Applied Psycholinguistics **38** (2017), 127–153
doi:10.1017/S0142716416000138

Reading between the words: The effect of literacy on second language lexical segmentation

NAOMI HAVRON and INBAL ARNON
Hebrew University of Jerusalem

Received: September 10, 2015 Accepted for publication: February 26, 2016

ADDRESS FOR CORRESPONDENCE
Naomi Havron, Department of Psychology, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel. E-mail: naomi.havron@mail.huji.ac.il

ABSTRACT
There is evidence that the ability to segment an utterance into words improves with literacy, yet previous research makes it difficult to disentangle the effect of literacy from that of age or cognitive abilities. We tested the hypothesis that literacy increases lexical segmentation in a second language in a unique sample of adult illiterates learning to read in their second language, controlling for cognitive abilities and using a task that taps language processing rather than only metalinguistic awareness. Participants' segmentation was correlated with first language reading at the beginning of an intensive literacy course for illiterate adults. At the end of the course, those learning to read for the first time benefited more in terms of their segmentation abilities. We discuss implications for models of second language learning.

In the past few decades, it has been shown that literacy is not only a representation of spoken language but also impacts the perception and processing of spoken language (Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998; Cheung & Chen, 2004; Levin, Ravid, & Rappaport, 1999; Olson, 1996). Contrary to the classical view that the writing system is simply a representation of spoken language, learning to read and write has far-reaching cognitive and linguistic consequences. One domain that may be affected by literacy is the ability to locate word boundaries in speech, that is, to perform *lexical segmentation*. Unlike spoken language, where word boundaries are not consistently marked by pauses, written language makes word boundaries salient by graphically representing gaps between words (at least in many orthographies). Learning to read may therefore make speakers more aware of word boundaries in spoken language.

Accordingly, there is evidence that literacy impacts lexical segmentation in both children and adults (Bialystok, 1986; Correa & Dockrell, 2007; Holden & MacGinitie, 1972; Kurvers, van Hout, & Vallen, 2009). Lexical segmentation in children has been studied extensively. Studies have found enhanced lexical segmentation among literate children compared to preliterate children (Bialystok, 1986; Chaney, 1989; Holden & McGintie, 1972; Kurvers et al., 2009). However, it is unclear whether this is the effect of age or of literacy, because literate children are typically older than preliterate ones (Karmiloff-Smith, Grant, Sims, Jones, & Cuckle, 1996). A recurrent finding in adults is that illiterate adults have more difficulty than literate adults segmenting utterances into word units, supporting the suggestion that literacy plays a role independent of that of age in the development of lexical segmentation (Gombert, 1994; Kurvers, 2015; Kurvers & Uri, 2006; Kurvers et al., 2009; Veldhuis & Kurvers, 2012). However, there are many challenges involved in comparing literate and illiterate adults and in assessing language processing in illiterate populations, which make it hard to isolate the effect of literacy on lexical segmentation.

In the present study, we will look at the effect of first language (L1) and second language (L2) literacy on the lexical segmentation abilities of illiterate and literate adults learning to read in their L2, using a novel task and controlling for cognitive and spoken language measures. This will enable us to isolate the effect of literacy on lexical segmentation from that of age or cognitive abilities, while tackling some of the challenges in testing illiterate populations. We will also expand previous findings by testing both participants who can read in their L1, but not in their L2 (L1 literate), and participants who cannot read in their L1 and are thus fully illiterate (L1 illiterate). This will allow us to compare the effect of learning to read for the first time to that of learning to read in in L2–L2 literacy is an additional writing system for some participants, while for others it is the first writing system they learn. This comparison, to our knowledge, has not been previously explored.

LITERACY AND LEXICAL SEGMENTATION IN ADULTS

Previous work investigated the lexical segmentation abilities of illiterate adults both in their native language and in their L2. In both languages, lexical segmentation is often assessed by asking participants to divide sentences they hear into parts, in order to see how many word units they produce. These lines of investigation have shown an effect of L1 literacy on L1 lexical segmentation, such that literate adults show enhanced segmentation in their native language compared to illiterate adults (Kurvers et al., 2009; Morais, Bertelson, Cary, & Alegría, 1986), as well as an effect of L1 literacy on L2 segmentation, such that native language literacy is associated with enhanced segmentation in an L2 (Gombert, 1994; Kurvers et al., 2009). That L1 literacy impacts L1 lexical segmentation suggests that literacy changes, rather than mirrors, perception and processing of spoken language. That it also impacts L2 processing opens up interesting questions about a possible relationship to L2 learning, by suggesting that knowing how to read in an L1 "infects" the processing of linguistic input in a new language. Knowing to read in an L1 may lead learners to treat words as more salient units in the L2 they are learning.

This possibility is especially interesting in light of recent suggestions that learning from multiword units may lead to better learning of certain grammatical relations, compared to learning from individual words (Arnon, 2010). This suggestion is supported by several different lines of research. First, there is growing evidence that multiword units may support L1 acquisition (e.g., Abbot-Smith & Tomasello, 2006; Arnon & Clark, 2011; Lieven, Pine, & Barnes, 1992; Lieven, Salomo, & Tomasello, 2009) and that they play a role in native language processing (e.g., Arnon & Priva, 2013; Arnon & Snider, 2010; Tremblay, Derwing, Libben, &

Westbury, 2011; Wray, 2002, 2009). Second, they are an important component of L2 fluency (Arnaud & Savignon, 1997; Wray, 1999, 2013), and there is increasing interest in targeting them in L2 instruction (Wood, 2002, 2009; Wray, 2000; Wray & Fitzpatrick, 2010). Third, artificial language learning studies have shown that adults learn grammatical dependencies better when multiword units were used as building blocks for learning (Arnon & Ramscar, 2012; Siegelman & Arnon, 2015). If literacy affects the size of the units learners learn from, and leads them to be more aware of word boundaries, this may negatively impact the way they learn certain aspects of grammar in the L2. We return to the implications of lexical segmentation for language learning in the Discussion section of this article.

As mentioned, the body of evidence from research on illiterate adults suggests that literacy plays a role in the development of lexical segmentation in both L1 and L2 and that this role is independent from that of age. However, the studies that explored language processing with illiterate adults face several challenges having to do with testing this unique population. These challenges can be divided into the sample-related and the task-related challenges. In terms of the sample, it is difficult to find suitable comparison groups for illiterate adults. Illiterate adults differ from literate ones in many respects, only one of which is their knowledge of print. Any literate comparison group will probably differ from an illiterate comparison group in their cognitive abilities, specifically nonverbal intelligence and short-term memory (Kosmidis, Zafiri, & Politimou, 2011; Olivers, Huettig, Singh, & Mishra, 2014; Ostrosky-Solís & Lozano, 2006). These differences may also influence performance on segmentation tasks, making it hard to attribute differences found between these groups to literacy alone.

Though many impressive attempts have been made to find literate adults who will be comparable to an illiterate sample (e.g., Castro-Caldas et al., 1998; Kosmidis et al., 2011), most studies settle for using a low-educated sample of literates from a background similar to the illiterate sample as the comparison group. Because none of the lexical segmentation studies carried out on illiterate adults collected measures of nonverbal intelligence and short-term memory, it is hard to ensure that the differences in lexical segmentation between the groups did not reflect differences in cognitive abilities, rather than literacy. The sample-related challenge can be tackled in at least two ways: (a) by collecting cognitive measures to statistically control for differences between participants, and (b) by testing the same participants before and after they learn to read, thereby reducing differences in cognitive abilities and other individual differences between participants. Research designs such as this are much needed in the field (Ardila et al., 2010).

In addition to the sample-related challenge, studies with illiterate adults face the challenge of finding suitable tasks for illiterate adults to perform. Some of the tasks that have been used with illiterate adults place a burden on short-term memory. This is troubling because there are probably initial differences between the illiterate group and the control group in short-term memory (Ardila, Rosselli, & Rosas, 1989; Kosmidis et al., 2011). For example, Gombert (1994) asked participants to repeat a sentence and instill pauses between the words. He comments that illiterate adults might have had trouble remembering the sentences, which would negatively impact their ability to repeat them. Other tasks may also be too complex for illiterate adults to understand. Both Morais et al. (1986) and Kurvers et al. (2009)

asked participants to segment a sentence into parts. Such instructions sound fairly simple, but they may be difficult for illiterate adults to understand. Kurvers et al. give an illustrative example of how illiterate participants carried out this task. In their example, a participant is shown to break the sentence into units based on its meaning, rather than into words or multiword units. They specifically state that the participant in the example did not understand what was asked of her.

Of course, an inability to understand the tasks may, in itself, point to an inability to perform them, but this does not have to be the case. It is possible that illiterate adults could not understand the task because they lack *metalinguistic awareness*, that is, the ability to consciously reflect on language; but they are nevertheless processing language in the same manner as literate adults. While studies on children's lexical segmentation abilities have undergone a transition since the first wave in the 1970s, with tasks becoming more and more implicit in an attempt to tap online language processing rather than metalinguistic knowledge (Karmiloff-Smith et al., 1996; Veldhuis, 2012; Veldhuis & Kurvers, 2012), all but one (Onderdelinden, van der Craats, & Kurvers, 2009) of the studies on illiterate adults used very explicit measures. The findings that literacy enhances metalinguistic awareness (Olson, 1996; Levin et al., 1999; Veldhuis & Kurvers, 2012) further highlight the need to use more implicit tasks when assessing lexical segmentation among illiterate adults.

To summarize, the above-mentioned challenges point to the need to better address the problem of comparison groups and control for the differences in cognitive abilities between literate and illiterate adults. There is also a need to develop simpler tasks that will place lower or no demands on short-term memory. Finally, there is a need to assess lexical segmentation in illiterate adults in a less explicit, more online manner, which may reflect language processing rather than only metalinguistic understanding. The present study will attempt to fill in these gaps in the existing literature.

THE PRESENT STUDY

To achieve the two goals described above, the present study will focus on the effect of L1 and L2 literacy on L2 lexical segmentation. First, we will ask whether L1 literacy still shows an effect on L2 processing after we address the above-mentioned methodological challenges and remove some of the confounds found in previous research. Second, we will ask whether learning to read and write in L2, in adulthood, affects L2 spoken language segmentation. We will do this by comparing participants' performance on a lexical segmentation task at the beginning and end of an L2 literacy course. We will also expand previous findings to Arabic (L1) and Hebrew (L2), two Semitic languages with writing systems different from the Latin script.

Our research participants were native Arabic speakers learning to read in their L2, Hebrew. Some were fully illiterate: they could not read in their native language or in the L2. Others were literate in their L1: they read comparatively well in their L1, but less so in the L2. Hebrew and Arabic are both Semitic languages with different writing systems, but a similarly rich synthetic morphology and a deep orthography, which consistently represents its morphology (Ravid, 2001). These particular typological characteristics of the language make reading in Hebrew (e.g., Deutsch, Frost, & Forster, 1998; Frost, Forster, & Deutsch, 1997), and specifically reading acquisition in Hebrew (e.g., Levin et al., 2001; Ravid, 2001; Ravid & Malenky, 2001; Share & Levin, 1999), an interesting case for psycholinguistic research. Despite this, lexical segmentation in these languages was not previously studied.

To achieve our first goal, which is to deal with the sample-related and task-related challenges described above, we employed four strategies. First, we collected measures of short-term memory, nonverbal intelligence, and L2 vocabulary to control for their effects on task performance. Second, we tested participants at the beginning and at the end of an intense, 3-month L2 literacy course. This enabled us to both test the effect of L1 literacy on L2 processing (at the beginning of the course) and observe enhancement in L2 spoken language segmentation ability following L2 literacy acquisition. Third, to limit the possibility that improvement in lexical segmentation did not result from the improvement in literacy, we also tested our participants at the beginning and end of the course on a nonlinguistic cognitive control task. We did this in order to measure the course's impact on an ability that should not be affected by literacy, and support the claim that it is the acquisition of literacy rather than a general improvement in executive functions following the course that explains any change in task performance. Fourth, to overcome shortcomings of previously used tasks, we used a task that was simple, did not burden short-term memory, and was relatively implicit. We will show that this task taps online language processing rather than only metalinguistic awareness by manipulating psycholinguistic properties of the stimuli and measuring the effect of this manipulation on performance.

The task we chose was a word order reversal task, based on Huttelocher's (1964) study of what she called the *word–phrase relationship*. The participants hear a pair of words and are asked to reverse their order, for example, hear "little boy" and produce "boy little." The task of reversing word order requires participants to recognize the boundary between the two words and then manipulate their order. Its advantages over previously used tasks are that it is quite simple and easy to understand and it hardly burdens short-term memory, because there are only two words to remember. This task also seems to require lower levels of metalinguistic abilities, and is more implicit than other tasks used, because it does not require conscious reflection on language and its constituent units. In line with psycholinguistic research showing that language processing is sensitive to the frequency of two-word and larger multiword units (Arnon & Snider, 2010; Jurafsky, Bell, Gregory, & Raymond, 2001), the current study will also manipulate the frequency of co-occurrence of two words (bigrams), arguing that more frequent word pairs should be harder for participants to reverse. Confirming this hypothesis will support the claim that the task taps online language processing, rather than only metalinguistic knowledge (see Materials section for more details on the task).

To summarize, this study will test several predictions. First, in line with psycholinguistic literature (Arnon & Snider, 2010; Bannard & Matthews, 2008; Jurafsky et al., 2001), we expect the frequency of word pairs to affect lexical segmentation, so that all participants (regardless of literacy status) will have more difficulty segmenting frequent word pairs. Such a finding would indicate the task does measure language processing rather than metalinguistic knowledge only. Second, we expect L1 literate adults to show enhanced lexical segmentation in their L2 compared with L1 illiterates (after controlling for word pair frequency, and participants' L2 vocabulary and cognitive abilities), illustrating the effect of L1 literacy on L2 processing. Third, we expect that both literacy and spoken language lexical segmentation performance will improve following the literacy course, further supporting the claim that literacy (rather than age) impacts lexical segmentation. Because some of our participants were acquiring a second literacy system, we will compare the improvement in lexical segmentation achieved by our more literate participants with that achieved by our less literate participants.

Though intuitively one may expect that improvement in reading will be positively correlated with improvement in segmentation, we do not predict such an effect for two reasons. First, there may be a fundamental difference between becoming a reader and becoming a better reader. In the first case, learners gain access to an additional knowledge about how language is represented. In the second, they are merely strengthening and entrenching the knowledge they have. Consequently, the effect of L2 literacy acquisition on lexical segmentation may be large in the first scenario, but negligible in the second. Second, the relation between segmentation enhancement and literacy improvement may even be negative for L1 literates. These learners may have an easier time learning to read in an L2 than their illiterate peers (i.e., show greater improvement on the literacy measures), because they can build on their understanding of L1 literacy to acquire their L2 literacy. However, they are predicted to show a smaller enhancement in lexical segmentation (because of their existing knowledge of segmentation from their L1 literacy). For these reasons, we predict that the improvement in lexical segmentation will be greater among those participants who were illiterate in both languages compared to those who were illiterate only in the L2. We expect no improvement on the nonlinguistic cognitive control task, which should be less susceptible to the effect of literacy.

METHOD

Participants

The participants were 39 male soldiers who were enrolled in an intensive Hebrew literacy course during their service in the Israeli army. All soldiers were native Arabic speakers from different communities and backgrounds (Bedouins, Muslims, and Christians). They were 18 to 28 years old (mean age = 21). All participants learned Hebrew as their L2 and were proficient speakers: Hebrew is the dominant language in Israel and is taught to Arab speakers from second grade. All participants have been living in a dominantly Hebrew environment since enlisting in the army (at least 1 year). All participants were deemed fit for army service and did not have any diagnosed learning or cognitive impairments.

In terms of literacy, participants were illiterate or poor readers in Hebrew (L2): they were attending the course because they could not read well in Hebrew. Their literacy level ranged from partial familiarity with the Hebrew letters to an ability to read a text at a first-grade level. It is important to note that illiteracy is generally uncommon in Israel (3.4% illiteracy for the general population), though the rates are higher for native Arabic speakers. In 2012, 4.8% of Arab men were considered illiterate, compared with only 1.6% of Jewish men. Similarly, 11.4% of Arab women in Israel were considered illiterate, compared with only 2.9% of Jewish women (Israeli Central Bureau of Statistics, 2014). Participants differed in their ability to read in their native language (Arabic), as well as in their L2 (Hebrew). About half of the participants could not read in any language (L1 or L2), and the other half had some literacy skills.

The army course teaches L2 literacy through a program developed by the army specifically for this population. The program aims to improve reading and writing skills, as well as Hebrew language skills, and general knowledge. Students receive both classroom oral instruction and individual tutoring in writing and reading short texts, and in reading comprehension skills. They study about 9 hr per day. The course takes place twice a year, with about 15 soldiers starting each cycle. Our sample is made up of soldiers from three consecutive cycles. All of the soldiers who were present on the day of testing participated in the study (except for 1 soldier who declined to participate). The study was approved by the Institutional Review Board of Hebrew University. The participants were told that the study goal is to evaluate what happens to adults as they learn to read. They were told that their participation is not mandatory and that they can stop at any stage. We emphasized that any data we collect would not be transferred to anyone, including their commanders. Consent was obtained verbally.

Materials

Four types of measures were collected: (a) measures of online word segmentation in the lexical segmentation task, (b) cognitive and linguistic measures used to control for differences in L2 linguistic abilities and cognitive abilities between the participants (nonverbal intelligence, short-term memory, L2 vocabulary, and a language background questionnaire), (c) literacy measures used to assess the literacy level in the two languages (L1 Arabic and L2 Hebrew), and (d) a measure of cognitive control used as a control for the before–after comparison.

Lexical segmentation. The lexical segmentation task was based on a study conducted by Huttenlocher (1964). In this study, participants heard word pairs and were asked to reverse their order (e.g., hear "little boy," produce "boy little"). This task requires participants to locate the two words and then manipulate their order: it is impossible to perform successfully without differentiating between the two words. Huttenlocher used the task with children 4.5–5 years old and showed that it was sensitive to the linguistic properties of the word pairs. In the original task, frequently encountered pairs were contrasted with listlike word pairs (e.g., child–lady) and with anomalous pairs that are not usually encountered (e.g., table–goes), rather than with simply infrequent pairs. There were considerably fewer errors in listlike and anomalous pairs compared to frequently encountered word pairs.

We expanded the original task by closely examining the effect of word co-occurrence frequency on performance, wanting to test the prediction that participants will have a harder time reversing higher frequency word pairs. We thus used

Table 1. *Properties of frequent and infrequent word pairs*

	Word Pairs			
	Frequent	Infrequent	t	p
Mean frequency				
Per million	50.36	1.4	5.49	<.000***
Word1	2239	2196	0.08	.94
Word2	888	558	1.73	.09
Number of syllables				
Word1	2.12	1.96	0.99	.33
Word2	1.56	1.56	0	1.0
Total word pair syllables	3.68	3.52	0.61	.55

***$p = .001$.

only grammatical word pairs, all appearing in everyday speech, differing only in frequency. We also controlled tightly for single word frequency. We used the MILA Tapuz People Forum Corpus (Itai & Wintner, 2008) to construct our items and calculate word and word pair frequencies. This is a collection of around 1.4 million words taken from Israeli online forums in Hebrew. We used this corpus because (a) there is at the moment no large enough corpus of spoken Hebrew (e.g., the Corpus of Spoken Israeli Hebrew only contains 93,000 tokens) and (b) it is closer to spoken Hebrew that other available written corpora.

We constructed two lists of 50 word pairs, one for each session. Each list contained equal numbers of frequent and infrequent word pairs. Frequently used word pairs appeared over 15 times per million in the MILA Tapuz corpus. Infrequent word pairs appeared under 3 times per million. The word pairs in both lists were very simple and familiar ones. Even the infrequent word pairs were made up of two simple words (e.g., "my father" is frequent but "her father" is infrequent, but all the single words are frequent and familiar). The two types of word pairs did not differ significantly in the frequency of the individual words making up the pair. They also did not differ significantly in the number of syllables of the first word, second word, or word pair. The stimuli properties are summarized in Table 1. All the word pairs were recorded by a female native speaker of Hebrew who was unaware of the study goals and who had been formerly employed as a radio anchor. The order of presentation of items was randomized for each participant.

As we were running the experiment, we realized that in Arabic (participants' native language) some word pairs constitute a single word, rather than two words. For example, the Arabic equivalent of the Hebrew word pair "aba sheli" (my father) is one word, "abuy." For the third course, we therefore removed or replaced these items with equivalent ones that are two words in both languages. In the new lists, which consisted of 40 items, the frequent and infrequent word pairs did not differ significantly in the frequency of the individual words making up the pair, or the number of syllables of the first word, second word, or word pair. We added whether a word pair was one or two words in Arabic to the statistical analyses (see Results section).

CODING. Responses were coded as correct if the word pair was successfully reversed and as incorrect if it was not. If participants misheard the word pair (e.g., heard: "little boy," produced "*toy* little"), the response was excluded from the analysis. However, if participants produced words that were not phonetically similar (hear "little boy," produced "apple little"), the responses were coded as incorrect.

Cognitive measures. Participants completed two cognitive abilities measures: a measure of nonverbal intelligence and a measure of short-term memory. Nonverbal intelligence was assessed using the adult version of the block design subset of the Wechsler Adult Intelligence Scale—Fourth Edition (Wechsler, 2008). In this task, participants are asked to use colored blocks to match a pattern displayed to them. Their performance was scored based on the norms of the block design subset of the Wechsler Adult Intelligence Scale—Fourth Edition. Short-term memory was evaluated using the forward digit span task in Arabic (participants' native language). In this task, participants are asked to repeat a string of digits they hear, with a growing number of digits (Wechsler, 2008). We used only the digits one to five, because it was proposed that adult illiterates' worse performance on this task is partly due to them being less familiar with numbers higher than five (Petersson, Reis, & Ingvar, 2001). Participants' score on this task was the number of correctly recalled strings.

L2 linguistic measures. Our vocabulary measures were based on the Multilingual Naming Test, which is a validated scale for measuring vocabulary in an English–Spanish bilingual population (Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012), and on the Gitit Kave battery (Kavé, 2005), which is a validated Hebrew vocabulary measure. We used the standard administration procedures for these tasks. Performance was coded as percentage of correct responses. We also collected self-reported ratings of participants' ability to speak and understand Hebrew using the Language Experience and Proficiency Questionnaire, a validated and reliable scale for bilinguals and multilingual's self-rated proficiency (Marian, Blumenfeld, & Kaushanskaya, 2007). We only used items that were relevant for illiterate participants. Participants rated their proficiency of a scale of 1–7.

Literacy assessment. We assessed participants' ability to read in both Arabic (L1) and Hebrew (L2) in two ways: self-reported ratings (from the Language Experience and Proficiency Questionnaire) and rated paragraph reading. Participants rated their ability to read and write in Hebrew and Arabic on a scale from 1 to 7. In addition, we asked participants to read a short paragraph in the two languages. In both languages, the paragraph was at the level of first-grade readers (downloaded from an educational website). The paragraph reading was recorded and later rated by native speakers of Hebrew and Arabic who were blind to the study goals and had not met the participants. We used a scale developed for assessing literacy in illiterate populations (Tarone, Bigelow, & Hansen, 2007). Reading was rated on a scale of 0–6 for two measures: fluency (0–3) and confidence (0–3). To make it comparable to the self-rating scale, 1 point was added to the aggregate score, to create a scale of 1–7.

Cognitive control measure. We assessed participants' nonlinguistic cognitive control using a validated and reliable spatial Stroop task, the *Simon task* (Bialystok, Craik, Green, & Gollan, 2009; Lu & Proctor, 1995; Mor, 2012). In this task, a dot appeared on a screen, sometimes on the left side of the screen and other times on the right side of the screen. Participants were asked to press the left key of a response box if the dot appeared on the left and the right key if it appeared on the right. After performing a few practice trials, they were told that now an arrow will appear in the middle of the screen before the dot appears on either side. They were told to ignore the direction the arrow is pointing at, and only take notice of the location of the dot. Half of the trials were congruent: the arrow pointed in the same direction as the location of the dot. The other half were incongruent: the arrow pointed in the opposite direction. Participants performed a practice block before they began the task, during which they received feedback for their response (we used a smiley face if they succeeded and a red X if they failed).

Participants' reaction time for each trial was recorded. We calculated two cognitive control scores: a reaction time difference score and an accuracy difference score. For the reaction time difference score, we subtracted participants' mean reaction time on accurate congruent trials from their mean reaction time on accurate incongruent trials. For the accuracy difference score, we subtracted their accuracy on incongruent trials from their accuracy in congruent trials. Inhibiting the automatic shift of attention to the direction of the arrow requires cognitive control, and therefore the smaller the difference between accuracy and speed in congruent and incongruent trials, the better cognitive control the participant is estimated to have.

Procedure

Participants were tested twice: once in the second week of the course and once in the penultimate week of the course. They were tested in a quiet room using a laptop. We used the experimental software E-Prime (Schneider, Eschman, & Zuccolotto, 2012) to run the different tasks and record responses. All sessions were also recorded using a portable recorder. During the first session, participants completed the tasks in the following order: (a) the lexical segmentation task, (b) L2 vocabulary assessment, (c) nonverbal intelligence task, (d) short-term memory task (digit span), (e) cognitive control task, (f) L1 and L2 literacy assessment, and (g) language background questionnaire. In the first course only ($N = 16$), participants completed the cognitive measures on a separate testing session due to the course's time constraints. In addition, the third course did not complete the cognitive control task due to the limited testing time we were allocated by the course managers. Because we wanted to examine participants' improvement on literacy, lexical segmentation, and cognitive control following the course, participants completed only these tasks during the second session. At the end of the course, we thanked participants for their participation by organizing a small party at the army base.

RESULTS

We collected data from 39 participants who took part in three consecutive cycles of the literacy course (participants who performed the course twice were not tested twice). The data from 6 participants was removed from the analysis due to the following reasons: 2 did not complete the cognitive and linguistic measures, 1 dropped out of the course after the first session, and 3 did not manage to complete the segmentation task (responded to under 50% of the trials). Only 17 participants out of the 39 were available for testing at the end of the course because the entire third course was stopped in the middle due to the geopolitical events of summer 2014 in the Middle East (i.e., we could not test any of the third course participants again at the end of the course). In total, the data from 33 participants was analyzed at the start of the course, and the data from 17 participants was analyzed for the before–after comparison. It is important to note that there were no apparent differences between the participants who completed both testing sessions and those who only completed the first: the reduction in sample size in the second testing was not because certain participants dropped out (which could have biased the sample), but because the whole course was stopped. To further ensure the similarity of samples between courses, we ran three analyses of variance to compare L1 reading, L2 reading, and task performance between the courses in the first session (where we had participants from all three courses): there were no significant differences between the courses in L1 reading: $F\,(30, 2) = 0.68$, $p = .52$; L2 reading: $F\,(30, 2) = 0.15$, $p = .87$; segmentation task performance: $F\,(30, 2) = 0.67$, $p = .52$.

Literacy

We collected two literacy measures for each language: a self-report measure and a paragraph reading measure (see Materials section for coding details). The interjudge reliability of the reading measure was high ($\kappa = 0.89$). The two measures were highly correlated in the L1 and the L2: L1: $r\,(31) = .82$, $p < .0001$; L2: $r\,(31) = .81$, $p < .0001$, which points to the convergent validity of the measures. We created two aggregated scores, one for each language, consisting of the average between self-report and paragraph reading score.

Correlations between L1 literacy and other measures

As discussed in the literature review, illiterate populations usually differ from their literate or low-literate comparison groups in terms of their cognitive abilities. In our case, it was also possible that they differ in their spoken L2 skills, which could affect their performance on a spoken L2 segmentation task. To examine whether L1 literacy was related to the cognitive and linguistic measures, we calculated correlations between these measures and L1 literacy (see Table 2). Because L1 literacy scores were quite evenly distributed across the literacy continuum, we chose not to split our participants into two groups, an illiterate group and a literate group, based on the median of the distribution, but rather to look at L1 literacy's correlations with cognitive and linguistic variables across all levels of literacy (see MacCallum, Zhang, Preacher, & Rucker, 2002).

As expected, L1 literacy was positively correlated with nonverbal intelligence, $r\,(31) = .36$, $p = .04$, and short-term memory, $r\,(31) = .35$, $p = .04$. L2 reading scores were also correlated with L1 reading scores, $r\,(31) = .47$, $p = .005$. It is important that there was no correlation between L2 vocabulary and L1 reading,

Table 2. *Means, standard deviations, and correlations between the cognitive and linguistic measures and L1 literacy score*

Variable	*M*	*SD*	r^a	*t*	*p*
L2 reading	3.18	1.78	.47	3	.005
L2 vocabulary	33.02	8.47	.00	0	.99
Block design (nonverbal intelligence)	6.42	2.17	.36	2.16	.04
Digit span (short-term memory)	6.44	1.64	.35	2.1	.04
L2 comprehension self-report (1–7)	5.45	1.42	−.46	−2.86	.01
L2 speaking self-report (1–7)	4.79	1.30	−.27	−1.53	.14
Cognitive control task accuracy	0.06	0.04	.09	0.39	.7
Cognitive control task reaction time	25.19	59.99	−.28	−1.22	.24

Note: L1, First language; L2, second language.
[a]$n = 33$.

$r(31) = .0001, p = .99$, or between self-reported L2 speaking skills and L1 reading, r (31) = –.27, p = .14. In other words, being illiterate in the L1 did not impede learning to speak the L2. It is interesting that there was even a significant negative correlation between L2 comprehension (from self-report) and L1 literacy, r (31) = –.46, p = .01. As mentioned, we were able to collect cognitive control measures only for 21 participants. For these participants, there was no significant correlation between L1 literacy and the cognitive control measures: accuracy score: r (19) = .09, p = .7; reaction time score: r (19) = –.28, p = .24.

To summarize, while cognitive abilities and L2 literacy were related to L1 literacy, as was found in previous studies, spoken L2 measures were not. All measures will be controlled for in the following analyses.

General performance on lexical segmentation task

The percentage of correct responses at the beginning of the L1 literacy course ranged from 54% to 100%, with a mean score of 86%. This shows that despite using Hebrew in their daily lives, our participants were not at ceiling on our lexical segmentation task. For comparison, in a pilot study, a group of native Hebrew-speaking students who completed the same task did perform at ceiling (n = 16, mean performance 99% correct). However, the participants were able to perform correctly on most trials, indicating that they understood what was required of them. Because this task has not been used before with illiterate populations, we wanted to assess its reliability. The test–retest reliability of the reversal task was high, r (16) = .87, p < .001. The Cronbach α scores were computed separately for the first session, second session, and third course's first session (because slightly different items were used, see Materials section). All coefficients were high (first session, 50 items, α = 0.92; second session, 50 items, α = 0.88; third course, 40 items, α = 0.89), indicating that the task's internal reliability was high.

In addition to correct responses, we also looked at error patterns. Incorrect responses included repetition of the word pair, content errors (e.g., hearing "my name," producing "Ahmed"), and reversing meaning (e.g., hearing "little boy," producing "big boy"). Such errors do not show evidence of lexical segmentation, and they accounted for 59% of incorrect responses. Other incorrect responses can be said to exhibit some level of segmentation, but not the full ability to recognize and manipulate the two words. These included repetition of only one of the words (e.g., hearing "little boy," producing "little," "little boy boy," or "little boy little"), changing one of the words (e.g., hearing "little boy," producing "small boy"), and changing or adding function words (e.g., hearing "which one," producing "what one"). Partial segmentation responses such as these accounted for 41% of incorrect responses. We further examine the distinction between these two types of errors in the next section.

The effect of L1 literacy and word pair frequency on lexical segmentation

Next, we tested our two main predictions: (a) that L1 literacy predicts lexical segmentation in an L2, and (b) that more frequent word pairs will be harder to reverse, suggesting the task taps into online processing. We used a mixed-effects logistic regression model to test both predictions, focusing on task performance at the beginning of the course (we used the glmer function in R program; Bates, Maechler, Bolker, & Walker, 2015). A logistic regression provides the probability of a correct response at each trial (in logit space) depending on the value of the independent variables. While only 33 participants were tested in the first session, because we are predicting the result at each trial, we have 1,500 observations in the model. This kind of analysis is increasingly used in the psycholinguistic literature, among other things, because of the advantages it has compared to standard analyses of variance (see Baayen, Davidson, & Bates, 2008; Jaeger, 2008).

We had logged word pair frequency and centered L1 literacy as our variables of interest. We had several additional fixed effects in the model: L2 literacy as the beginning of the course; short-term memory and nonverbal intelligence scores (both centered) to control for cognitive differences. We also added L2 vocabulary scores (also centered) to control for differences in spoken L2 knowledge, which may well affect performance on spoken L2 lexical segmentation. As was mentioned in the Method section, we learned that some of the word pairs ($n = 11$ in the first session) constitute one word in Arabic, the participants' L1. We added this variable to our model to control for any effect it may have on participants' performance (because it may be harder for participants to reverse word pairs that are one word in their native language). The model included random intercepts for participant and item. The collinearity of the model was low (maximal variance inflation factor = 1.47).

The full model is highly significant (log likelihood $\chi^2 = 47.28$, $df = 6$, $p < .001$). As predicted, the effect of word pair frequency on performance was significant: More frequent word pairs were harder to reverse ($\beta = -0.23$, $SE = 0.07$, $p = .001$, model comparison $p = .002$). It is important that the effect of L1 literacy was also significant: participants who were better readers in the L1 were more likely to correctly reverse the word pairs ($\beta = 0.99$, $SE = 0.27$, $p = .0003$, model comparison $p = .001$). This held after controlling for L2 vocabulary and the two cognitive measures. L2 vocabulary also predicted performance with participants

Table 3. *Logistic regression of frequency, L1 literacy, and control variables*

Variable	*B*	*SE*	*Z*	*p*	95% CI
Intercept	3.1	0.28	10.93	<.001	(2.54, 3.65)
Word pair frequency (log)	**−0.23**	**0.07**	**−3.18**	**.001**	**(−0.23, −0.09)**
Literacy in L1 (centered)	**0.36**	**0.09**	**4.09**	**<.001**	**(0.19, 0.53)**
L2 vocabulary (centered)	0.04	0.02	1.8	.07	(−0.003, 0.07)
Nonverbal intelligence (centered)	0.04	0.09	0.51	.61	(−0.13, 0.21)
Short-term memory (centered)	0.22	0.12	1.87	.06	(−0.01, 0.46)
Item is one word in L1	−1.07	0.35	−3.06	.002	(−1.75, −0.38)

Note: The variables of interest are in bold. L1, First language; 95% CI, 95% confidence interval; L2, second language.

with larger L2 vocabularies more likely to reverse the word pair correctly (β = 0.04, *SE* = 0.02, *p* = .05, model comparison *p* = .06). The effect of short-term memory was also significant (β = 0.27, *SE* = 0.12, *p* = .02, model comparison *p* = .03), such that better short-term memory was associated with higher probability of a correct response. It is interesting that word pairs that constitute one word in participants' L1 were harder to reverse compared to items that are two words in both languages (β = –1.1, *SE* = 0.35, *p* = .003, model comparison *p* = .003). None of the other factors was significant (see Table 3).

We conducted an additional analysis to put our predictions to a stricter test, and to take into consideration that some errors (as described above) can be thought of as displaying partial segmentation. We wanted to ensure that better L1 literacy and lower frequency still predict better performance on the task when taking the difference between the two types of errors into account. In this analysis, we treated performance on each trial not as a binary variable (correct vs. incorrect) but as an ordinal one ("no segmentation," "partial segmentation," or "full segmentation"). We used a cumulative link model using the R command clm from the package "ordinal" (Christensen, 2015). Cumulative link models are ordinal regression models, which estimate the probability that a response will be included in each response category depending on the independent variables. Our cumulative link model used the exact same fixed factors as the binary logistic regression we conducted, but with the ordinal dependent variable.

As in the previous analysis, in this model the frequency and L1 literacy were both significant predictors of lexical segmentation: higher frequency was associated with worse performance, while higher literacy was associated with better performance (see full model in Table 4). That is, even when we take into account partial segmentation, literacy and frequency are significant predictors of success on the task.[1]

Separating the effect of L1 literacy from that of L2 literacy. To further ensure that the effect of L1 literacy is not an effect of L2 literacy in disguise, we ran an additional analysis only on those participants who could not read in the L2 (scored under 4 on the aggregate L2 literacy scale). This left us with 19 participants. We reran the logistic regression with the logged word pair frequency and logged L1

Table 4. *Logistic regression of frequency, L1 literacy, and control variables, using only participants who could not read in the L2 (scores < 4)*

Variable	*B*	*SE*	*z*	*p*	95% CI
Intercept	2.54	0.29	8.85	<.001	(1.97, 3.10)
Word pair frequency (log)	**−0.26**	**0.07**	**−3.57**	**.0004**	**(−0.41, −0.12)**
Literacy in L1 (centered)	**0.25**	**0.08**	**2.87**	**.004**	**(0.08, 0.42)**
L2 vocabulary (centered)	0.03	0.02	1.53	.13	(−0.008, 0.06)
Nonverbal intelligence (centered)	0.02	0.09	0.22	.82	(−0.15, 0.19)
Short-term memory (centered)	0.08	0.11	0.66	.51	(−0.15, 0.30)
Item is one word in L1	−0.99	0.36	−2.74	.006	(−1.70, −0.28)

Note: The variables of interest are in bold. L1, First language; L2, second language; 95% CI, 95% confidence interval.

literacy as our variables of interest, and short-term memory, nonverbal intelligence, L2 vocabulary, and whether the word was one word in Arabic as fixed effects. The model included random intercepts for participant and item. The collinearity of the model was low (maximal variance inflation factor = 1.38).

Despite the loss of more than a third of our data, the results were very similar to those we received when the regression was run with all 33 participants. The effect of word pair frequency was significant: more frequent word pairs were harder to reverse ($\beta = -0.26$, $SE = 0.07$, $p = .0003$, model comparison $p = .0005$). The effect of L1 literacy was significant: higher L1 literacy scores were associated with better performance at lexical segmentation in L2 ($\beta = 0.7$, $SE = 0.26$, $p = .008$, model comparison $p = .01$). The effect of whether the word pair was one word in Arabic was also still significant: word pairs that constitute one word in Arabic were harder to reverse ($\beta = -0.97$, $SE = 0.36$, $p = .008$, model comparison $p = .01$). None of the other factors was significant (see Table 5).

To summarize, our variables of interest, L1 literacy and frequency of word pair, significantly affected performance on the task even when looking only at participants who could not read in their L2, and thus reducing our sample size to 19 participants.[2]

Effect of L2 literacy acquisition on L2 lexical segmentation

Finally, we looked at the effect the L2 literacy course had on task performance. Unfortunately, as mentioned above, geopolitical events allowed us access to only 17 participants for the before–after comparison, but we report the finding for these 17.

To examine the effect of literacy acquisition on lexical segmentation, we wanted to make sure L2 literacy improved after the intensive course. If it improved, we wanted to also see whether task performance improved. To support the claim that the improvement in segmentation is due to the improvement in literacy, we sought to examine whether improvement was larger the less literate a participant was at the start of the course. We reasoned that participants who were being introduced to the graphical representation of word boundaries for the first time should show

Table 5. *Cumulative link model of frequency, L1 literacy, and control variables*

Variable	*B*	*SE*	*z*	*p*	95% CI
Word pair frequency (log)	−0.22	0.04	−5.26	<.001	(−0.30, −0.14)
Literacy in L1 (centered)	0.3	0.04	7	<.001	(−1.27, −0.58)
L2 vocabulary (centered)	0.03	0.01	3.4	<.001	(0.21, 0.38)
Nonverbal intelligence (centered)	0.07	0.04	1.65	.1	(−0.01, 0.15)
Short-term memory (centered)	0.19	0.05	3.52	<.001	(0.01, 0.05)
Item is one word in L1	−0.92	0.18	−5.23	<.001	(0.09, 0.30)

Threshold Coefficients[a]

Threshold (Intercepts)	*B*	*SE*	*Z*
No segmentation | partial segmentation	−3.67	0.18	−20.84
Partial segmentation | full segmentation	−2.67	0.15	−17.35

Note: The variables of interest are in bold. L1, First language; 95% CI, 95% confidence interval; L2, second language.
[a]Threshold coefficients are similar to an intercept in a linear regression, but indicate the cutoff points between the different ordinal categories in the latent (hypothetically continuous) dependent variable.

greater enhancement in lexical segmentation. Finally, to support the claim that the improvement in the lexical segmentation task is due to the improvement in literacy and not in other cognitive factors that might have benefited from this intensive course, we examine whether participants' nonlinguistic cognitive control also improved following the course.

We ran a paired-sample *t* test, looking at L2 literacy scores at the beginning and end of the course. As could be expected, L2 literacy improved significantly from the beginning to the end of the course ($M = 3.18$, $SD = 1.78$, at the beginning, $M = 4.76$, $SD = 1.64$, at the end of the course), $t\ (16) = 5.13$, $p < .0001$, showing that the course was overall successful in improving our adult participants' literacy skills. However, not all participants improved, and not all of them emerged from the course as literates. Three participants did not improve at L2 reading at all. Moreover, the great bulk of participants (8 out of 17) are now focused between scores 3 and 4, which is still quite low.

Turning to task performance, we ran a paired-sample *t* test, comparing segmentation task performance at the beginning and end of the course. As predicted, performance on the segmentation task improved significantly from the beginning to the end of the course ($M = 81\%$, $SD = 14\%$, at the beginning of the course, $M = 86\%$, $SD = 12\%$, at the end of the course), $t\ (16) = 2.87$, $p = .006$.

Next, we examined whether L1 literacy had an effect on the amount of improvement in task performance. We calculated a simple linear regression with improvement in our segmentation task as the dependent variable (performance at second session minus performance at the first session) and L1 literacy score at the start of the course as our independent variable. Literacy scores at the start of the course significantly predicted task improvement, indicating that learning to read in the L2 was less beneficial for lexical segmentation the more literate the participant was in his L1 ($\beta = -6.11$, $SE = 2.6$, $p < .005$).

Finally, we looked at participants' improvement on the nonlinguistic cognitive control task. One participant did not complete this task, leaving us with 16 participants for the before–after comparison. We ran two paired-sample *t* tests, comparing both reaction time difference scores and accuracy difference scores at the beginning and end of the course. As predicted, unlike the improvement in the segmentation score, performance on both measures (accuracy and reaction time) of the nonlinguistic cognitive control task did not improve from the beginning to the end of the course. The mean reaction time difference score was 61 ms ($SD = 23$) at the beginning of the course and 54 ms ($SD = 34$) at the end, $t(15) = 0.71$, $p = .48$. The mean accuracy difference score was 0.04 ($SD = 0.07$) at the beginning of the course and 0.04 ($SD = 0.05$) at its end, $t(15) = -0.03$, $p = .98$.

To summarize, literacy scores improved after the course; however, not all participants could read in their L2 at the end of the course. Task performance also significantly improved, and this improvement was smaller the more L1 literate the participant was, as could be expected if learning to read in an L2 only had an additive effect for persons who could already read in their L1. Whereas both reading and lexical segmentation significantly improved, nonlinguistic cognitive control abilities, as measured by a spatial Stroop task, did not improve following the course.

DISCUSSION

The main goals of this study were to examine the effect of L1 and L2 literacy on spoken L2 lexical segmentation, and to do this in a way that addresses some of the challenges arising from studying illiterate populations. These challenges include controlling for inherent differences between illiterate adults and their literate comparison group, and finding a task that will be suitable to use with an illiterate population (i.e., a task that is simple, does not burden short-term memory, and is relatively implicit, so that it will reflect not only metalinguistic awareness but also more on-line language processing). We attempted to accomplish this in four ways: by assessing participants' cognitive abilities and adding them as control variables to our analyses; by comparing participants' lexical segmentation abilities at the beginning and end of an L2 literacy course; by choosing a task that was never used with illiterate population and is simple, places little burden on short-term memory, and is relatively implicit; and by demonstrating that our task taps into language processing and not only metalinguistic awareness by manipulating psycholinguistic properties of the stimuli. We predicted that both the L1 and the L2 literacy skills of adult learners will be related to increased lexical segmentation in spoken L2, a finding that would support the idea that the effect of literacy on spoken language lexical segmentation is independent of that of age.

Our results show a clear effect of L1 literacy on spoken L2 lexical segmentation: higher L1 literacy was associated with better task performance, even after controlling for individual differences in nonverbal intelligence, short-term memory, L2 vocabulary, and stimuli linguistic properties. The results also show an enhancement of segmentation ability after an L2 literacy course, indicating that learning to read in an L2 effects lexical segmentation in this language. These findings are in line with previous studies on illiterate adults showing that L1 and L2 literacy impacts L2 processing (Gombert, 1994; Kurvers et al., 2009). We extend the current literature by showing that the results hold after controlling for cognitive and linguistic abilities, and when using a more implicit task, and by showing that improvement was greater for those learning to read for the first time. Our improved task characteristics and the controls we added make it easier to attribute the difference we found between literate and illiterate adults to real differences in processing rather than to an inability to understand the task, remember the items, or make use of metalinguistic understanding. Our before–after design, combined with our control of cognitive measures and choice of task, therefore allows us to cautiously infer that the relationship between literacy and spoken language segmentation is a causal one: literacy seems to enhance lexical segmentation.

Our results are also relevant for the debate on the relative role of age and literacy in lexical segmentation. They lend strong support to the claim that literacy plays a role independent from that of age in increasing sensitivity to word boundaries: adults who could not read had greater difficulty segmenting word pairs, and lexical segmentation in adults was enhanced following a literacy course. This finding joins the growing literature on the effect of literacy on linguistic processing (Levin et al., 1999; Ravid & Malenky, 2001; Ravid & Tolchinsky, 2002).

Returning to the effect of L2 literacy acquisition on lexical segmentation, out of our 33 participants, we were only able to test 17 participants at the end of the course (the course was stopped in the middle because of geopolitical reasons). This remaining group of participants improved in their literacy score as a whole, though not all participants improved and many still struggled with reading. Participants also improved in lexical segmentation: they showed better task performance at the end of the course. As hypothesized, this improvement was greater for participants who were less literate in their L1, suggesting that L2 literacy acquisition was less beneficial for participants who were already literate in their L1. This result highlights the difference between being exposed to the graphical representations of words for the first time and learning to read in an L2. However, it is possible that this differential improvement was partly due to a ceiling effect because three of our more literate participants received very high scores (>95%) in the first testing, which did not leave them much room for improvement.

There are two additional alternative explanations that could account for the improvement in task performance. First, it is possible that the improvement was the result of a test–retest effect: that participants had already performed the task in the first session caused their better performance in the second session. However, this would not explain the differential improvement of the L1 illiterate participants and seems less likely because we used different stimuli in the two sessions and because the two sessions were separated by an interval of 2 months. Second, there is a possibility that the improvement in task performance might have stemmed not from improvement in literacy but from a more general effect the course had on participants. The current study's context, like that of any educational setting, makes it hard to separate the effect of literacy acquisition from that of schooling (Kosmidis, Tsapkini, & Folia, 2006), because during the course participants gained additional L2 knowledge and academic skills. In other words, as a by-product of learning to read, our participants might have also learned how to learn, and how to better focus on the task. However, participants' cognitive control abilities, which were also tested at the start and end of the course, did not improve following the course. Though this in no way precludes the possibility that improvement was due to other factors rather than literacy acquisition, it provides some converging evidence in support for this claim.

The effect of multiword frequency on processing

Our findings also have implications for the psycholinguistic literature on the processing of multiword sequences. There is growing evidence that the frequency of multiword units affects language processing, much like the frequency of single words. Frequently used multiword units are processed faster, produced faster, and their phonetic duration is shorter than less frequently used units (Arnon & Priva, 2013, 2014; Arnon & Snider, 2010; Bannard & Matthews, 2008; Tremblay et al., 2011), whether because they are stored as wholes (Bannard & Matthews, 2008) or because connections between their components strengthen with increased use (Arnon & Priva, 2013, 2014). Given these findings, we predicted that more frequent word pairs would be harder to reverse, thus validating that our lexical segmentation task was measuring language processing rather than only metalinguistic awareness. This prediction was confirmed: the more frequent the word pair, the harder it was for our participants to reverse the order of the two words. In other words, participants found it hard to override frequency induced word order preferences, especially for high-frequency word pairs (e.g., Siyanova-Chanturia, Conklin, & van Heuven, 2011). These findings extend previous findings on the role of multiword information in processing by demonstrating such effects also exist in illiterate adults, and by using a novel behavioral measure of language processing: the word order reversal task.

The finding that both literacy and frequency play a role in participants' success in reversing the word pairs also relates to the large literature on formulaic language (e.g., Wray, 2002; Kapatsinski & Radicke, 2009). Wray (2002) suggested that learning to read and write is instrumental in segmenting previously stored formulaic language, but that the segmented words are stored alongside the original formulaic sequences. This may explain why some of the participants' incorrect responses revealed what can be regarded as partial access to word units. These participants could have had access to the words making up the word pair, but because the word pair is also stored, failed to override the automatic access to them. Alternatively, the responses showing incomplete knowledge of the units making up the word pair can be viewed as a stage in the development of full segmentation, an explanation that could be tested by examining lexical segmentation longitudinally, as participants become fluent readers.

Extending previous findings to Hebrew as an L2

Our findings also extend previous findings by examining the spoken lexical segmentation of Hebrew (L2) in a sample of Arabic native speakers. Hebrew is a

Semitic language, which is much more synthetic than the Indo-European L2s previously studied (Dutch: Kurvers et al., 2009; Onderdelinden et al., 2009; French: Gombert, 1994 [3] ; Portuguese: Morais et al., 1986). Synthetic languages tend to express more morphemes per word than analytic languages, meaning that many Hebrew or Arabic words would be translated into several words in more analytic languages. In addition, the two different Arabic and Hebrew scripts are not strictly alphabetic, but rather Abjad scripts, which means that they contain symbols for consonants, but vowels are only optionally marked. However, the languages previously studied and the languages explored in the current study do share an important feature: they all place spaces between words, graphically marking word boundaries. This is not true for all language. In logographic scripts, such as Chinese Hanzi and Japanese Kanji, there are no conventions for determining word boundaries (Bassetti, 2007). It is consistently found that participants who are literate in these languages have more difficulty segmenting a sentence into words than do participants who are literate in an alphabetic script (Bassetti, 2007; Miller, Frosch, Kelly, & Zhang, 2001; Tsai, McConkie, & Zheng, 1998). That Hebrew and Arabic literacy showed the same effect on lexical segmentation as alphabetic scripts in analytic Indo-European languages suggests that it is the graphic representation of word boundaries, rather than any other morphological or graphical characteristic of these languages, that causes lexical segmentation to be enhanced.

Related to this, an interesting finding was that whether a word pair was one or two words in participants' native language had an effect on their performance in the L2. Word pairs that constitute one word in participants' L1 were harder for participants to reverse. The difficulty participants had in reversing the order of the words in word pairs that constituted only one word in their native language can be seen as a form of *lexical transfer*, a process by which a person's knowledge of one language influences their processing of another language (Jarvis, 2009). In our case, participants can be said to have experienced *processing interference*, which can arise when words in one language are activated though the person is attempting to use (in our case to manipulate) another language (Jarvis, 2009).

Implications for models of L2 learning

We believe that our findings may hold important implications for the study of differences between child and adult language learning. Despite having a larger memory capacity, more knowledge and experience with language, and sharper analytic skills, adults learning a new language rarely reach native proficiency, unlike young children (DeKeyser, 2000; DeKeyser, Alfi-Shabtay, & Ravid, 2010; Johnson & Newport, 1989; Long, 1990; Vanlancker–Siditis, 2003). Many factors may underlie these differences, including cognitive abilities, neurological makeup, the linguistic input children and adults receive from the environment, and also the different linguistic units they learn from (Arnon, 2010; Elman, 1993; Hoff, 2006; Meisel, 2009; Newport, 1990; Wong-Fillmore, 1991). In the present study, we construe literacy as an additional factor that can affect L2 learning. We found that knowing to read in the L1 was associated with improved lexical segmentation in an L2. This could mean that L1 literacy is affecting the processing of a new L2 input, while it is being learned. Put differently, the input learners learn from may be more easily broken down into word units by literate adults (and children) compared to illiterate adults and preliterate children. This would cause them to learn from smaller units, possibly affecting their learning outcomes.

As mentioned in the introductory section, there is growing interest in the role of multiword units in language learning. Multiword units are seen as important building blocks for L1 learning in usage-based approaches (e.g., Abbot-Smith & Tomasello, 2006; Arnon & Christiansen, 2014; Bod, 2006), and they are predicted to facilitate grammar learning as well as the use of collocations and formulaic language (Arnon, 2010; McCauely & Christiansen, 2011, 2014; Wray, 2002). In line with these predictions, learning from multiword units rather than from separate words was found to improve some aspects of grammar learning in artificial languages (Arnon, 2010; Arnon & Ramscar, 2012; Siegelman & Arnon, 2015; Valian & Levit, 1996), as well as contribute to discourse level production for beginner L2 learners (Taguchi, 2007).

In this context, it is intriguing that we did not find any correlation between L1 literacy and L2 vocabulary or (self-reported) speaking skills. There was even a *negative* correlation between (self-reported) L2 comprehension and literacy, by which higher levels of literacy were actually correlated with lower levels of L2 comprehension. While it is beyond the scope of the present study to evaluate the impact of literacy on L2 proficiency, it is thought provoking that L1 literacy did not seem in any way to aid L2 learning, despite its positive effect on other cognitive and linguistic abilities (e.g, Castro-Caldas et al., 1998; Cheung & Chen, 2004; Levin et al., 1999). Pettitt and Tarone's (2015) study of literacy acquisition in an L1 illiterate adolescent also provides some anecdotal evidence on how illiteracy, otherwise considered an impediment, does not hinder mastering an additional language. Their illiterate adolescent learner already spoke seven languages when he began learning to read and write. They note that limited literacy is not necessarily a barrier to effective oral communication. Somewhat contrary, Tarone and Bigelow (2005) suggest that L1 literacy may improve L2 acquisition through an improvement in memory and metalinguistic awareness. Although there are findings that support the claim L1 illiteracy hinders learning of morphology (Vainikka & Young-Scholten, 2007; van der Craats, 2011), there are also findings that show that higher L1 metalinguistic awareness, and a conscious attempt to extract grammatical rules (both associated with literacy), may actually harm L2 learning in literate adults (Robinson, 2005; Xhafaj & Mota, 2011). It is possible that literacy aids some areas of language learning (e.g., vocabulary and maybe morphology), but hinders others (e.g., collocations and gender agreement).

Future studies will test the hypothesis that literates not only process language differently but also learn languages differently and might be learning languages in a manner that makes it more difficult to learn aspects of grammar that relate to relations between words. We are currently exploring this possibility in a series of artificial language learning studies with both literate and preliterate children and literate adults, hoping to inform the theoretical and practical debate on the role of multiword units in language learning and teaching (e.g., Boers & Lindstromberg, 2012; Meunier, 2012). Finding that literacy leads learners to rely less on multiword units and hinders learning of certain grammatical relations could have implications for L2 pedagogy. In particular, pre- and postliteracy L2 pedagogy should be constructed differently, building on preliterate students' strengths and targeting literate students' difficulties.

To conclude, we showed that L1 literacy affects L2 lexical segmentation, even when controlling for cognitive and linguistic variability, and when using a simple, online task that relies on short-term memory less than previously used tasks. We also showed an improvement in L2 lexical segmentation following an L2 literacy course, with improvement being greater the less L1 literate the participant was, as would be expected if literacy in any language affects lexical segmentation. We add to the growing literature on the role of multiword units in language processing and expand pervious research to cover lexical segmentation in Hebrew, a Semitic language, much more synthetic than previously studied L2s. Our findings may have important implications for models of L2 learning, and ongoing research performed in our laboratory is currently examining the possibility that L1 literacy may affect not only L2 lexical segmentation but also, through this effect, L2 learning.

ACKNOWLEDGMENTS

This work was supported by ISF Grant 52712 (to I.A.). The authors thank the Haskala course commanders, teachers, and soldiers for their heartfelt cooperation. We thank Anat Prior and Tamar Dgani for useful discussions and sharing of their computerized cognitive tests and Walaa Hajajra for her great help with all aspects of the study relating to Arabic. We also thank Holly Krech Thomas and two anonymous reviewers for their helpful comments. Finally, we thank the competent and devoted research assistants who administered the tasks: Yaron Shapira, Dikla Koren, Tamar Lan, Maya Inbar, and Limor Raviv.

NOTES

1. One of the reviewers proposed that rather than measuring lexical segmentation, our task may reflect cognitive control, with participants who are more able to suppress automatic responses performing better on the task. To rule out the effect of cognitive control on task performance, we ran the model again, and added cognitive control as an additional fixed factor. We did not include it in the original model because this would restrict our sample to the first two courses ($n = 21$). We removed the random effect for participant to allow our model to converge (it accounted for less of the variance than the random effect for items). As in the previous model, the results show a significant effect of word pair frequency ($\beta = -0.27$, $SE = 0.09$, $p = .004$, model comparison $p = .005$), L1 literacy ($\beta = 0.40$, $SE = 0.008$, $p < .0001$, model comparison $p < .001$), L2 vocabulary ($\beta = 0.04$, $SE = 0.1$, $p < .001$, model comparison $p < .001$), and whether word pairs constituted one word in participants' L1 ($\beta = -1.07$, $SE = 0.37$, $p = .005$, model comparison $p = .003$). None of the other factors was significant, indicating that task performance variance cannot be attributed to cognitive control.
2. Here, too, a cumulative link model produces virtually the same results, with both frequency ($\beta = -0.26$, $SE = 0.06$, $p < .0001$) and L1 literacy ($\beta = 0.37$, $SE = 0.07$, $p < .0001$) predicting the probability of a response displaying more segmentation.
3. Some of the illiterate and literate adults previously studied spoke Arabic as their native language (Gombert, 1994; Kurvers et al., 2009; Onderdelinden et al. 2009).

REFERENCES

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *Linguistic Review, 23*, 275–290.

Ardila, A., Bertolucci, P. H., Braga, L. W., Castro-Caldas, A., Judd, T., Kosmidis, M. H., et al. (2010). Illiteracy: The neuropsychology of cognition without reading. *Archives of Clinical Neuropsychology, 25*, 689–712.

Ardila, A., Rosselli, M., & Rosas, P. (1989). Neuropsychological assessment in illiterates: Visuospatial and memory abilities. *Brain and Cognition*, *11*, 147–166.

Arnaud, P. J. L., & Savignon, S. J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 157–173). Cambridge: Cambridge University Press.

Arnon, I. (2010). *Starting big: The role of multiword phrases in language learning and use*. Unpublished doctoral dissertation, Stanford University.

Arnon, I., & Christiansen, M. H. (2014). Chunk-based language acquisition. In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of language development* (pp. 88–91). Thousand Oakes, CA: Sage.

Arnon, I., & Clark, E. V. (2011). Why brush your teeth is better than teeth—Children's word production is facilitated in familiar sentence-frames. *Language Learning and Development, 7*, 107–129.

Arnon, I., & Priva, U. C. (2013). More than words: The effect of multiword frequency and constituency on phonetic duration. *Language and Speech*, *56*, 349–371.

Arnon, I., & Priva, U. C. (2014). Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *Mental Lexicon, 9*, 377–400.

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

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multiword phrases. *Journal of Memory and Language, 62*, 67–82.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19*, 241–248.

Bassetti, B. (2007). Bilingualism, biliteracy and metalinguistic awareness: Word awareness in English and Japanese users of Chinese as a second language. *Birkbeck Studies in Applied Linguistics, 2*, 1–21.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8 [Computer software]. Retrieved from http://CRAN.R-project.org/package=lme4

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 15*, 498–510.

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*, 89–129.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *Linguistic Review*, *23*, 291–320.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32*, 83–110.

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*, 1053–1063.

Chaney, C. (1989). I pledge a legiance tothe flag: Three studies in word segmentation. *Applied Psycholinguistics*, *10*, 261–281.

Cheung, H., & Chen, H. C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes, 19*, 1–28.

Christensen, R. H. B. (2015). Ordinal-regression models for ordinal data. R package version 2015.6-28 [Computer software]. Retrieved from http://www.cran.r-project.org/package=ordinal/

Correa, J., & Dockrell, J. E. (2007). Unconventional word segmentation in Brazilian children's early text production. *Reading and Writing, 20*, 815–831.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499–533.

DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31*, 413–438.

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1238–1255.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71–99.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 829.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition, 15*, 594–615.

Gombert, J. E. (1994). How do illiterate adults react to metalinguistic training? *Annals of Dyslexia, 44*, 250–269.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*, 55–88.

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

Huttenlocher, J. (1964). Children's language: Word–phrase relationship. *Science, 143*, 264–265.

Israeli Central Bureau of Statistics. (2014). *Statistical abstract of Israel* (chap. 8). Jerusalem: Author.

Itai, A., & Wintner, S. (2008). Language resources for Hebrew. *Language Resources and Evaluation, 42*, 75–98.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Jarvis, S. (2009). Lexical transfer. *Bilingual Mental Lexicon: Interdisciplinary Approaches, 70*, 99.

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*, 60–99.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language, 45*, 229–254.

Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. *Formulaic Language, 2*, 499–520.

Karmiloff-Smith, A., Grant, J., Sims, K., Jones, M. C., & Cuckle, P. (1996). Rethinking metalinguistic awareness: Representing and accessing knowledge about what counts as a word. *Cognition, 58*, 197–219.

Kavé, G. (2005). Standardization and norms for a Hebrew naming test. *Brain and Language*, *92*, 204–211.

Kosmidis, M. H., Tsapkini, K., & Folia, V. (2006). Lexical processing in illiteracy: Effect of literacy or education? *Cortex*, *42*, 1021–1027.

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology, 26*, 575–582.

Kurvers, J. (2015). Emerging literacy in adult second-language learners: A synthesis of research findings in the Netherlands. *Writing Systems Research*, *7*, 1–21.

Kurvers, J., & Uri, H. (2006). Metalexical awareness: Development, methodology or written language? A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *35*, 353–367.

Kurvers, J., van Hout, R., & Vallen, T. (2009). Print awareness of adult illiterates: A comparison with young pre-readers and low-educated adult readers. *Reading and Writing, 22*, 863–887.

Levin, I., Ravid, D., & Rapaport, S. (1999). Developing morphological awareness and learning to write: A two-way street. In T. Nunes (Ed.), *Learning to read: An integrated view from research and practice* (pp. 77–104). Amsterdam: Springer.

Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential–expressive distinction. *Journal of Child Language, 19*, 287–310.

Lieven, E. V., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics, 20*, 481–507.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–285.

Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2*, 174–207.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940–967.

McCauley, S. M., & Christiansen, M. H. (2011). *Learning simple statistics for language comprehension and production: The CAPPUCCINO model*. Paper presented at the 33rd Annual Conference of the Cognitive Science Society, Boston.

McCauley, S. M., & Christiansen, M. H. (2014). *Language learning as language use: A computational model of children's language comprehension and production*. Ithaca, NY: Cornell University Press.

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift Für Sprachwissenschaft*, *28*, 5–34.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32*, 111–129.

Miller, K., Frosch, C., Kelly, M. K., & Zhang, H. (2001). *What words look like: Preschool children's understanding of orthographic regularities in Chinese and English*. Unpublished manuscript.

Mor, B. (2012). *The combined effect of bilingualism and ADHD on inhibitory functioning*. Unpublished MA dissertation, University of Haifa.

Morais, J., Bertelson, P., Cary, L., & Alegría, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45–64.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Olivers, C. N. L., Huettig, F., Singh, J. P., & Mishra, R. K. (2014). The influence of literacy on visual search. *Visual Cognition*, *22*, 74–101.

Olson, D. R. (1996). *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.

Onderdelinden, L., van der Craats, I., & Kurvers, J. (2009). Word concept of illiterates and low-literates: Worlds apart? *LOT Occasional Series*, *15*, 35–48.

Ostrosky-Solís, F., & Lozano, A. (2006). Digit span: Effect of education and culture. *International Journal of Psychology, 41*, 333–341.

Petersson, K. M., Reis, A., & Ingvar, M. (2001). Cognitive processing in literate and illiterate subjects: A review of some recent behavioral and functional neuroimaging data. *Scandinavian Journal of Psychology, 42*, 251–267.

Pettitt, N. M., & Tarone, E. (2015). Following Roba: What happens when a low-educated adult immigrant learns to read. *Writing Systems Research*, *7*, 20–38.

Ravid, D. (2001). Learning to spell in Hebrew: Phonological and morphological factors. *Reading and Writing, 14*, 459–485.

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

Ravid, D., & Tolchinsky, L. (2002). Investigating literacy development and language acquisition. Reply to commentaries on "Developing linguistic literacy: A comprehensive model." *Journal of Child Language*, *29*, 489–494.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *27*, 235–268.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Share, D. L., & Levin, I. (1999). Learning to read and write in Hebrew. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 89–111). Cambridge: Cambridge University Press.

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.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *64*, 429–458.

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 776.

Taguchi, N. (2007). Chunk learning and the development of spoken discourse in a Japanese as a foreign language classroom. *Language Teaching Research*, *11*, 433–457.

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

Tarone, E., Bigelow, M., & Hansen, K. (2007). The impact of alphabetic print literacy level on oral second language acquisition. In N. Faux (Ed.), *Low-educated second language and literacy acquisition: Research, policy, and practice. Proceedings of the second annual forum of the Literacy Institute* (pp. 99–122). Richmond, VA: Virginia Commonwealth University, Literacy Institute.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*, 569–613.

Tsai, C.-H., McConkie, G. W., & Zheng, X. J. (1998). *Lexical parsing by Chinese readers*. Poster presented at the Advanced Study Institute on Advances in Theoretical Issues and Cognitive Neuroscience Research of the Chinese Language, University of Hong Kong. Retrieved from http://research.chtsai.org/papers/cht_asi.html

Vainikka, A., & Young-Scholten, M. (2007). The role of literacy in the development of L2 morphosyntax from an Organic Grammar perspective. In *Low-educated second language and literacy acquisition: Research, policy, and practice: Proceedings of the second annual forum of the Literacy Institute* (pp. 123–148). Richmond, VA: Virginia Commonwealth University, Literacy Institute.

Valian, V., & Levitt, A. (1996). Prosody and adults' learning of syntactic structure. *Journal of Memory and Language, 35*, 497–516.

van der Craats, I. (2011). A LESLLA corpus: L1 obstacles in the learning of L2 morphosyntax. In C. Schöneberger, I. van de Craats, & J. Kurvers (Eds.), *Low-educated adult second language and literacy acquisition. Proceedings of the 6th symposium* (pp. 33–48). Nijmegen, Netherlands: Centre for Language Studies.

Vanlancker-Sidits, D. (2003). Auditory recognition of idioms by native and nonnative speakers of English: It takes one to know one. *Applied Psycholinguistics, 24*, 45–58.

Veldhuis, D. (2012). *Effects of literacy and frequency of occurrence on the offline segmentation and online processing of language*. Paper presented at the Anéla 2012 Applied Linguistics Conference.

Veldhuis, D., & Kurvers, J. (2012). Offline segmentation and online language processing units: The influence of literacy. *Written Language & Literacy, 15*, 165–184.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition (WAIS–IV)*. San Antonio, TX: NCS Pearson.

Wong-Fillmore, L. (1991). Second language learning in children: A model of language learning in social context. In E. Bialystok (Ed.), *Language processing by bilingual children* (pp. 49–69). Cambridge: Cambridge University Press.

Wood, D. (2002). Formulaic language in acquisition and production: Implications for teaching. *TESL Canada Journal, 20*, 1–15.

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, *12*, 39–58.

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching, 32*, 213–231.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21*, 463–489.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A. (2009). Future directions in formulaic language research. *Journal of Foreign Languages, 32*, 2–17.

Wray, A. (2013). Formulaic language. *Language Teaching, 46*, 316–334.

Wray, A., & Fitzpatrick, T. (2010). Pushing speakers to the extreme: The artificial use of prefabricated material in conversation. *Innovation in Language Learning and Teaching, 4*, 37–51.

Xhafaj, D., & Mota, M. B. (2011). Can metalinguistic awareness explain a connection between L1 literacy and L2 proficiency level in literate adults? *Ilha do Desterro, 60*, 189–246.