now say two utterances: One is correct in the language, and one has a mistake. They were then asked to decide which of the utterances was correct. Half of the trials tested vocabulary knowledge: the incorrect utterance was wrong because it had the wrong label for the object appearing on the screen (noun trials). The other half tested article-noun agreement: the incorrect utterance was wrong because it contained the wrong article for that label (agreement trials). There were 16 test trials (8 for vocabulary and 8 for agreement). After completing all tasks, children received a sticker and a certificate of participation.

Results

Descriptive statistics

Literacy

First graders had an average reading score of 1.82 (SD = 1.29) in the first session and an average reading score of 5.23 (SD = 0.64) in the second session, showing that their literacy skills improved significantly between the two sessions [t(60) = 13.06, p < 0.001]. All third graders had the maximal score of 6 in the first session, indicating they could already read.

Control measures

Third graders were significantly better at the digit span than first graders (first graders mean = 6.03, SD = 1.26; third graders mean = 7.23, SD = 1.22; t(54.87) = 3.53, p < 0.001). Both groups' performance on the digit span was higher than the norms, probably because we only used digits 1-5. Third graders also had larger productive vocabulary (first graders mean = 34.61, SD = 4.47; third graders mean = 38.25, SD = 4.29; t(55.53) = 3.23, p = 0.002). These scores are in line with existing norms (e.g., Kavé 2006), and will be used as control measures in all analyses.

Task performance

In the first session, first-graders showed learning of the article-noun agreement but were not above chance in learning the noun-labels (mean = 50%, SD = 13.69%; t(30) = 0, p = 1). By the second session, first graders showed learning of both vocabulary and agreement patterns (see Table 1). The third graders, in contrast, showed similar learning of both vocabulary and agreement patterns in both sessions (see Table 1).

First session: comparing preliterate and literate children

We predicted that preliterate first graders (n = 31) will show better learning of article-noun agreement than of nouns (unlike what has been found for adults in previous studies, Arnon and Ramscar 2012; Havron and Arnon, under revision; Paul and Grüter 2016; Siegelman and Arnon 2015), while third graders (n = 27), who can already read, will not (they would either be better at nouns or equal in both trial types). To test this prediction, we calculated a *difference score* for each participant by subtracting their score on agreement trials from their score on noun trials. A negative score means that a participant showed better learning on article trials (agreement advantage), and a positive score means they showed better learning on noun trials (noun advantage). See Fig. 2 for the distribution of difference scores in the first session.

Next, we ran a linear regression to predict this difference score as a factor of AGE GROUP (first vs. third grade, dummy coded with third graders as the reference level), Productive Vocabulary Score

Table 1 Mean scores (SDs in brackets) in the artificial language learning task in the two sessions (asterisks mark difference from chance, which was calculated with a single sample *t* test)

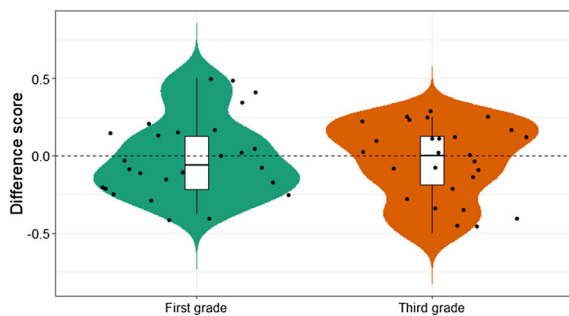| | Session 1 | | Session 2 | |
	Noun	Agreement	Noun	Agreement
Third graders	66.5%** (20.64%)	68.5%*** (14.93)	65.28%** (20.31%)	68.5%*** (14.03%)
First graders	50.00% (13.69%)	64.11%*** (16.69%)	56.85%* (17.63%)	60.48%** (18.57%)
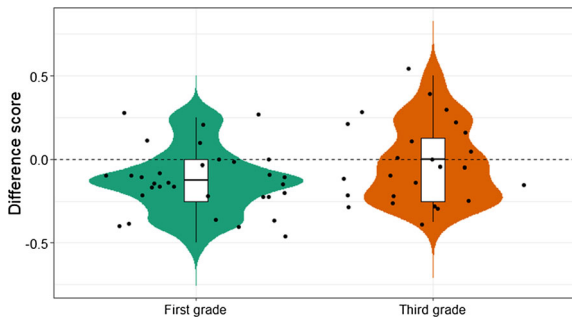
Fig. 2 difference score for the two age groups in session 1. Negative scores mean that participants were better on agreement trials than on noun trials (i.e., an agreement advantage). Each individual point is a participant. The lower and upper hinges correspond to the first and third quartiles, the black lines within the box plots represent the means. The top whiskers denote the maximum value, and the bottom whiskers the minimum value. First graders have a negative difference score, which for third graders there is no difference between the two trial types

(continuous, centered) and Short-Term Memory Score (continuous, centered). The model showed a significant effect of AGE GROUP, with preliterate first graders showing a bigger difference than literate third graders ($\beta = 0.145$, SE = 0.067, $p = 0.03$). No other predictor was significant. Indeed, the third graders were equally good on article-noun agreement and noun trials [2% difference, t(48) = 0.39, $p = 0.35$], while preliterate first graders were significantly better on article-noun agreement trials than on noun trials [13.42% difference, t(48) = 3.56, $p = 0.0004$]. Note however, that preliterate children were not above chance on noun trials: they learned which article appears with each noun label, but had difficulty matching the nouns to the correct label. To conclude, we found that preliterate children showed an agreement advantage while literate children did not.

Second session: comparing children before and after learning to read

In the second session, after first graders have learned to read, we predicted that the agreement advantage found in the first session will disappear. Third graders, who did not show substantial gains in reading during that time, were hypothesized to show a similar pattern of results to that found in the first session. That is, first graders and third graders are now expected to show the same pattern of results and not differ in the size of the agreement advantage.

To examine this, we ran the exact same model as in session one. This time, AGE GROUP was not a significant predictor of the difference score between agreement and noun trials ($\beta = 0.032$, SE = 0.076, $p = 0.68$). None of the other factors were significant. This means that unlike in the first session, first graders and third graders did not differ in the size of their agreement advantage. As predicted, third graders did not show a significant change in their performance across sessions: the difference between noun trials and agreement trials was 2% on the first session and 4.32% in the second session [t(24) = 0.33, $p = 0.63$, Cohen's d = 0.06]. In contrast, first graders showed a significant change with the difference between noun trials and agreement trials decreasing from 12.89% on the first session to only 3.63% in the second session [t(30) = 2.03, $p = 0.03$, Cohen's d = − 3.18]. First graders also showed evidence of learning the nouns in the second session (see Fig. 3). This pattern is consistent with the suggestion that first grader's learning styles changed as a function of learning to read.

Fig. 3 difference score for the two age groups in session 2. Negative scores mean that participants were better on agreement trials than on noun trials (i.e., an agreement advantage). Each individual point is a participant. The lower and upper hinges correspond to the first and third quartiles, the black lines within the box plots represent the means. The top whiskers denote the maximum value, and the bottom whiskers the minimum value. For both age groups there is now no difference between the two trial types

Discussion

In the current paper, we tested the hypothesis that literacy can impact language learning. Previous studies conducted on adults consistently found that they learn nouns (vocabulary) better than article-noun agreement (Arnon and Ramscar 2012; Havron and Arnon, under revision; Paul and Grüter 2016; Siegelman and Arnon 2015). These findings are in line with the L2 literature, which shows that adults struggle with learning grammatical gender agreement (e.g. Dewaele and Véronique 2001), but have less difficulty with learning vocabulary (MacWhinney 2005). Adults' greater difficulty in learning such grammatical relations has been attributed to differences in linguistic-unit size (Arnon 2010; Arnon and Christiansen 2017): Children tend to learn language from larger units (multiword phrases), while adults tend to learn language from smaller units (single words). Learning from smaller units is postulated to hinder learning of opaque dependencies such as gender agreement, while possibly facilitating vocabulary learning. Here, we hypothesized that some of the difference in unit size between children and adults is related to literacy acquisition, as literacy has been shown to affect processing unit size by increasing the salience of single words compared to multiword phrases (Gombert, 1994; Havron and Arnon 2017a, b; Holden and MacGinitie 1972; Kurvers et al. 2007; Veldhuis and Kurvers 2012). In the current paper we extended this

work by asking whether literacy also impacts learning of a novel language.

Since learning to read is related to less reliance on multiword units (e.g., Havron and Arnon 2017a, b; Veldhuis and Kurvers 2012), we predicted that preliterate children will show better learning of article-noun agreement compared to learning nouns, while literate children will not show such an agreement advantage. We tested this prediction by examining children's learning outcomes on an artificial language learning task, before and after they learn to read. We compared preliterate first graders to literate third graders at the beginning of the year, and then compared first graders to themselves after they learned to read. We predicted that preliterate first-graders will be better at learning the article-noun agreement than at learning the nouns (i.e., an article-agreement advantage), while third graders who can already read will not show this pattern. We also predicted that following literacy acquisition, first graders will show a reduction in the article-noun agreement advantage, while third graders will show no change in their learning patterns.

As predicted, we found that literacy impacted children's learning patterns. Preliterate first graders showed an article-noun agreement advantage in the first session, but third graders did not: while the preliterate first graders were better at learning the article-noun agreement, literate third graders learned both equally well. In the second session, first graders who became literate showed a large reduction in the article-agreement advantage, but third graders showed no change. This pattern of improvement in noun learning but not in agreement learning can therefore be associated with the acquisition of literacy. We suggest that this is because literacy impacts the size of the units used in learning by biasing learners' attention towards smaller units (single words rather than multiword phrases), which in turn affects their segmentation of a new language. Supporting this claim, studies have shown that vocabulary learning is aided by segmentation: Adults show signs of lexicalization in an artificial language only when there are clear cues to segmentation (Fernandes et al. 2009). Similarly, infants' early segmentation abilities predict their vocabulary at 24 months and other language outcomes at 4 years (Newman et al. 2006a, b; Singh et al. 2012). In contrast, under-segmentation aids learning of opaque article-noun agreement, but not of vocabulary (Siegelman and Arnon 2015). Our findings are in line

with the proposal that literacy affects segmentation. They suggest that literacy contributes to a transition in learning style, and that its acquisition underlies (some of) the documented differences between children and adults in language learning. Literacy is often confounded with age in studies of L2 acquisition: adult L2 learners are almost always literate, while child L2 learners are often not (if tested before entering school). Consequently, some of the observed differences may be related to literacy acquisition, rather than age.

Even though research shows that literacy has a positive effect on certain aspects of L2 learning (e.g., recall of oral recasts and elicited imitation, Tarone and Bigelow 2005), our results suggest that learning to read can also have detrimental consequences for learning certain aspects of language, namely semantically-opaque grammatical relationships (such as gender agreement). Since literacy contributes to many aspects of cognitive development such as memory (Reis and Castro-Caldas 1997), prediction (Mishra et al. 2012), and even oral repetition of pseudowords (Castro-Caldas et al. 1998), it may be counter-intuitive to consider its downsides. However, in line with this hypothesis, a recent study with illiterate and literate adults learning to read in an L2 found a negative correlation between their literacy levels and their self-reported comprehension proficiency in the L2 (Havron and Arnon 2017a). In that study, adults that had better L1 reading skills actually reported lower levels of proficiency in their second language. While that study did not control for factors such as age of first exposure or amount of exposure, it provides preliminary support for the idea that illiteracy or pre-literacy could actually assist L2 learning. This might come as a surprise not only in light of the cognitive difference between literate and illiterates, but also since illiterate people do not have access to many important language resources (e.g., book reading is an important contributor to L2 ability, Warwick and Francis 1983). However, this is less surprising when we remember that the most prominent group of language learners are infants and young children who cannot read, yet reach much higher levels in their L2 than adults do. That is, illiteracy is the most common context in which languages are naturally learned. Moreover, non-literate societies seems to have no problem acquiring not only their native language, but also multiple languages (Aikhenvald 2008).

Interestingly, even the literate children in our study did not show the vocabulary advantage previously found with adults in similar tasks (Arnon and Ramscar 2012; Havron and Arnon, under revision; Paul and Grüter 2016; Siegelman and Arnon 2015). Even after the acquisition of literacy, children learned the nouns and article-noun agreement equally well. The fact that the vocabulary advantage continues to develop into adulthood may stem from a plurality of factors, including the development of memory abilities (Gupta and MacWhinney 1997), adults' superior ability to focus attention on parts of the utterance and ignore others (Ramscar and Gitcho 2007; Thompson-Schill et al. 2009), or their metalinguistic awareness (e.g. Kurvers and Uri 2006; Ravid and Malenky 2001). As children develop metalinguistic skills and start to consciously reflect on their L1, they may become more attuned to patterns and linguistic entities in their L2 or in an artificial language. The development of memory and metalinguistic awareness continues after the age of seven, well into the upper-elementary years (Nagy 2007; Schneider 2010). The fact that executive functions and metalinguistic awareness continue to develop after the third grade might explain why third graders, while different than preliterate first graders, do not show the same result pattern as adults.

We assessed unit size only in an indirect way: children's segmentation was assessed through the pattern of results obtained in the learning task, yet no direct measure of segmentation or unit size was collected. This limits the conclusions that can be drawn from our study about the relation between literacy and segmentation. For example, performance may be influenced not only by the segmentation abilities, but also by the quality and nature of the phonological representations, which are themselves affected by literacy (e.g., Cheung and Chen 2004; Kolinsky et al. 1987). Future work is needed to substantiate the link between literacy, changes in unit size, and learning outcomes. Another limitation of the current study is that in the first session (but not in the second), first graders were not above chance in learning the meanings of the nouns, although they did learn the agreement between the nouns and their articles. It is possible that the relatively short exposure led children to learn 'long words' rather than article-noun agreement patterns. We are currently exploring this possibility by increasing children's exposure, and by examining their performance when exposed to

segmented utterances. Importantly, the process of acquisition whereby one starts by associating a multiword 'chunk' with meaning, and only then segments it, is one that can be seen in child language learning. Children's early speech includes multiword utterances like 'what-is-this' or 'how-are-you' (Peters 1983), and many of their early multiword productions are relatively frozen, showing little evidence of productivity of their parts (Lieven et al. 1997; Lieven et al. 2003). Young infants show memory of multiword phrases, long before they can produce them, another indication of their use, at least at the perceptual level (Mandel et al. 1994).

In sum, the present study sought to examine whether literacy acquisition affects language learning by impacting the linguistic building blocks children learn from. We found that literacy was associated with a smaller agreement advantage—both when comparing preliterate first graders to literate third graders; and when comparing first graders before and after they learn to read. Since learning agreement patterns has been shown to be facilitated by larger units of learning, this suggests that preliterate children rely on larger units than literate children. This change in learning-style may be part of what causes older children and adults' poorer results in learning semantically opaque relations between words in a second language such as gender agreement and collocations (Laufer and Waldman 2011; Sugiura 2002; Wray 1999, 2000; Yorio 1989; Dewaele and Véronique 2001; Lew-Williams and Fernald 2010). While literacy may hinder the learning of such relations, it does benefit learning nouns, consistent with studies which show an adult advantage in vocabulary acquisition (Snow and Hoefnagel-Höhle 1978). Thus, literacy has important advantages that cannot be overlooked for both language processing and for cognition more generally (e.g., Ehri 1979; Olson 1996; Ravid and Tolchinsky 2002).

Compliance with ethical standards

Appendix 1: The guidelines for scoring childrens paragraph reading from recordings

Scores were given in two scales, fluency and confidence. For each scale, the minimum score was 0 and the maximum score was 3, amounting to an overall score ranging from 0 to 6.

Fluency

0—Did not read the paragraph.

1—Read slowly, backtracked, did not understand what they were reading (this is sometimes obvious from how the experimenter is reacting), asked for help from the experimenter. Read words incorrectly and did not try to correct themselves.

2—Started slowly but picked up the pace. Still found it hard to decipher word meanings. Sometimes read the same thing twice, faster the second time.

3—Read comfortably, relatively fast, little difficulty in understanding.

Confidence

0—Did not read the paragraph.

1—Said they do not want to or cannot read (but still agreed to try). If a child did not finish reading the paragraph they will also be given 1, even if they did not hesitate to try.

2—Tried, but was uncertain of their abilities. Asked questions throughout the process. Tried and read (not fluently), but showed no confidence.

3—Approached the paragraph reading without hesitation.

Appendix 2: Stimuli used for the artificial language learning task

Session 1			Session 2		
Meaning (Hebrew gender)	Word	Article	Meaning (Hebrew gender)	Word	Article
Carrier phrase - "there is"	Kamek		Carrier phrase	Notok	
Hat (m)	Bigo	Ga	Bath (f)	Dino	Bu
Bike (f)	Tomu	Ga	Brush (f)	Muka	Bu
Chair (m)	Nadi	Ga	Cup (m)	Kebi	Bu
Pan (f)	Kule	Ga	Plane (m)	Lote	Bu
Clock (m)	Keba	Do	Sock (m)	Melu	Na
Scissors (f)	Tibo	Do	Spoon (f)	Guni	Na
Bed (f)	Mula	Do	Kettle (m)	Gako	Na
Table (m)	Nita	Do	Car (f)	Dota	Na

References

Aikhenvald, A. Y. (2008). Multilingual imperatives: the elaboration of a category in northwest Amazonia. *International Journal of American Linguistics, 74*(2), 189–225.

Arnon, I. (2010). *Starting big—The role of multi-word phrases in language learning and use*. Stanford: Stanford University.

Arnon, I., & Christiansen, M. H. (2014). Chunk-based language acquisition. In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of language development* (pp. 88–90). Thousand Oaks, CA: Sage Publications.

Arnon, I., & Christiansen, M. H. (2017). The role of multiword building blocks in explaining L1–L2 differences. *Topics in Cognitive Science, 9*(3), 621–636. https://doi.org/10.1111/tops.12271.

Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition, 122*(3), 292–305.

Bassetti, B. (2005). In V. Cook & B. Bassetti (Eds.), *Second language writing systems* (pp. 335–356).

Blair, C., & Peters Razza, R. (2007). Relating Effortful Control. *Executive function, and false belief understanding to emerging math and literacy ability in Kindergarten, 78*(2), 647–663.

Brunswick, N., Martin, G. N., & Rippon, G. (2012). Early cognitive profiles of emergent readers: A longitudinal study. *Journal of Experimental Child Psychology, 111*(2), 268–285. https://doi.org/10.1016/j.jecp.2011.08.001.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121*(6), 1053–1063.

Cheung, H., & Chen, H.-C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes, 19*(November), 1–28. https://doi.org/10.1080/01690960344000071.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*(6009), 1359–1364. https://doi.org/10.1126/science.1194140.

DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31*(3), 413.

Dellatolas, G., Braga, L. W., Souza, L. D. O. N., Filho, G. N., Queiroz, E., & Deloche, G. (2003). Cognitive consequences of early phase of literacy. *Journal of the International Neuropsychological Society, 9*(5), 771–782.

Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition, 4*(3), 275–297. https://doi.org/10.1017/S136672890100044X.

Ehri, L. (1979). Linguistic insight: Threshold of reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice*. New York: Academic Press.

Fernandes, T., Kolinsky, R., & Ventura, P. (2009). The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition, 112*(3), 349–366. https://doi.org/10.1016/j.cognition.2009.05.002.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78–104. https://doi.org/10.1006/jmla.1999.2638.

Gombert, J. E. (1994). How do illiterate adults react to metalinguistic training? *Annals of Dyslexia, 44*, 250–269. https://doi.org/10.1007/BF02648164.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: computational and neural

bases. *Brain and Language, 59*(2), 267–333. https://doi.org/10.1006/brln.1997.1819.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical period hypothesis for second language acquisition. *Psychological Science, 14*(650), 31–38.

Havron, N., & Arnon, I. (2017a). Minding the gaps: literacy enhances lexical segmentation in children learning to read. *Journal of Child Language, 44*(6), 1516–1538. https://doi.org/10.1017/S0305000916000623.

Havron, N., & Arnon, I. (2017b). Reading between the words: The effect of literacy on second language lexical segmentation. *Applied Psycholinguistics, 38*(1), 127–153. https://doi.org/10.1017/S0142716416000138.

Holden, M. H., & MacGinitie, W. H. (1972). Children's conceptions of word boundaries in speech and print. *Journal of Educational Psychology, 63*(6), 551.

Huba, M. E., & Ramisetty-Mikler, S. (1995). The language skills and concepts of early and nonearly readers. *The Journal of Genetic Psychology, 156*(3), 313–331. https://doi.org/10.1080/00221325.1995.9914826.

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind—A critical examination of theories and evidence. *Language and Linguistics Compass, 8*(10), 401–427.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60–99.

Kaufman, A. S., Balgopal, R., Kaufrnan, J. C., & McLean, J. E. (1994). WISC-III Short Forms: Psychometric Properties vs. Clinical Relevance vs. Practical Utility. Paper for presentation at the annual meeting of the Mid-South Educational Research Association, Nashville, TN

Kavé, G. (2006). The development of naming and word fluency: Evidence from Hebrew-speaking children between ages 8 and 17. *Developmental Neuropsychology, 29*(3), 493–508. https://doi.org/10.1207/s15326942dn2903.

Kolinsky, R., Cary, L., & Morais, J. (1987). Awareness of words as phonological entities: The role of literacy. *Applied Psycholinguistics, 8*(3), 223–232. https://doi.org/10.1017/S0142716400000278.

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence. *Archives of Clinical Neuropsychology, 26*(7), 575–582. https://doi.org/10.1093/arclin/acr063.

Kurvers, J., Hout, R., & Vallen, T. (2007). Literacy and word boundaries. In *Low-Educated Second Language and Literacy Acquisition: Research, Policy and Practice: Proceedings of the Second Annual Forum* (pp. 45–64).

Kurvers, J., & Uri, H. (2006). Metalexical awareness: Development, methodology or written language? A cross-linguistic comparison. *Journal of Psycholinguistic Research, 35,* 353–367. https://doi.org/10.1007/s10936-006-9019-6.

Kurvers, J., Vallen, T., & Van Hout, R. (2006). Discovering features of language: Metalinguistic awareness of adult illiterates. *Proceedings of the Inaugural Symposium*, (pp. 69–88).

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' english. *Language Learning, 61*(June), 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x.

Levin, I., Ravid, D., Rapaport, S., & Nunes, T. (1999). Developing morphological awareness and learning to write: A two-way street. *Neuropsychology and Cognition, 17,* 77–104.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003.

Lieven, E. V. M., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language, 30,* 333–370.

Lieven, E., Pine, J., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24,* 187–220.

MacWhinney, B. (2005). Emergent fossilization. *Studies of Fossilization in Second Language Acquisition*, pp. 134–156.

Mandel, D. R., Jusczyk, P. W., & Nelson, D. G. K. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition, 53*(2), 155–180.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye- movements are modulated by reading ability—Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24,* 45–64. https://doi.org/10.1016/0010-0277(86)90004-1.

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading* (pp. 52–77). New York, NY: Guilford Press.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006a). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology, 42*(4), 643–655. https://doi.org/10.1037/0012-1649.42.4.643.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006b). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology, 42*(July 2015), 643–655. https://doi.org/10.1037/0012-1649.42.4.643.

Olson, D. R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition, 60*(1), 83–104.

Paul, J. Z., & Grüter, T. (2016). Blocking effects in the learning of Chinese classifiers. *Language Learning, 66*(4), 972–999.

Peters, A. M. (1983). *The units of language acquisition*. Cambridge, UK: Cambridge University Press.

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends In Cognitive Science, 11*(7), 274–279.

Ravid, D., & Malenky, A. (2001). Awareness of linear and nonlinear morphology in Hebrew: A developmental study. *First Language, 21*(61), 25–56.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of child language, 29*(2), 417–447.

Reis, A., & Castro-Caldas, A. (1997). Illiteracy: a cause for biased cognitive development. *Journal of the International Neuropsychological Society, 3*(5), 444–450.

Schneider, W. (2010). Memory development in childhood. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 347–376). Malden, MA: Wiley-Blackwell.

Siegelman, N., & Arnon, I. (2015). The advantage of starting big: Learning from unsegmented input facilitates mastery of grammatical gender in an artificial language. *Journal of Memory and Language, 85,* 60–75. https://doi.org/10.1016/j.jml.2015.07.003.

Singh, L., Reznick, S. J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental science, 15*(4), 482–495. https://doi.org/10.1111/j.1467-7687.2012.01141.x.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development, 49*(4), 1114–1128.

Sugiura, M. (2002). Collocational knowledge of L2 learners of English: A case study of Japanese learners. *Language and Computers, 38*(1), 303–323.

Tarone, E., & Bigelow, M. (2005). Impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics, 25*(1), 77–97.

Tarone, E., Bigelow, M., & Hansen, K. (2007). The impact of alphabetic print literacy level on oral second language acquisition. LESLLA Proceedings (pp 99–122).

Thompson-Schill, S., Ramscar, M., & Chrysikou, M. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science, 8,* 259–263.

Veldhuis, D., & Kurvers, J. (2012). Offline segmentation and online language processing units: The influence of literacy. *Written Language and Literacy, 15,* 165–184. https://doi.org/10.1075/wll.15.2.03vel.

Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., & Dehaene, S. (2013). Literacy acquisition reduces the influence of automatic holistic processing of faces and houses. *Neuroscience Letters, 554,* 105–109. https://doi.org/10.1016/j.neulet.2013.08.068.

Warwick, E. B., & Francis, M. (1983). The impact of reading on second language learning. *Reading Research Quarterly, 19*(1), 53–67.

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching, 32,* 213–231. https://doi.org/10.1017/S027226310000629X.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21,* 463–489. https://doi.org/10.1093/applin/21.4.463.

Yorio, C. (1989). *Idiomaticity as an indicator of second language proficiency* (pp. 55–72). Bilingualism across the Lifespan: Aspects of Acquisition, Maturity, and Loss.