Journal of Memory and Language 85 (2015) 60–75

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The advantage of starting big: Learning from unsegmented input facilitates mastery of grammatical gender in an artificial language

Noam Siegelman, Inbal Arnon *

Department of Psychology, The Hebrew University, Jerusalem 91905, Israel

ARTICLE INFO

Article history:
Received 2 April 2015
revision received 26 July 2015
Available online 10 August 2015

Keywords:
Language learning
Multiword units
Second language learning
Grammatical gender
Adult L2 learning

ABSTRACT

Why is it difficult to learn a second language as an adult? We focus on the way adults' existing knowledge of words impacts L2 learning. We suggest adults' prior knowledge leads them to rely less on multiword units, and that this hinders learning certain grammatical relations between words. We test this prediction in two artificial language learning studies of grammatical gender with adult learners. In the first study, we manipulate whether the artificial language is segmented during initial learning and show that learning from unsegmented input leads to more article–noun units, and to better learning. Individually, learners who were more likely to treat the article and noun as one unit showed better learning on an off-line measure, confirming the link between unit-size and learning outcomes. In the second study, we show this pattern does not hold when the article conveys semantic information (animate vs. inanimate), illustrating the way unit size interacts with informativity during learning. The findings provide novel evidence for the advantage of learning grammar from multiword units, highlight the benefit of learning segmentation and structure simultaneously, and offer an experience-based explanation for L1–L2 differences that relates building blocks to learning outcomes.

Introduction

Why is acquiring a second language as an adult so difficult? Unlike children, adults rarely reach native-like proficiency in a second language. Numerous studies have documented adult learners' difficulty in a range of linguistic domains (e.g., Clahsen & Felser, 2006; Johnson & Newport, 1989; Moyer, 1999). Given that adults perform better on a wide range of learning tasks, the fact that children seem to be better language learners is puzzling. Many existing accounts attempt to explain this puzzle by focusing on the biological, neural, and cognitive differences between children and adults (Kuhl, 2000; Lenneberg, 1967; Neville & Bavelier, 2001), and the way early neural commitment shapes subsequent learning (e.g., Werker & Tees, 1984). Differences in cognitive abilities—such as cognitive control or working memory—have also been invoked to explain the different learning outcomes in children and adults (Elman, 1993; Newport, 1990; Ramscar & Gitcho, 2007; Thompson-Schill, Ramscar, & Chrysikou, 2009).

Such accounts can explain the overall qualitative difference between children and adults in language learning, but have a harder time explaining the particular pattern of learning outcomes observed in adult learners. Specifically, adults are not always worse than children in learning a novel language, and seem to struggle with certain domains more than others. Adult learners, for instance, are generally faster and more efficient than children in the early stages of learning (e.g., Krashen, Long, & Scarcella, 1979), and are often better in experimental settings (Ferman & Karni, 2010). Moreover, adults do not find all aspects of the novel

* Corresponding author.
E-mail address: inbal.arnon@mail.huji.ac.il (I. Arnon).

http://dx.doi.org/10.1016/j.jml.2015.07.003
0749-596X/

language equally hard: vocabulary and word order are learned with more ease than morphological and syntactic regularities (e.g., DeKeyser, 2005). Even within morpho-syntax, certain relations are harder to master than others: domains such as grammatical gender, article use, and classifiers pose difficulty even for highly proficient adult speakers (Johnson & Newport, 1989), including those classified as near-native (Sorace, 2003). The currently unresolved challenge is to link specific learner characteristics to the particular learning outcomes we see in adult L2 learners. That is, to explain what adults will find particularly hard, and why.

In this paper, we approach this language-learning puzzle by focusing on an additional factor that may affect L2 learning: the effect of experience and existing knowledge on learning outcomes. We focus in particular on the way adults' existing knowledge of words influences the linguistic units they learn from – their early building blocks – and how those early units shape subsequent learning. That is, we ask how adult's prior knowledge impacts the units they focus on and attend to in learning a novel language. Learners are affected by the properties of their L1 when learning a second language: for better and worse, L1 knowledge changes how a novel language is processed by adult learners. Recent studies using artificial languages illustrate such effects: L1 phonotactics impact how learners segment a novel speech stream (Finn & Hudson Kam, 2008; Toro, Pons, Bion, & Sebastiá n-Gallés, 2011), which in turn affects morphological learning (Finn & Hudson Kam, 2015). On the positive side, having certain properties in your L1 (e.g., tone) can enhance learning of similar relations in a novel language (Onnis & Thiessen, 2013; Wang & Saffran, 2014).

These studies show how the specific content of L1 influences L2 learning. Here, we focus on a broader 'burden' that adults bring to the task of learning a second language: their knowledge of words. Unlike infants – who are born into a "blooming, buzzing confusion" (William James, 1890) and have to discover the different units of language – adults already know that words exist. This knowledge, we suggest, impacts the linguistic units adults attend to and use in learning a novel language. In particular, we propose that adults' existing knowledge of words leads them to rely less on multiword units in learning a second language, and that this hinders learning certain grammatical relations by *reducing the strength of the association* between the linguistic elements in question (we elaborate on this below; see Arnon & Ramscar, 2012, for details). If some of adults' difficulty is related to the units they attend to and learn from, then manipulating their input in a way that affects these units should lead to changes in learning outcomes. In particular, manipulating adults' early experience in a way that directs their attention to multiword units should enhance learning in certain domains, but not others (as will be detailed below).

This proposal diverges from commonly held assumptions in several ways. First, it focuses on the role of existing knowledge about language in general on L2 difficulty, rather than on knowledge of specific L1 properties. Second, it predicts a link between early building blocks – the linguistic units used during learning – and specific L2 outcomes, and in doing so, it tries to offer an explanation for why adults struggle more with certain aspects of the novel language compared to others. Third, it assumes that multiword units are building blocks for language, that is, that speakers are sensitive not only to the properties of sounds, morphemes and words (the linguistic elements traditionally seen as building blocks of language), but also to those of multiword sequences. This idea has been gaining theoretical and empirical support in recent years. The assumption that multiword units are integral units of language follows from single-system models of language, where all linguistic experience – sounds, words, multiword sequences – is processed by the same cognitive mechanisms (e.g., Bybee, 1998; Christiansen & Chater, 2015; Elman, 2009; McClelland, 2010). Such models predict that speakers will be sensitive to distributional information computed at multiple granularities (between sounds, words, and sequences of words) and draw on units of varying sizes in language. Accordingly, there is growing empirical evidence illustrating the parallels in processing words and larger sequences, and showing that children and adults are indeed sensitive to the distributional properties of multiword units in learning and processing (e.g., Arnon & Cohen Priva, 2013; Arnon & Snider, 2010; Bannard & Matthews, 2008; Reali & Christiansen, 2007; Tremblay & Baayen, 2010). Finally, this perspective on L2 difficulty has the advantage of looking at something that can be manipulated: while we cannot undo the way a brain or mind have developed, we can influence the units adults attend to in learning a novel language.

We examine this proposal by looking at adult learning of grammatical gender in an artificial language learning (ALL) paradigm. Specifically, we use this paradigm to experimentally manipulate whether adults' early experience with a novel language is segmented (leading to more single word units) or unsegmented (leading to more multiword units). A previous study provided initial support for the idea that learning from multiword units can facilitate mastery of certain grammatical relations between words, such as the association between articles and nouns in languages with grammatical gender (Arnon & Ramscar, 2012, described in detail in the next section). In the current paper, we significantly expand and qualify this proposal. We provide novel, and more direct evidence for the role of multiword units in learning, while also showing that learning from larger units is beneficial for learning certain grammatical relations, but not others. Together, the studies offer an experience-based theory of the relation between adult's units of learning and their learning outcomes that aims to capture and predict the particular profile of learning outcomes observed in the literature.

Grammatical gender as a case study

Grammatical gender is a system found in many languages. It assigns all nouns (including inanimate ones) to noun classes, usually marking neighboring words for agreement (Corbett, 1991). In Spanish, for instance, articles have to agree in gender with the noun, while in Hebrew only verbs and adjectives carry gender agreement. The assignment is arbitrary in the sense that the same object can be assigned a different gender in different languages.

In languages that mark gender, knowing a noun's gender is essential for correct sentence construction.

Grammatical gender agreement provides a fruitful domain to study the effect of multiword building blocks on learning outcomes since learning differs markedly between native speakers and L2 learners. While children master this system relatively early (Lew-Williams & Fernald, 2007; Slobin, 1985), adult learners struggle with it even after extensive exposure (e.g., Lew-Williams & Fernald, 2010; Scherag, Demuth, Roesler, Neville, & Roeder, 2004). L2 learners also differ from native speakers in their ability to use the gender information conveyed by the article predictively. While native speakers (including young children) can use the article to guide lexical access – they anticipate a feminine noun following a feminine article (e.g., Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009)—non-native speakers do not show these effects (Dussias, Valdes Kroff, Guzzardo Tamargo, & Gerfen, 2013; Lew-Williams & Fernald, 2010). Generally speaking, L2 learners seem to treat the article and noun as a less cohesive unit: they have more difficulty selecting the correct article in production and cannot use it to facilitate noun recognition.

Learning grammatical gender from multiword units

Why do L2 learners have such difficulty in learning the association between the article and the noun? We hypothesize that this has to do with their smaller reliance on multiword units (Arnon & Ramscar, 2012). One can imagine two ways of learning the article–noun pairing. In the 'adult-like' way, learners' existing knowledge of words leads them to treat the article and noun as two separate units. They start out by associating the object, say 'hand' with its label *mano* (hand in Spanish), and only later learn that it is a feminine noun and must therefore appear with the feminine article 'la'. In the 'child-like' way, learners – who are learning segmentation and meaning simultaneously – initially associate the entire article–noun sequences 'la-mano' with the object (hand), and then learn, over time, that the two elements can appear separately. Importantly, we predict that the 'child-like' way of learning leads to a greater association between the article and the object (hand), which will lead to better learning of the article–noun pairing.

In a recent study, Arnon and Ramscar (2012) explored this idea computationally and experimentally. The learning advantage of the 'child-like' way was formalized within the framework of error-driven learning, specifically using the concept of blocking (Kamin, 1969).[1] In these theories, prior learning can *block* the learning of later information by making it redundant: when an existing cue fully predicts a response, introducing a new cue will not add information, and will consequently make it harder to learn the association between the novel cue and the response (see Ramscar et al., 2010; Ramscar et al., 2013). In the case of grammatical gender, the initial association between the noun-label and the object in the 'child-like' way will block later learning of the association between the article and the object (the article does not add information about the object), and reduce the association between the article and the noun. In the 'adult-like' way, in contrast, the initial association between the article–noun sequence and the object will be preserved, even after they are dissociated (see paper for full details). This framework makes two clear predictions (a) that having adults learn from larger units (making them be more 'child-like') will lead to better learning of the relation between the article and the object when that relation is arbitrary (as in many grammatical gender systems), and (b) that this advantage will disappear when the article itself adds information about the object, that is, when it carries semantic information.

The first prediction was tested by looking at adult learning of grammatical gender in an artificial language and manipulating learners' early input in a way that increased initial unit size (Arnon & Ramscar, 2012). As predicted, participants who were exposed first to a block of full sentences (including the article–noun sequence) and then to a block of single nouns showed better learning of the article–noun pairing compared to participants who were exposed first to nouns and only then to full sentences: they were better at selecting the correct article in a forced-choice task, and better at producing the correct article for a given noun. Even though both groups had received exactly the same input (just in different order), the sentence-first group showed better learning. Similar results were obtained in simulated acquisition of the language using a Rescorla & Wagner, 1972 model: despite equal exposure, the association between the article and object was better when exposed first to full sentences.

These results provide preliminary evidence for the effect of existing knowledge and early experience on learning outcomes: they show that early exposure to larger units (full sentences, including the article–noun sequence) leads to better learning of grammatical gender in adult learners. However, they have one major limitation: since the units learners used were not directly evaluated during learning, the study cannot establish a direct link between the size of early building blocks and learning outcomes. Put differently, we do not know if participants in the sentence-first condition showed better learning *because* they were more likely to treat the article and noun as a single unit. Moreover, because the full sentences contained cues to word segmentation (there were pauses between words), it is possible that participants in both conditions segmented the speech stream into words from the start. The previous study also did not test the second prediction made by this framework: that the learning advantage will be found only in cases when the article does not carry semantic information.

[1] Note that this learning advantage can be formally captured in several other modeling approaches, including Bayesian and connectionist frameworks. However, discriminative learning models like the one presented here have been successfully applied in recent years to both L1 and L2 learning (Arnon & Ramscar, 2012; Ellis, 2006; Ellis & Sagarra, 2010; Ramscar, Yarlett, Dye, Denny & Thorpe, 2010; Ramscar, Dye, & McCauley, 2013; Ramscar, Dye, Popick, & O'Donnell-McCarthy, 2011) and seem particularly well-suited to capture the way early unit size affects what is subsequently learned through interference, blocking and other mechanisms.

The current study

The current study will extend and qualify the theoretical proposal outlined above in several important ways: (a) by demonstrating a more direct relation between the use of multiword units during learning and better learning outcomes, (b) by showing how individual differences in the tendency to extract multiword units affects learning of grammatical relations, and (c) by showing this effect is limited to cases where the article does not carry semantic information. Experiment 1 will address the possible shortcomings of previous work by providing a better and more direct manipulation of unit size, and, more importantly, by introducing a new online measure of unit size, which allows us to identify the units learners used during learning. Using this measure we can both (1) validate the manipulation to see whether there is a difference in unit size between the conditions, and (2) extend the findings to show a direct relation between unit size and learning outcomes: at the individual level, we predict that more reliance on multiword units will lead to better learning.

Experiment 2 will test a novel prediction about the limits of the multiword advantage by asking whether learning from multiword units helps also when the division of nouns into noun classes is not arbitrary, but rather semantic (as in animate–inanimate systems). We use the exact same language as in Experiment 1 to assess learning when the article distinguishes between animate and inanimate nouns. Our account predicts that learning such relations will *not* be facilitated by greater reliance on multiword units, because the article in this case carries semantic information about the object, and is consequently predictive of it. Put differently, the benefit of learning from multiword units will disappear when there is a stronger (semantic) cue available to map the article to the object. In terms of blocking, the association between the article and the object will not be blocked by an earlier association between the noun and the object. Together, the studies serve to expand and qualify the role of multiword units in learning, and the effect of early units on learning outcomes. They illustrate when multiword units are and are not advantageous in L2 learning, and demonstrate the power of asking—given very simple assumptions about how learning takes place—how basic features of prior knowledge exert a profound influence on learning outcomes.

Experiment 1

This study is modeled on the language used by Arnon and Ramscar (2012) with several important modifications. Instead of comparing sentences-first learning to words-first learning, we introduce two novel conditions that both involve exposure to full sentences. The difference between the conditions is in the degree to which early stimuli are segmented into words and, consequently, how likely learners are to extract word vs. multiword units. In the unsegmented-first condition, participants are exposed first to unsegmented sentences (with no pauses between words) and only then to segmented sentences (with pauses between words). We compare this to a condition where participants hear segmented sentences first (with pauses between words), and only then unsegmented stimuli. By the end of the study, participants in both conditions heard exactly the same input, just in a different order. After training, we assess learning using a forced-choice task and a production task, both of which test participants' knowledge of the article–noun pairings (the learning measures are identical to those used in Arnon & Ramscar, 2012). If learning segmentation and structure simultaneously results in more multiword units, and if this leads to better learning, we should see better outcomes in the unsegmented-first condition, despite the additional burden of having to segment and learn structure at the same time.

Importantly, we introduce a new measure of building blocks/segmentation by asking participants to type in the stimuli they heard during learning. This measure allows us to see if participants treat the article and noun as one unit or two in the course of learning, and to test several predictions. First, if early exposure to unsegmented speech leads to more multiword units, there will be more article–noun 'words' produced by subjects in this condition. Second, if multiword building blocks lead to better learning, then participants who were more likely to treat the article and noun as one unit should show better learning of the pairing between them. Confirmation of these two predictions would provide novel evidence for the link between unit size and learning of grammatical relations, and for the relation between building blocks and learning outcomes more generally.

Method and materials

Participants

Forty-four students at the Hebrew University (25 females and 19 males) participated in the study for course credit or payment. Their age ranged from 19 to 37 (mean age 25.9). Subjects were all native Hebrew speakers, with no history of speech, reading or hearing disability, ADD or ADHD. Subjects were randomly assigned to one of the two conditions, segmented-first or unsegmented-first (22 participants in each).

Materials

The artificial language had 12 novel labels for concrete nouns, 2 articles (*fo* and *se*), and a carrier phrase (*osferpel-ti*, see the Appendix A for item list). All objects had high-frequency, early-acquired Hebrew labels. The nouns were divided into two "classes" and each noun only appeared with one article. Hebrew has grammatical gender, though there is no gender agreement with articles, unlike in our artificial language. To ensure that learning is not affected by the gender of nouns in Hebrew (the L1), the two noun 'classes' were matched on the gender of the nouns in Hebrew: each had an equal number of masculine and feminine nouns. There were no semantic or phonological cues to class membership – the only cue was distributional (which article the noun appears with). All noun labels were two syllables long. All objects had high frequency labels in Hebrew, were all concrete, and

all had early-acquired labels (Based on the English Age-of-Acquisition Bristol norms, Stadthagen-Gonzalez & Davis, 2006, which have recently been validated cross-linguistically, Luniewska, 2014). The artificial language had a fixed word order: articles always followed the carrier phrase and preceded the nouns (see example 1).

(1) os ferpel ti	**fo**	**etkot**
carrier phrase	article	noun

A female Hebrew speaker recorded the carrier phrase, the articles, and the nouns separately. The elements were synthesized to a frequency of 170 Hz to remove possible prosodic cues to word boundaries, and concatenated using Praat (Boersma, 2001) to create the full sentences for the two conditions. The same recorded token of each noun, article and carrier phrases was used throughout the experiment. The duration of the two articles was identical to ensure they were equally prominent. Segmented sentences had a 250 ms pause inserted between each element (one pause between the carrier phrase and the article, and one between the article and the noun-label) while unsegmented sentences had no pauses between the elements.

In addition to the experimental items, a second set of phrases was constructed to serve as a distractor block between training and testing in order to ensure that the last block before testing was identical in the two conditions. This "distractor block" had the same carrier phrase, but had six different objects, referred to using different nouns, and two different articles (*tid* and *gob*). The stimuli for this set were recorded by a male speaker, making it easy to distinguish between the test language and the distractor block. Unlike the test language, the mapping between articles and nouns was not consistent and each noun could appear with both articles. Note that all the objects, noun-labels and articles in the distractor block were different from the ones used in the test language, in order to minimally interfere with learning.

Procedure

The experiment was divided into two parts: learning trials and test trials. Participants were told they would be learning a novel language by seeing pictures and hearing their description in that language. The experiment lasted around 25 min (training: 15–20 min; testing: 5–7 min). Participants were instructed to listen carefully as they would later be tested on the language they just heard. They were also told they would periodically be asked to type in the last sentence they heard.

Learning trials. Fig. 1 presents the learning procedure in the two conditions. As can be seen, participants in the two groups were exposed to the same stimuli, but in a different order. Participants in the unsegmented-first condition heard unsegmented sentences first while participants in the segmented-first condition heard segmented sentences first. Overall, participants heard 120 sentences during learning: 60 segmented and 60 unsegmented (each of the 12 nouns was repeated five times in each condition). Segmented sentences had 250 ms pauses between words while unsegmented sentences had no pauses between words. Each sentence type (segmented and unsegmented) appeared in a separate block – the only difference between the conditions was which block appeared first. Following the two learning blocks, participants in both conditions were exposed to a distractor block with 36 sentences (six novel objects, each repeated six times). The distractor block was introduced to control for any recency effects in testing, and ensure that the last block before testing was identical in the two learning conditions.

Typing trials. On fifteen trials in each block, participants were instructed to type the last sentence they heard (total of 30 typing trials over the two blocks). The serial position of typing trials in each block was the same for all participants (taken from a pre-generated sequence of 15 position that did not allow more than two typing trials in a row). However since stimuli presentation was randomized across participants, each participant typed in different sentences. Responses were recorded and later coded by a researcher blind to experimental condition. Responses were coded according to whether the article and noun were typed as one or two words.

Test trials. Test trials were identical in the two conditions. Participants completed a two alternative forced choice task, followed by a production task. The forced-choice test consisted of 24 trials. On each trial, participants saw a picture, heard two sentences and had to say which sentence was the correct one in the language. They were told that only one sentence was correct. Half of the forced-choice trials tested participants' knowledge of the article–noun pairing (henceforth – *article* trials, where the incorrect sentence had the right noun but wrong article). The other half of the trials tested knowledge of the noun–object mapping (*noun* trials, where the incorrect sentence had the wrong noun but the right article). Each of the 12 objects appeared once in an article trial and once in a noun trial, yielding 24 forced-choice trials. Each of the twelve noun labels appeared four times over the course of the test: twice in the same article trial (once as a target and once as a distractor, with different articles) and twice in two noun trials (again, once as a target and once as a distractor). In each trial type, half of the sentences consisted of stimuli with 250 ms breaks between words, and the other half of sentences without breaks to ensure both stimuli types were represented at testing. The order of the trials and the order of the two options within each trial were randomized for each participant.

Participants performed a production task following the forced-choice task. In this task, participants saw a picture and were asked to produce a full sentence describing it. Each object appeared once, yielding 12 production trials. Participants' responses were recorded and coded by a researcher blind to experimental conditions. For each trial, the article and the noun were separately coded as correct or incorrect. Nouns and articles were coded as correct if they did not differ from the target in more than one sound (e.g., *efkot* instead of *etkot* was coded as correct, see Arnon & Ramscar, 2012 for full details). Productions that differed

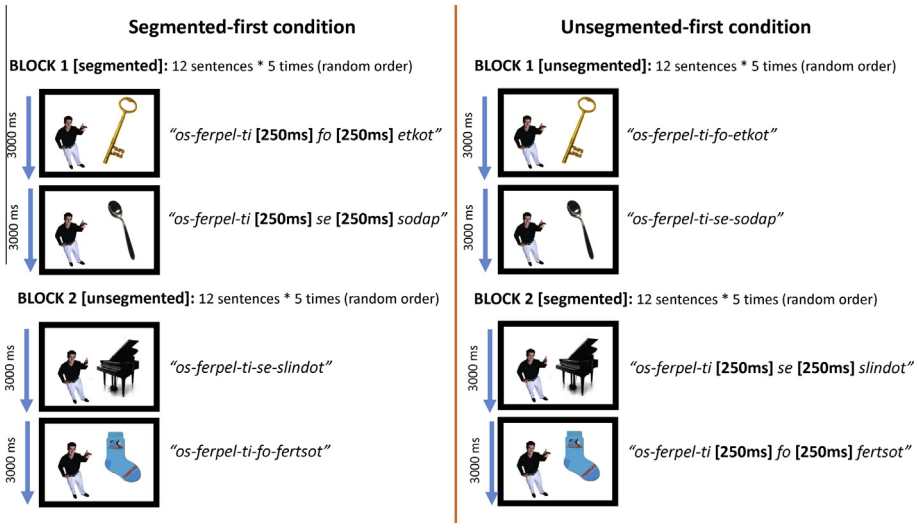

Fig. 1. Learning procedure in the two conditions (segmented-first on the left, unsegmented-first on the right). In the unsegmented-first condition, participants heard a block of un-segmented sentences and then a block of segmented ones. In the segmented-first condition the order of blocks was reversed. Each trial lasted for 3000 ms. The order of trials within blocks was randomized. Objects were changed on each trial (see Appendix A for a list of nouns and their meanings).

from *both* articles in one sound (e.g., *fe*) were coded as incorrect. The production of the carrier phrase was also coded for accuracy, on a 1–3 scale (3 – fully accurate, 2 – partially accurate, 1 – not accurate).

Results and discussion

We had three measures of learning: forced-choice and production measures from the test trials, as well as typing measures collected during learning. In order to remove participants who were not showing any learning, we excluded participants who scored less than 2SD from the mean in the noun trials in the forced-choice task, a measure on which performance was high (86% success before exclusion). Six participants (three in each condition) displayed such low performance, and were therefore excluded from further analyses. We use mixed-effect regression models to analyze the results, using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015). All models had the maximal random effects structure justified by the design (Barr, Levy, Scheepers, & Tily, 2013). Collinearity in all reported models was small (<0.3). Categorical variables were effect coded in all analyses. Results were qualitatively identical using ANOVA (i.e., significant effects remained significant and non-significant effects remained non-significant). We report mixed-effects models because ANOVA is less appropriate for use with binary responses (Jaeger, 2008).

Manipulation validation

We looked at the typing measures to see if indeed participants in the unsegmented-first condition relied more on multiword units during learning. Each of the 30 typing trials was coded for whether the participant typed the article and noun as one word (=1) or two words (=0). We then calculated an overall score for each participant based on the percentage of times that he/she typed the article and the noun as one word (one subject was excluded from this analysis because he/she did not type full sentences as instructed, but only the object label). Before testing our manipulation, we first used the typing measure to make sure that participants were learning article–noun pairings (a grammatical gender system) in both conditions, and were not just learning to map longer words (article + noun 'words') onto objects, without separating the two. If participants were simply learning longer words, then they should type the article–noun as one unit throughout the entire learning phase (even after exposure to segmented sentences), suggesting they had not learned there are two separate elements. Only one participant (in the unsegmented-first condition) displayed this pattern (typing the article–noun as one word throughout the study), and he was removed from subsequent analyses. All remaining participants ($N = 36$) typed in the article and noun as two separate words in (at least) the last eight trials of the learning phase. This shows that by the end of the learning phase, all remaining participants were treating the article and noun as two words, indicating that they were not just learning long words.

Turning to our manipulation test, as predicted, participants in the unsegmented-first condition were more likely to treat the article and noun as one word ($M = 31\%$, $SD = 23\%$) compared to participants in the segmented-first condition ($M = 2\%$, $SD = 5\%$, $t(39) = 4.42$, $p < 0.001$). This finding validates our manipulation: participants exposed first to sentences without breaks perceived them as less segmented and were more likely to treat the article and noun as one unit.

Forced-choice trials

We ran a mixed-effect logistic regression model, with learning condition (segmented-first vs. unsegmented-first),

trial type (article vs. noun) and article-type (*fo* vs. *se*) as fixed effects. The model included random intercepts for participant and item, as well as a by-subject random slope for article-type and a by-item random slope for trial type (the maximal random effect structure justified by the data that converged). Overall, the effect of learning condition was not significant (84% vs. 79%, $\beta = .13$, $SE = .14$, $p = .34$). However, there was a marginally significant interaction of condition and trial type ($\beta = -.19$, $SE = .1$, $p = .07$). As predicted, learning condition affected article trials, but not noun trials (see Fig. 2): exposure to unsegmented input first led to better learning of the article–noun pairing, but not of the noun labels themselves. Not surprisingly, given the difficulty of grammatical gender, both sets of participants selected the correct noun-label more often than they selected the correct article (91% vs.72% correct, $\beta = .61$, $SE = .12$, $p < .001$). The effect of article was also significant, with *se* articles being learned better than *fo* articles ($B = .35$, $SE = .15$, $p < .05$). This unexpected effect could have been caused by the similarity of *se* to the Hebrew demonstrative article *ze* ('this').

Since we predicted that learning condition will specifically enhance learning of article–noun pairings, we ran two additional models, conducted separately for article and noun trials, to test this prediction. The models had learning condition (segmented-first vs. unsegmented-first), and article-type (*fo* vs. *se*) as fixed effects. The models included random intercepts for participant and item, as well as a by-subject random slope for article-type and a by-item random slope for condition (the maximal random effect structure justified by the data that converged). Learning condition significantly affected accuracy in choosing the correct article (77% vs. 66%, $\beta = .30$, $SE = .15$, $p < .05$), but did not affect the accuracy in selecting the correct noun (90% vs. 92%, $\beta = -.11$, $SE = .21$, $p = .58$). As predicted, learning from unsegmented input led to better learning of the article–noun pairings, but not of the noun-labels themselves. This effect was also confirmed in post hoc *t*-tests: the two groups differed in their accuracy on articles ($t(35) = 2.09$, $p < .05$), but not nouns ($t(35) = -0.32$, $p = .74$).

Production trials

We ran two mixed-effects logistic regression models for the coded responses of articles and nouns on the production task. Both models included condition (segmented-first vs. unsegmented-first) and article-type (*fo* vs. *se*) as fixed effects, as well as random intercepts for participant and item. Three participants (one in the segmented-first condition and two in the unsegmented-first condition) were excluded from the analysis since they failed to understand the instructions and tried to type their answers rather than to say them. The subsequent analysis is based on the responses of the remaining 34 participants. Fig. 3 presents the mean percentage of correct responses for articles and nouns in the two conditions. The overall accuracy rates in the production task were lower than in the forced-choice task, which is not surprising given the short exposure and number of nouns learned (12). The effect of learning condition on noun and article accuracy was not significant. While participants in the unsegmented condition were somewhat better at producing the correct article

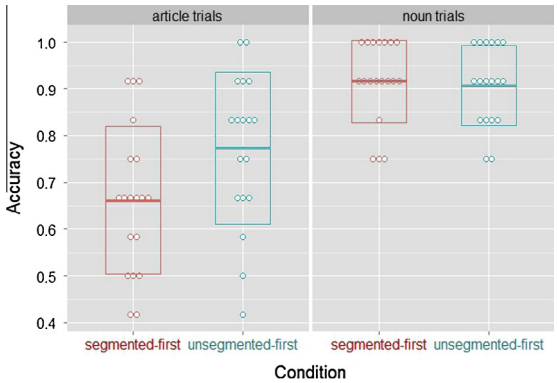

Fig. 2. Distribution of scores, means, and SDs in the forced-choice test by condition and trial type.

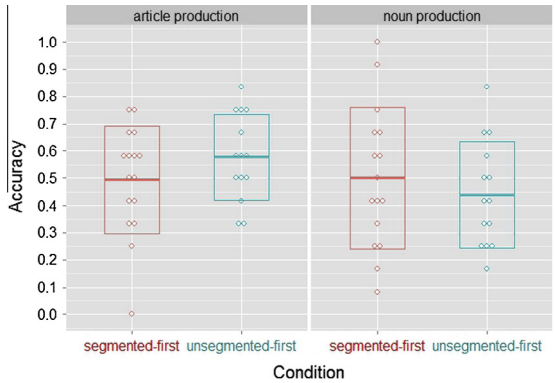

Fig. 3. Distribution of scores, means and SDs in the production test by condition and production type.

(58% vs. 49%, see Fig. 3), this difference was not significant ($B = .25$, $SE = .16$, $p = .14$). The two groups showed equal accuracy on producing the nouns ($B = -.12$, $SE = .2$, $p = .54$).

Because we were interested in the effect of learning condition on the production of the correct article–noun pairing, we conducted a follow-up analysis focusing only on trials where the noun was produced correctly. For these trials ($N = 182$), we asked whether the produced article was correct or incorrect. We calculated the ratio of correct articles given a correct noun in the two conditions. Participants were more likely to produce the correct article for a given noun in the unsegmented-first condition compared to the segmented-first group (77% vs. 57%, Chi-square = 6.86, $p < .01$; We used Chi-square and not *t*-test to better control for the different number of correct nouns produced by each participant). Importantly, there was no difference between the two conditions in the accuracy of the carrier phrase ($p > 0.3$): what participants learned better in the unsegmented-first condition was the pairing of the article and noun.

Typing trials

To test the relation between segmentation rates and learning, we calculated a measure of under-segmentation for each individual participant, based on the percentage

of typing trials (out of 30) where the article–noun were typed as one word. As predicted, those who were worse at segmenting the speech stream were actually better in learning the article–noun pairing, as reflected by a significant correlation between this measure and success in the forced-choice article trials ($r = 0.45$, $p < 0.01$): The scatter plot for these two measures is shown in Fig. 4. Importantly, all other correlations between this measure and learning outcomes (forced-choice noun trials, and production measures) were smaller than ±0.25, suggesting that treating the article and noun as one unit was specifically linked to the ability to correctly pair the articles and nouns.

In sum, all three measures illustrate the effect of early experience on building blocks, and document the advantage of utilizing multiword building blocks in learning grammatical gender. In line with our predictions, we found that participants were more likely to treat the article and noun as one unit when they were exposed first to unsegmented sentences. More importantly, we found this had an effect on learning: participants in the segmented-first condition showed better learning of the article–noun pairings. This pattern was also found on an individual level: participants who were more likely to treat the article–noun sequence as one word showed better learning of the article–noun pairing. Taken together, the findings document the predicted link between multiword building blocks and better learning outcomes.

Experiment 1 set out to expand on the findings in Arnon and Ramscar (2012) by providing a more direct manipulation of unit size, and by monitoring actual unit size during learning. The two studies – which use a similar artificial language – converge on the positive effect of using multiword units as early building blocks. However, the results of the studies also differ in several interesting respects. First, learning of the article–noun pairing was generally better in the current study: participants in this study were better at selecting the correct article compared to ones in the Arnon and Ramscar study (72% vs. 58% in the previous study, averaging over learning conditions). Second, the effect of the manipulation on learning was also stronger in the current study: the difference in correct article selection between the two learning conditions was only 7% in the previous study compared to 11% in the current one.

Two factors may have contributed to these differences. First, participants in the current study were all native Hebrew speakers – a language with grammatical gender, while participants in the previous study were monolingual English speakers – a language with no grammatical gender. Although Hebrew does not mark articles for gender, it does classify nouns into masculine and feminine and marks verbs and adjectives for gender agreement. The better performance of our Hebrew speakers on a grammatical gender learning task is consistent with findings that speakers whose L1 has grammatical gender have an easier time learning such systems in a second language (e.g., Sabourin, Stowe, & de Haan, 2006). The generally better performance, as well as the larger difference between the learning conditions may also reflect the strength of the

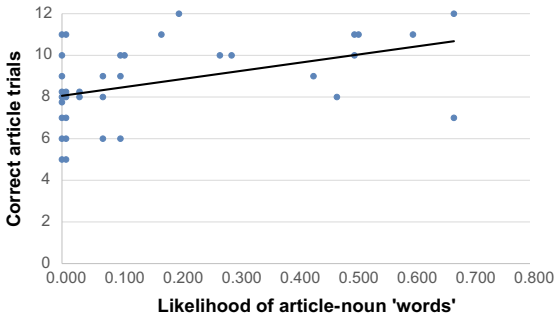

Fig. 4. The correlation between the tendency to treat the article and noun as one unit and success on forced-choice article trials.

current manipulation. Unlike the previous study, participants were not exposed to individual words but only to sentences. Moreover, exposure to unsegmented speech (rather than just to full sentences) may have led to more multiword units overall, which in turn would lead to better learning.

Overall, the two studies provide strong support for the idea that learning from multiword units can lead to better learning of article–noun pairings in adult learners. But what is the mechanism that underlies this effect? We hypothesized (here and in Arnon & Ramscar, 2012) that learning from multiword units is beneficial *because* it increases the association between the article and the object, an association that is hard to establish when the mapping of objects into classes is arbitrary. In such cases, associating an object with an article–noun sequence will make the article more predictive of the noun than it would otherwise be. If this explanation is correct, then the effect of building blocks on learning should change when the article (or the grammatical element more generally) is predictive of the object in its own right, that is, when it carries semantic information. In these cases, multiword units are not needed to establish (or facilitate) the association between the article and the object.

Some classifier systems provide an example of a noun class system that reflects semantic distinctions. In Japanese for example, different classifiers are used for humans, animals, and inanimate objects (with subclasses of each type; Yamamoto & Keil, 2000). Even within grammatical gender systems where the division of nouns into classes is mostly arbitrary (as in many Romance languages, Corbett, 1991), there is a clear semantic component to the classification for a subset of nouns, mainly those having biological gender like boy, girl, man, woman. Interestingly, L2 learners seem to show a greater association between the article and noun for such nouns: they show similar processing patterns as native speakers only for nouns that have biological gender (but not for ones that do not). In these cases, they can use the gender information on the article to predict an upcoming noun (Lew-Williams & Fernald, 2009). These results are consistent with the prediction of a differential effect of building blocks on different grammatical relations, a prediction that has not yet been tested and will be tested in Experiment 2.

Experiment 2

Experiment 1 demonstrated an effect of building blocks on learning outcomes. Learning from unsegmented speech led participants to rely more on article–noun units, which resulted in better learning. The grammatical gender system in Experiment 1 was fully arbitrary: there were no phonological or semantic cues to class membership. In Experiment 2 we wanted to examine what happens in a system that assigns nouns to classes based on their semantic properties. In this study, nouns were divided into classes based on animacy (animate vs. inanimate). We use the same learning conditions (unsegmented-first vs. segmented-first) but we predict a different pattern of learning. We still expect participants to extract more article–noun 'words' in the unsegmented-first condition. However, we predict that since the grammatical element (i.e., the article) carries inherent semantic information, we will *not* find better learning in the unsegmented-first condition, despite a greater reliance on multiword units in learning.

Methods and materials

Participants

Forty-four students at the Hebrew University, who did not participate in the previous study, participated in the study for course credit or payment (30 females and 14 males). Their age ranged from 19 to 30 (mean age 23.6). All participants were native Hebrew speakers, with no history of speech, reading or hearing disability, ADD or ADHD. As in Experiment 1, participants were randomly assigned to one of the two conditions: segmented-first or unsegmented-first (22 in each condition).

Materials and procedure

In Experiment 2, we used the exact same procedure as in Experiment 1. There were only two differences in the materials. First, we replaced the article *se* with the article *si* to minimize similarity to the Hebrew demonstrative *ze*. If the difference between the two articles in Experiment 1 was driven by this similarity, it should not appear in Experiment 2. Second, and more importantly, the division of objects into noun 'classes' was not arbitrary, but semantically based: the two articles distinguished between animate and inanimate nouns. In Experiment 1, all 12 nouns referred to inanimate objects. In Experiment 2, half of the objects ($N = 6$) were animate (animals) and half inanimate ($N = 6$, the full list of nouns and their meaning can be found in Appendix B). In contrast to Experiment 1, class membership was now entirely semantic – nouns referring to animate objects appeared with the article *si*, while nouns referring to inanimate objects appeared with the article *fo*.

As in the previous study, all objects had high frequency and early Age-of-Acquisition labels in Hebrew (all animals names were ones learned by children under age five). All the novel labels had two syllables. As in Experiment 1, to ensure that, learning is not affected by the gender of nouns in Hebrew (the L1), the two noun 'classes' were matched on the gender of the nouns in Hebrew: each had an equal number of masculine and feminine nouns.

The procedure was identical to that of Experiment 1, with the addition of a generalization test after the two learning tests (the forced-choice task and the production task). Since now the mapping between nouns and objects was not arbitrary, we could assess participants' learning of the article–noun pairing not only for familiar nouns (ones that appeared during training), but also for novel items. If participants learned the semantically-based generalization, they should be able to select the correct article for novel objects, based on their animacy (select *si* for animate objects and *fo* for inanimate ones). The generalization test consisted of 12 forced-choice trials with novel nouns (half animate and half inanimate), all examining the ability to choose the correct article for a novel noun (see Appendix C for the full list of novel nouns and meaning).

Results

As in Experiment 1, we excluded participants who scored less than 2SD below the mean in the noun trials of the forced choice test. Three such participants were excluded from all reported analyses below (one in the unsegmented-first condition and two in the segmented-first condition).

Manipulation validation

Again, we first used the typing measure to ensure that participants learned the grammatical gender system and not just to map longer 'words' to objects. If they did, then they should have treated the article–noun as one word throughout the experiment (even after exposure to segmented sentences). Three participants (all in the unsegmented condition) shows this pattern, and were removed from further analyses, since we could not know that they were learning article–noun pairings, and not just long words. The analysis below was conducted on the remaining 38 participants.

We then tested our manipulation. As in Experiment 1, early exposure to unsegmented sentences led to more article–noun 'words': participants were more likely to treat the article–noun as one unit in the unsegmented-first condition compared to the segmented-first condition (unsegmented-first: $M = 35\%$, $SD = 34\%$; segmented-first: $M = 3\%$, $SD = 10\%$; $t(39) = 3.9$, $p < .001$). As in Experiment 1, early exposure affected building blocks. We next test the prediction that this will not affected learning in the same way when the article carries semantic information.

Forced-choice trials: familiar items

As in Experiment 1, we ran a mixed-effect logistic regression model, with learning condition (segmented-first vs. unsegmented-first), trial type (article vs. noun) and article-type (*fo* vs. *se*) as fixed effects. The model included random intercepts for participant and item, as well as by-participants random slopes for trial type, and by-item random slopes for trial type and condition. As predicted, and in contrast to Experiment 1, there was no effect of condition on learning (82% vs. 89%, $\beta = -0.3$, $SE = .24$,

$p > .2$) and no interaction between condition and trial type ($\beta = -0.06$, $SE = 0.1$, $p > .5$). Importantly, as can be seen in Fig. 5, participants in the unsegmented-first condition were not better than those in the segmented-first condition at selecting the correct article. Also in contrast to Experiment 1, no effect of article (*si* vs. *fo*) was found ($p > .2$), suggesting that the advantage of the article *se* in the previous experiment was indeed driven by its' similarity to the Hebrew demonstrative *ze*.

We ran two additional models separately on the article and noun trials with condition and article-type as fixed effects, and with by-item and by-subject random intercepts and by-item random slope for condition. These models also revealed no difference between the two conditions in selecting the correct article (79% vs. 85%, $\beta = -.25$, $SE = .27$, $p > .3$) or in selecting the correct noun (86% vs. 92%, $\beta = -.44$, $SE = .3$, $p > .1$). These results stands in contrast to the results of Experiment 1, and show that learning from unsegmented sentences does not lead to better performance when the article–noun pairing is semantic rather than arbitrary. Importantly, the lack of difference between the learning conditions does not reflect a ceiling effect: participants were not at ceiling for article (average 82% correct) or nouns (average 89% correct).

Forced-choice trials: novel items

This part asked if participants learned the animacy-based generalization. If they did, they should be able to correctly select an article for a novel object (not seen in training). We tested generalization by having participants select articles for 12 novel objects, half animate and half inanimate. We used a mixed-effect logistic regression model with learning condition (segmented-first vs. unsegmented-first) and article-type (*fo* vs. *si*) as fixed effects, and random intercepts for participant and item. As predicted, we did not find any difference in performance between the conditions: performance in the unsegmented-first condition did not differ from that of the segmented-first group (unsegmented-first: 80% vs. segmented-first: 85%, $B = -0.3$, $SE = .47$, $p > 0.5$). This result shows that there was no difference in the groups' ability to learn the generalization and apply it to novel items. Importantly, participants in both groups were able to apply the generalization to novel items (significantly above chance performance: $M = 83\%$, $SD = 23\%$, t-test $(37) = 8.6$, $p < .0001$). This shows that participants did not just learn the specific article–noun pairings presented in learning, but learned the animacy-based classification.

Production trials

A mixed-effect logistic regression analysis similar to the one reported for Experiment 1 was performed on the data of the production test, which is depicted in Fig. 6. Five subjects (three from the unsegmented-first condition and three from the segmented-first) were removed from this analysis as they did not understand the instructions properly and tried to type in the sentences instead of saying them. As predicted, there was no significant effect of condition on the production of articles (74% vs. 85%, $B = -1.75$, $SE = 1.15$, $p = .13$) or on the production of nouns (31% vs. 44%, $B = -0.86$, $SE = .56$, $p = .12$). Participants in

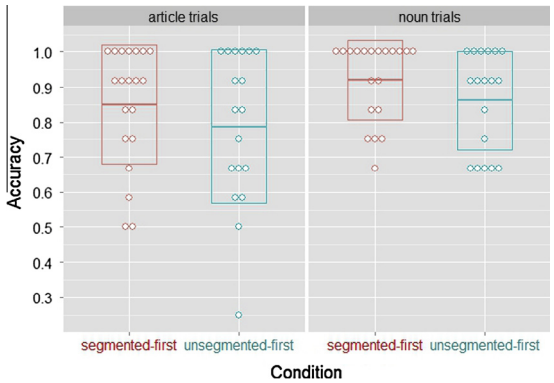

Fig. 5. Distribution of scores, means, and SDs in the forced-choice test by condition and trial type.

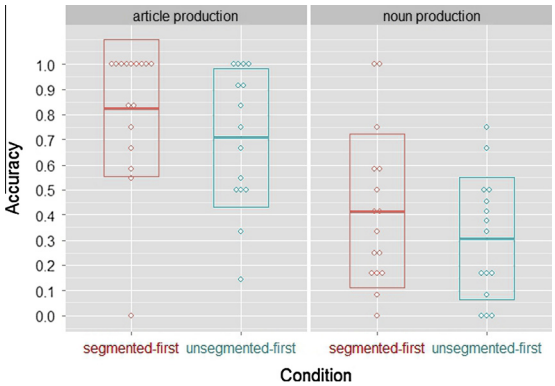

Fig. 6. Distribution of scores, means and SDs in the production test by condition and production type.

both conditions were equally likely to produce the article and the noun correctly. Note that they are much better at producing the articles compared to Experiment 1 – which is not surprising given that the semantic classification of nouns enabled subjects to produce the correct article based on the animacy of a given item. This kind of information was not available for participants in Experiment 1, who had to learn a completely arbitrary mapping of items to classes.

We also conducted the additional analysis reported above for Experiment 1, computing the ratio of correct articles given a correctly produced noun. This analysis focused only on trials where the noun was produced correctly ($N = 147$), and asked whether the article was produced correctly or not. Unlike Experiment 1, there was no difference between the conditions: Participants in the unsegmented-first condition were as likely to produce the correct article for a given noun as those in the segmented-first group (92% vs. 91%, Chi-square = 0.06, $p > .8$). Note that participants in both conditions were very accurate in producing the correct article, which is again not surprising given the semantic classification.

Typing trials

Following Experiment 1, we calculated the correlation between the likelihood to treat the article and noun as

one unit (of each individual) and the various learning measures. All correlations were smaller than ±0.2. Most importantly, in contrast to Experiment 1, we did not find any evidence for a correlation between the reliance on article–noun units and performance in the forced-choice article trials ($r = 0.05$, $p > .7$; see Fig. 7). This result means that in contrast with Experiment 1, individuals who relied more on multiword units were not better learners of the article–noun pairing compared to those who relied more on single words.

In sum, in contrast with Experiment 1, all learning measures (forced-choice for familiar and novel items, production, and typing) revealed no difference between the two learning conditions. Participants in the unsegmented-first condition did not show better learning than those in the segmented-first condition.

Comparing the results of Experiment 1 and 2

The two experiments share a very similar design: in both, participants learned an artificial language where nouns were classified into classes, and this classification determined article choice. In the two experiments, participants' early exposure was manipulated in a way that led them to learn more from multiword units (the unsegmented-first condition) or single words (segmented-first condition). The fundamental difference between the experiments was whether the classification of nouns into classes was arbitrary (Experiment 1) or semantic (Experiment 2). As predicted, participants in the unsegmented-first condition showed better learning of the article–noun pairings in Experiment 1, but not in Experiment 2. In order to directly compare the results, we conducted additional analyses using the data from the two studies combined. We focus on the measures looking at how well participants learned the article–noun pairing given this is the predicted difference between the studies.

Forced-choice trials

We ran a mixed-effect logistic regression model on accuracy in the article trials to see if the effect of learning conditions was different in the two studies. We had experiment (1 vs. 2), learning condition (unsegmented-first vs. segmented-first), and the interaction between them as fixed factors. The model included random intercepts for subject and item. Across the two experiments, the effect of condition was marginally significant ($B = .34$, $SE = .19$, $p = .07$). Importantly, and as predicted, we found a significant interaction between experiment and condition ($B = -.29$, $SE = .14$, $p < .05$): starting from unsegmented input facilitated learning only in the first experiment. The effect of experiment was also significant ($B = .37$, $SE = .14$, $p < .01$): unsurprisingly, it was easier to learn the article–noun pairings when they were animacy based (72% vs. 82% in Experiment 2).[2]

[2] We ran the exact same model on the noun trials, and as expected, none of the factors were significant. There was no effect of learning condition, experiment and no interaction between them (all p's > 0.1).

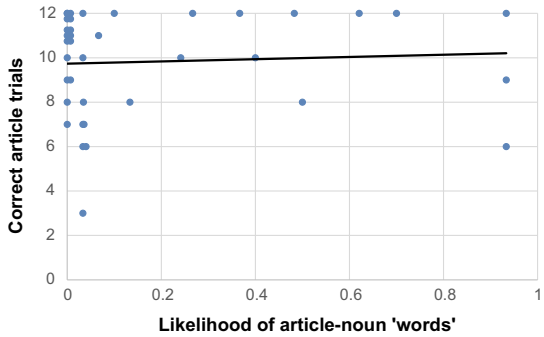

Fig. 7. The correlation between the tendency to treat the article and noun as one unit and success on forced-choice article trials.

Production trials

The comparison of the production trials revealed a similar pattern. Participants in Experiment 2 were better at producing the correct article than those in Experiment 1, reflecting its' semantic basis (79% vs. 54%, t-test (63) = 4.33, $p < .001$). More importantly, there was a different in the effect of learning condition on accuracy. Participants in Experiment 1, but not in Experiment 2, were more likely to produce a correct article given a correct noun in the unsegmented first condition (77% vs. 57%, Chi-square = 6.86, $p < .01$ in Experiment 1; 92% vs. 91%, Chi-square = 0.06, $p > .8$, in Experiment 2). This again reflects the difference in the influence learning condition in the two experiments.

Relation between building blocks and learning outcomes

Individuals who relied more on article–noun 'words' showed better learning of the article–noun pairing in Experiment 1 ($r = 0.45$), but not in Experiment 2 ($r = 0.05$). We used Fisher's r-to-z transformations to compare the magnitude of the effect. As predicted, the difference in the size of the correlation coefficients between Experiment 1 and Experiment 2 was significant ($Z = 1.79$, $p < .05$). A bigger reliance on article–noun units led to better learning of the article–noun pairing in Experiment 1, but not in Experiment 2.

Taken together, the results of the forced-choice, production and typing measures from the two experiments converge to demonstrate that manipulating early unit size facilitated learning of article–noun pairings when those pairings were arbitrary (Experiment 1), but did not affect learning when the pairings were animacy based (that is, when the article carried semantic information).

General discussion

This paper set out to examine the effect of existing linguistic knowledge on subsequent language learning. We hypothesized that adults' existing knowledge of words leads them to rely less on multiword units and, therefore, to worse learning outcomes in specific language-learning tasks. We tested this by examining the relation between early building blocks and learning outcomes, manipulating whether the initial exposure to an artificial language was segmented or not, and whether the relation to be learned

was arbitrary (Experiment 1) or semantically-based (Experiment 2). In line with our predictions, the studies showed that (a) learning from unsegmented input results in more multiword units, (b) learning from multiword units leads to better learning of arbitrary article–noun pairings manifested in better group performance and in a significant individual correlation between under-segmentation and better learning (Experiment 1), and (c) this advantage is affected by the information conveyed by the article, and is not found when the article carries semantic information (Experiment 2).

These results replicate and significantly extend previous findings on the way early units affect L2 learning outcomes (Arnon & Ramscar, 2012). Experiment 1 provides direct evidence for the relation between unit size and learning of grammatical gender on both a group and individual level. By assessing units of learning during training (using a typing measure), we show that exposure to unsegmented input leads to more multiword units and to better learning even though participants in this condition had to simultaneously segment and learn the pairings. The typing measure also ensures that participants were not just learning long words, but were instead learning about the relation between the article and noun: by the end of training, all participants differentiated between the article and the noun. Experiment 2 further examined the mechanisms underlying the advantage of learning from multiword units, and specifically, the prediction that multiword units are not needed to establish (or facilitate) the association between the article and the object when the article carries semantic information and is therefore predictive of the object. The results of Experiment 2 (which had an identical design to Experiment 1 except for the animacy-based classification of nouns) confirmed this prediction. Participants were still more likely to treat the article and noun as one unit when they were exposed first to unsegmented sentences (based on the online typing measure), but this had no effect on learning.

Finding an individual correlation between segmentation and learning outcomes raises interesting questions about the relation between individual differences in segmentation and learning of grammatical relations. Several previous studies have shown a *positive* relation between segmentation abilities and some aspects of language learning and use, both in L1 and L2. In infants, for instance, better segmentation of an artificial speech stream was correlated with larger vocabularies (Junge, Kooijman, Hagoort, & Cutler, 2012; Singh, Steven Reznick, & Xuehua, 2012) while in adults, performance on a similar task was correlated with better sentence comprehension (Misyak & Christiansen, 2012; Misyak, Christiansen, & Tomblin, 2010). This study shows the opposite effect: individuals who segmented less (treated the article and noun as one unit) were *better* learners of grammatical gender. How can this finding be reconciled with previous ones? One intriguing possibility is that the correlation between segmentation and learning depends on the domain to be learned: being good at identifying words enhances learning about word properties (labels, morphological features, as in previous studies), while being good at chunking frequently co-occurring words together improves learning about the relations between them (e.g., grammatical gender, as in the current one). If so, we predict that segmentation abilities will be negatively correlated with improved performance on L2 grammatical measures, yet may be positively correlated with other L2 measures such as vocabulary size.

Taken together, the results illustrate the differential effect of unit-size on learning grammatical relations – starting from multiword units facilitates learning in some cases, but not in others – and enhance the claim that multiword units facilitate learning *by* increasing the association between the grammatical element and the object. If so, their advantage will vary depending on the information the grammatical element conveys. Given this view of learning, we can now make concrete predictions about how multiword units will affect L2 learning. Learning from multiword units should be particularly beneficial for a host of grammatical relations that are (a) not semantically transparent in that the grammatical element does not make an independent semantic contribution (e.g., gender-marked articles), and (b) hold between adjacent words, such that the multiword unit will contain the grammatical element and the word it modifies.

Language presents us with such relations in the form of gender agreement patterns, certain classifier systems, and verb–preposition pairings, all of which are necessary for native-like fluency and all of which are difficult for L2 learners to master (grammatical gender: e.g., Grüter, Lew-Williams, & Fernald, 2012; subject–verb agreement: Jiang, 2004; classifiers: Hansen & Chen, 2001; see Clahsen, Felser, Sato, & Silva, 2010; DeKeyser, 2005, for reviews highlighting such aspects as challenging for L2 learners). We predict that manipulating learners' exposure in a way that directs them to multiword units should enhance performance in such domains. In contrast, a greater reliance on multiword units should not enhance learning vocabulary items or grammatical relations like case marking (when it is not accompanied by gender distinctions) or plural marking, all of whom seem to pose less difficulty for L2 learners (Birdsong, 2009; Johnson & Newport, 1989; see DeKeyser, 2012, for a review). These predictions, which are unique to the proposal we outlined here and seem to be consistent with patterns of L2 difficulty, need to be tested in future research.

Different accounts of L2 learning

More broadly, the studies presented here aim to offer a first step in articulating an experience-based account of L2 difficulty. Understanding why adults struggle in learning a second language is one of the most studied issues in language research (see DeKeyser, 2005; Mitchell, Myles, & Marsden, 2013 for comprehensive reviews). Multiple factors have been found to influence adult learning of a second language (cognitive factors, motivational factors, learning environment, and many more). Consequently, any comprehensive theory of L2 learning will have to incorporate multiple factors, explain how they interact, and how together, they can account for the complex pattern of L2 learning outcomes. As a first step, any account of L2 learning has to explain *why* adult learners generally

fail to reach native-like proficiency, *why* they find certain domains harder than others and *how* this is linked to their cognitive make-up and language-learning situation.

Here, we approach this long-standing puzzle by emphasizing the relation between the linguistic units employed in learning a novel language (early building blocks) and learning outcomes. In particular, we suggest that adults' existing knowledge of words leads them to rely less on multiword units in learning, and that this negatively impacts mastery of certain grammatical relations (but not others). This view naturally addresses two challenges that face theories of L2 learning: the challenge of how to link specific features of the learner to the learning outcomes observed, and the challenge of why L2 learners find certain linguistic domains harder than others. Our proposal predicts differential performance on varying aspects of the novel language, and suggests that learning outcomes can be changed (to some degree) by manipulating learners' early input in a way that changes the units used during learning. Moreover, the proposal provides an initial set of testable constrains and limitations on the role of multiword units in learning, an important step in the development of any theoretical framework.

Importantly, this proposal is not intended to provide a comprehensive theory of L2 learning. It does not explain many patterns of L2 learning (e.g., the difference between phonology and syntax, Granena & Long, 2013), and does not incorporate factors known to affect learning (e.g., adults' ability to draw on declarative learning mechanisms, Ullman, 2001). Instead, the proposal aims to demonstrate the effect of an additional and novel factor on L2 learning: the way prior knowledge of words leads to fewer multiword units and impacts learning outcomes. The idea that adults may differ from children in the units from which they learn, and that this impacts learning outcomes builds on Newport's seminal less-is-more hypothesis (Newport, 1988, 1990) and Elman's starting small approach (Elman, 1993), both of which highlight the link between the way new input is represented and what can be learned from it (see also, Monner, Vatz, Morini, Hwang, & DeKeyser, 2013). Our account differs from previous ones in emphasizing the role of larger, multiword units in the process of learning.

Multiword units in language learning

The finding that multiword building blocks can facilitate learning has important theoretical implications. First, it illustrates the effect of early experience on subsequent learning: knowledge of word boundaries was found to affect the seemingly unrelated domain of grammatical gender, with unsegmented speech serving as a better starting point for learning. Second, the findings provide novel support for the idea that learning from multiword units is beneficial, an idea originally proposed for explaining first language acquisition (Abbot-Smith & Tomasello, 2006). Finally, while not tested directly, the results are also consistent with the claim that adults' greater difficulty in learning novel grammatical systems is related, at least to some extent, to their smaller reliance on multiword building blocks.

The idea that children make use of multiword units in learning language, and that adults are less likely to do so, is consistent with several bodies of findings. Multiword units are seen as important building blocks in usage-based models of first language learning, where grammatical knowledge is learned by abstracting over stored exemplars (Abbot-Smith & Tomasello, 2006; Lieven & Tomasello, 2008). In such approaches, children are predicted to use both words and multiword units in learning grammar. Analyses of children's early speech provide support for this prediction: children's early multiword utterances include many "frozen" sequences (up to 50%, e.g., give-me, get-it-out) whose components are not used productively in other combinations (Lieven, Pine, & Baldwin, 1997; Lieven, Salomo, & Tomasello, 2009); similar patterns are found in computational analyses of child speech (Borensztajn, Zuidema, & Bod, 2009) and in studies of children's error patterns (e.g., Kirjavainen, Theakston, & Lieven, 2009). Recent work provides more direct evidence for the role of multiword units in learning (Arnon, McCauley, & Christiansen, 2015, in preparation) by showing that, like words, multiword phrases also show Age-of-Acquisition effects. Phrases that are acquired earlier in childhood show processing advantages in adult speakers, illustrating their role as units of learning.

Unlike the reports of 'chunked' production in child language, the language of L2 learners is often characterized as non-formulaic, and is fraught with non-native idioms and collocations (e.g., *put more attention to*, *watch up*, Conklin & Schmitt, 2012; Pawley & Syder, 1983; Wray, 2002, 2008). Even advanced learners produce fewer formulaic sequences than native speakers in both written and spoken production (e.g., Howarth, 1998). Even when adults do acquire multiword sequences, they seem to use them differently in learning, memorizing a few fixed sequences (greetings, etc.) but not using them to further grammatical development (Wray, 1999; Wray & Fitzpatrick, 2008). Interestingly, unlike adults, children learning a second language use multiword chunks early on, and seem to use them to further grammatical development (e.g., Hakuta, 1974; Karniol, 1990). In a particularly detailed study, Wong Fillmore (1976, 1979) followed five Spanish speaking preschoolers learning English in an immersion setting (following immigration) for a year. The study documented extensive use of multiword chunks in early production, a pattern that stands in contrast with the ones reported for adult L2 learners. In sum, analyses of L2 and child speech point to a possible difference between them in the reliance on multiword units in learning.

This idea receives further support from recent computational work that underscores the difference in building blocks between children, adult native speakers, and L2 learners (Christiansen & McCauley, 2013). Christiansen and McCauley (2013) use the Chunk-Based-Learner (CBL, McCauley & Christiansen, 2011, 2014) to model language learning. This model provides a computational implementation of children's use of chunking in language acquisition. The model gradually builds an inventory of chunks consisting of one or more words—a "chunkatory"—used for both language comprehension and production. The same model can be used to compare unit size – operationalized as the proportion of multiword units in the "chunkatory" – between different speakers and populations. Consistent

with our predictions, the productions of L2 learners of German and English are less 'chunked' than matched productions by child and adult native speakers of those two languages (Christiansen & McCauley, 2013). Taken together, the findings reviewed provide support for the idea that infants and adults do differ in their reliance on multiword units, and that this may explain (some of) the differences between them. Much more research is needed, however, to substantiate this claim.

Learning multiple levels of language simultaneously

The results also raise interesting questions about the possible benefit of learning segmentation, meaning and grammar simultaneously. Participants in the unsegmented condition had to segment the speech while learning its components. Their better performance points to the possible advantage in learning multiple levels of language at the same time (in this case, segmentation and grammar). Learning from this kind of multilayered stimuli may also be more similar to what infants do during actual language acquisition (Romberg & Saffran, 2013; Sahni, Seidenberg, & Saffran, 2010). Nonetheless, much research in the field of language learning typically examines each process in isolation and in the absence of information from other levels (segmentation: e.g., Newport & Aslin, 2004; rule-learning: e.g., Marcus, Vijayan, Rao, & Vishton, 1999; word learning: e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998), ignoring the possible interactions between the different levels. However, recent work has explored the interactions between different levels of information. In the domain of word learning and segmentation, studies show that segmentation and word–object mapping can be learned in parallel after just a brief exposure, and that exposure to visual objects facilitates segmentation (Cunillera, Camara, Laine, & Rodriguez-Fornells, 2010). Similarly, recent studies document the role of feedback between levels and the importance of big to small processes in phonetic learning: It was shown that both adults and infants use word-level information in forming phonetic categories, by assigning vowels to different categories when they appear consistently in different words (Feldman, Griffiths, Goldwater, & Morgan, 2013; Feldman, Myers, White, Griffiths, & Morgan, 2013). The current study joins existing ones in highlighting the importance of learning multiple domains simultaneously and provides novel evidence for the beneficial interaction between segmentation and grammatical learning.

In sum, we have provided evidence for the claim that early exposure to unsegmented input makes learners more likely to treat the article and noun as one unit, which in turn leads to better learning of article–noun pairing in certain cases, but not others. These findings illustrate the effect of experience on building blocks and the advantage of learning certain grammatical relations from multiword units. The findings highlight the complex interaction between experience, building blocks, and learning outcomes, offering a new perspective as to why adults struggle with learning a second language. From an applied perspective, our proposal has the potential to significantly alter how we teach adults a second language. By investigating the way early units affect learning outcomes, and highlighting the role of larger chunks in learning, we may gain better understanding of the process of second language learning, and the way it differs from first language learning.

Acknowledgments

We would like to thank Alex B. Fine for statistical advice and extensive feedback, Ram Frost, Stewart M. McCauley & Morten H. Christiansen for helpful discussions, and Tamar Johnson for her help in conducting the experiments. This work was supported by ISF grant number 527/12 awarded to IA.

Appendix A. Items used in Experiment 1

1. Carrier phrase: *os-ferpel-ti*
2. Articles: *se*, *fo*
3. Nouns:
   3.1. Nouns following the article *fo*:

Word	Meaning
gorok	bike
panjol	television
toonbot	clock
fertsot	sock
perdip	house
etkot	key

   3.2. Nouns following the article *se*:

Word	Meaning
hekloo	bath
hertin	iron
geesoo	hat
slindot	piano
jatree	cup
sodap	spoon

Appendix B. Items used in Experiment 2

1. Carrier phrase: *os-ferpel-ti*
2. Articles: *si*, *fo*
3. Nouns:
   3.1. Nouns following the article *fo*:

Word	Meaning
gorok	bike
panjol	television
toonbot	clock
fertsot	sock
perdip	house
etkot	key

   3.2. Nouns following the article *se*:

Word	Meaning
hekloo	camel
hertin	cow
geesoo	lion
slindot	hen
jatree	sheep
sodap	turtle

Appendix C. Novel Items used in the generalization test in Experiment 2

Word	Meaning
romdee	bath
nemlik	iron
sidpal	hat
tilmoo	piano
famtog	cup
bagdee	spoon
fitloo	donkey
orteep	dove
dalku	parrot
pugtee	ladybug
gertom	snail
pridmos	ant

References

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

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech, 56*, 349–373.

Arnon, I., McCauley S.M., Christiansen, M.H. (2015). Age-of-Acquisition effects for multiword phrases, *poster presented at the International Workshop on Statistical Learning*, BCBL, June 2015.

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.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning. *Psychological Science, 19*, 241–248.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

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

Birdsong, D. (2009). Age and the end state of second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *The new handbook of second language acquisition*. Amsterdam: Elsevier.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*, 341–345.

Borensztajn, G., Zuidema, W., & Bod, R. (2009). Children's grammars grow more abstract with age. *Topics in Cognitive Science, 1*, 175–188.

Bybee, J. L. (1998). The emergent lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Proceedings of the Chicago Linguistic Society 34: The panels* (pp. 421–435). Chicago: Chicago Linguistic Society.

Christiansen, M. H., & McCauley, S. M. (2013). Computational investigations of multiword chunks in language learning. *Paper presented at a symposium on multiword sequences as building blocks for language: Insights into first and second language learning at the 35th annual conference of the Cognitive Science Society, Berlin, Germany*.

Christiansen, M. H., & Chater, N. (2015). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Clahsen, H., Felser, C., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning, 60*, 21–43.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45–61.

Corbett, G. G. (1991). *Gender*. Cambridge, UK: Cambridge University Press.

Cunillera, T., Camara, E., Laine, M., & Rodriguez-Fornells, A. (2010). Speech segmentation is facilitated by visual cues. *The Quarterly Journal of Experimental Psychology, 63*, 260–274.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*, 465–480.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*, 1–25.

DeKeyser, R. M. (2012). Age effects in second language learning. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 442–460). London: Routledge.

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition, 35*, 353–387.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*, 164–194.

Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition, 32*(4), 553–580.

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

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science, 33*, 547–582.

Feldman, N. H., Griffiths, T. L., Goldwater, S., & Morgan, J. L. (2013a). A role for the developing lexicon in phonetic category acquisition. *Psychological Review, 120*, 751–778.

Feldman, N. H., Myers, E. B., White, K. S., Griffiths, T. L., & Morgan, J. L. (2013b). Word-level information influences phonetic learning in adults and infants. *Cognition, 127*, 427–438.

Ferman, S., & Karni, A. (2010). No childhood advantage in the acquisition of skill in using an artificial language rule. *PLoS ONE, 5*, e13648.

Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition, 108*, 477–499.

Finn, A. S., & Hudson Kam, C. L. (2015). Why segmentation matters: Experience-driven segmentation errors impair "morpheme" learning. *Journal of Experimental Psychology: Language, Memory, and Cognition*.

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*, 311–343.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*, 191–215.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning, 24*, 287–297.

Hansen, L., & Chen, Y. L. (2001). What counts in the acquisition and attribution of numeral classifiers? *JALT Journal, 23*, 90–110.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*, 24–44.

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.

James, W. (1890). *The principles of psychology* (Vol. I). New York: Holt.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.

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.

Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science, 15*, 463–473.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. Campbell & R. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts.

Karniol, R. (1990). Second language acquisition via immersion in daycare. *Journal of Child Language, 17*, 147–170.

Kirjavainen, M. M. M., Theakston, A. L., & Lieven, E. V. (2009). Can input explain children's me-for-I errors? *Journal of Child Language, 36*, 1091–1114.

Krashen, S. D., Long, M. A., & Scarcella, R. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly, 13*, 573–582.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science, 97*, 11850–11857.

Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.

Lew-Williams, C., & Fernald, A. (2009). Fluency in using morphosyntactic cues to establish reference: How do native and non-native speakers differ. In *Proceedings of the 33rd annual Boston University conference on language development*.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*, 193–198.

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*, 447–464.
Lieven, E. V., Pine, J., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24*, 187–220.
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.
Lieven, E., & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 168–196). New York and London: Routledge.
Luniewska, M. (2014). Age of acquisition norms for nouns and verbs in 22 languages. *Poster presented at the International Association for the Study of Child Language, Amsterdam*.
Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80.
McCauley, S. M., & Christiansen, M. H. (2014). Acquiring formulaic language: A computational model. *The Mental Lexicon, 9*, 419–436.
McCauley, S. M., & Christiansen, M. H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 1619–1624). Austin, TX: Cognitive Science Society.
McClelland, J. L. (2010). Emergence in cognitive science. *Topics in Cognitive Science, 2*, 751–770.
Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62*, 302–331.
Misyak, J. B., Christiansen, M. H., & Tomblin, B. J. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science, 2*, 138–153.
Mitchell, R., Myles, F., & Marsden, E. (2013). *Second language learning theories*. New York: Routledge.
Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKeyser, R. (2013). A neural network model of the effects of entrenchment and memory development on grammatical gender learning. *Bilingualism: Language and Cognition, 16*, 246–265.
Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age motivation and instruction. *Studies in Second Language Acquisition, 21*, 81–108.
Neville, H. J., & Bavelier, D. (2001). Variability of developmental plasticity. In J. McClelland & R. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives*. Riverton, NJ: Foris.
Newport, E. L. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences, 10*, 147–172.
Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.
Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127–162.
Onnis, L., & Thiessen, E. (2013). Language experience changes subsequent learning. *Cognition, 126*, 268–284.
Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics, 7*, 551–579.
Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of 'mouses' in adult speech. *Language, 89*, 760–793.
Ramscar, M., Dye, M., Popick, H. M., & O'Donnell-McCarthy, F. (2011). The enigma of number: Why children find the meanings of even small number words hard to learn and how we can help them do better. *PLoS ONE, 6*, e22501.
Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Science, 11*, 274–279.
Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science, 34*, 909–957.
Reali, F., & Christiansen, M. H. (2007). Word-chunk frequencies affect the processing of pronominal object-relative clauses. *Quarterly Journal of Experimental Psychology, 60*, 161–170.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science, 37*, 1290–1320.
Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research, 22*, 1–29.
Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development, 81*, 727–736.
Scherag, A., Demuth, L., Roesler, F., Neville, H. J., & Roeder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition, 93*, B97–B108.
Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science, 15*, 482–495.
Slobin, D. I. (1985). *The crosslinguistic study of language acquisition: The data* (Vol. 1). NJ: Erlbaum.
Sorace, A. (2003). Near-nativeness. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 130–151). Oxford, UK: Blackwell.
Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38*, 598–605.
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.
Toro, J. M., Pons, F., Bion, R. A., & Sebastián-Gallés, N. (2011). The contribution of language-specific knowledge in the selection of statistically-coherent word candidates. *Journal of Memory and Language, 64*, 171–180.
Tremblay, A., & Baayen, H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151–173). London: The Continuum International Publishing Group.
Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and cognition, 4*, 105–122.
Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science, 12*, 419–425.
Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5.
Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289.
Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75*, 1866–1878.
Wong Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Unpublished PhD dissertation, Stanford University, Stanford, CA.
Wong Fillmore, L. (1979). Individual differences in second language acquisition. In C. J. Fillmore, D. Kempler, & S-Y. W. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 203–228). New York: Academic Press.
Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching, 32*(4), 213–231.
Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.
Wray, A., & Fitzpatrick, T. (2008). Why can't you just leave it alone? Deviations from memorized language as a gauge of native-like competence. In F. Meunier & S. Granger (Eds.), *Phraseology in language learning and teaching* (pp. 123–148). Amsterdam: John Benjamins.
Yamamoto, K., & Keil, F. (2000). The acquisition of Japanese numeral classifiers: Linkage between grammatical forms and conceptual categories. *Journal of East Asian Linguistics, 9*, 379–409.