

Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness

Patricia J. Brooks · Vera Kempe

Published online: 9 October 2012
© Psychonomic Society, Inc. 2012

Abstract In this study, we sought to identify cognitive predictors of individual differences in adult foreign-language learning and to test whether metalinguistic awareness mediated the observed relationships. Using a miniature language-learning paradigm, adults ($N = 77$) learned Russian vocabulary and grammar (gender agreement and case marking) over six 1-h sessions, completing tasks that encouraged attention to phrases without explicitly teaching grammatical rules. The participants' ability to describe the Russian gender and case-marking patterns mediated the effects of nonverbal intelligence and auditory sequence learning on grammar learning and generalization. Hence, even under implicit-learning conditions, individual differences stemmed from explicit metalinguistic awareness of the underlying grammar, which, in turn, was linked to nonverbal intelligence and auditory sequence learning. Prior knowledge of languages with grammatical gender (predominantly Spanish) predicted learning of gender agreement. Transfer of knowledge of gender from other languages to Russian was not mediated by awareness, which suggests that transfer operates through an implicit process akin to structural priming.

Keywords Language acquisition · Individual differences · Language production · Syntactic processing · Psycholinguistics

P. J. Brooks (✉)
Department of Psychology, College of Staten Island,
City University of New York,
2800 Victory Boulevard 4S-108,
Staten Island, NY 10314, USA
e-mail: patricia.brooks@csi.cuny.edu

V. Kempe
University of Abertay,
Dundee, Scotland

In first language acquisition, learning of sequentially distributed linguistic patterns is implicit: Learners come to distinguish grammatical from ungrammatical patterns without any explicit instruction, and usually without awareness of the underlying grammar. In contrast, second language (L2) learners often fail to discover grammatical patterns on their own in the absence of formal instruction. Over the past decade, our research program has examined individual differences in adult language learning under implicit conditions in which learners are given exposure to spoken dialogues in a foreign language (Russian) but no information about the underlying patterns in the input (i.e., the grammatical gender categories and case-marking paradigms). Our miniature language-learning paradigm was designed to emulate naturalistic language learning of richly inflected languages in which the acquisition of meaning and morphology are inseparable, due to the fact that uninflected content words do not exist. After several sessions of training on a representative set of phrases embedded in dialogues (referring to contexts instantiated using pictures), learners were asked to generalize what they had learned to a new set of items (i.e., to produce phrases that were not in the training input in response to new contexts). Generalization performance was linked to a number of factors, including prior exposure to languages with similar patterns, verbal working memory capacity, and nonverbal intelligence (Brooks, Kempe, & Sionov, 2006; Kempe & Brooks, 2008; Kempe, Brooks, & Kharkhurin, 2010). However, these studies failed to account for two potentially important predictors of adult language learning: learners' ability to extract patterns from sequential input and their awareness of the extracted patterns. In the present study, we aimed to fill this gap by adding auditory sequence learning and metalinguistic awareness of the underlying patterns to the set of predictor variables. We were specifically interested to see whether gaining explicit awareness of the underlying patterns would mediate the effects of

the other cognitive predictors. The results would help to clarify which learner variables facilitate explicit versus implicit learning of foreign languages by adults.

Sequence learning

Language learning requires sensitivity to patterns such as the grammatical (morphological and syntactic) regularities that govern how word forms are modified and combined to form larger constructions. These patterns are hierarchical and distributed over temporal sequences of linguistic units: Each linguistic unit is perceived or produced as part of a sequence, with one unit processed at a time. To acquire a language, learners must be capable of tracking how linguistic units are arranged in sequences; this requires sensitivity to adjacent dependencies between consecutive elements, as well as to nonadjacent dependencies spanning multiple units. For example, to acquire vocabulary, learners have to remember the order of phonemes and syllables within words or phrases; to acquire grammatical patterns, such as subject–verb agreement (e.g., *I am sleepy, He is sleepy; That child always wants ice cream, Those children always want ice cream*), learners must keep track of longer-distance dependencies between members of grammatical categories.

While sequence learning has been linked to individual differences in first language acquisition (Estes, Evans, & Else-Quest, 2007; Evans, Saffran, & Robe-Torres, 2009; Lum, Gelgic, & Conti-Ramsden, 2010; Plante, Gómez, & Gerken, 2002; Tomblin, Mainela-Arnold, & Zhang, 2007) as well as to adult sentence processing (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Misyak & Christiansen, 2007; Misyak, Christiansen, & Tomblin, 2010; Mueller, 2006), its role in explaining individual differences in L2 learning has yet to be established, as existing studies using a variety of foreign language outcome measures have generated equivocal findings (Frost, Narkiss, Afek, & Siegelman, 2011; Kaufman et al., 2010; Robinson, 2005). Sequence learning is typically measured using the artificial-grammar-learning task, in which participants make grammaticality judgments about visual or auditory sequences of elements (e.g., letters or nonsense syllables) after being exposed to a representative set of such sequences during a training phase. A related task, the serial reaction time task, measures the time that it takes for participants to complete sequences, with most variants using visual stimulus presentation (but see Misyak et al., 2010, for a novel task variant utilizing auditory and visual presentation of sequences of nonsense syllables).

Both artificial-grammar and serial reaction time tasks are based on sequences generated by an underlying grammar—often a finite-state grammar with rules that are typically too complex to be accessible to explicit awareness (Reber, 1989). Robinson (2005) found no evidence of a link

between an artificial-grammar-learning task (using sequences of printed letters) and performance on an L2 (Samoan) grammaticality judgment task, with the two tasks showing different patterns of correlations with measures of intelligence, working memory, and language-learning aptitude. Kaufman et al. (2010), however, found performance on a serial reaction time task to be predictive of individual differences in academic performance on foreign-language exams. These contradictory findings suggest the need for further investigations of the relationship between sequence learning and individual differences in foreign-language learning. Furthermore, there is evidence that sequence-learning abilities may be constrained by modality (Conway & Christiansen, 2005, 2006, 2009), with auditory input supporting better learning of the ends of sequences relative to visual input (Conway & Christiansen, 2005, 2009); such findings are especially relevant for the learning of grammatical morphemes positioned at the ends of words and phrases. Therefore, we selected for our study an auditory sequence-learning task that had been shown to be sensitive to individual differences in adult sentence processing (Misyak & Christiansen, 2007).

Metalinguistic awareness of grammatical patterns

The other factor that had not been considered in our previous research on implicit learning of grammatical patterns was whether the learning actually remained implicit or became explicit through the learners' noticing of regularities. The aim of the present study was therefore to examine whether learners' emergent metalinguistic awareness of the underlying patterns contributed to explaining the observed individual differences in their ability to generalize. In other words, is generalization performance in adult learners linked to their ability to describe what they have noticed in the input? Schmidt (1990, 1993) and others (e.g., Ellis, 2005; Lemhöfer, Schriefers, & Hanique, 2010; Lemhöfer, Schriefers, & Indefrey, 2011) have suggested that adult language learners must notice grammatical patterns in order to apply them productively. Robinson (1995, 1996) has further suggested that individual differences in verbal working memory capacity affect the extent of noticing. Thus, as a test of Schmidt's (1990, 1993) noticing hypothesis, we asked participants at the end of the present study to reflect on what patterns, if any, they had noticed. We hypothesized that participants' emergent awareness of Russian gender agreement and case marking would predict their learning and generalization of the patterns. In addition, as a further test of the role of metalinguistic awareness in adult language learning, we explored whether learners' noticing of the grammatical patterns mediated the effects of other cognitive predictors on learning outcomes.

The present study

We used a miniature language-learning task to expose participants to Russian gender agreement and case marking under implicit-learning conditions without formal instruction. In our task, we examined how learners extract patterns of long-distance agreement—that is, between adjectives and nouns agreeing in gender, or between prepositions and nouns marked for case—on the basis of limited input; this topic had previously been explored in an artificial-language study by Gómez (2002). In contrast to the artificial-language methodology, we used a naturalistic language-learning task in which participants were exposed to meaningful dialogues. Following the procedures of our prior investigations, we exposed participants to Russian phrases that were paired with corresponding pictures. To illustrate gender agreement, we used Russian adjective–noun phrases paired with colored objects; Russian color adjectives have distinct endings when modifying masculine and feminine nouns, and thus exemplify gender agreement patterns (cf. Kempe & Brooks, 2001). For case marking, we presented phrases utilizing dative and genitive case marking, as these two cases have distinct vowel endings for masculine and feminine nouns and can be depicted as the relationships “going toward” (dative) and “going away from” (genitive; cf. Brooks et al., 2006). As in previous studies, we measured a set of cognitive abilities (described below) to predict learning outcomes and generalization performance. These cognitive measures were selected on the basis of existing literature documenting relationships to adult language-learning outcomes, which we will briefly review below.

Phonological short-term memory capacity The ability to maintain phonological information in short-term memory has been shown to be predictive of word learning and vocabulary size in children (Gathercole & Baddeley, 1989), as well as foreign-language vocabulary acquisition in adults (Gupta, 2003; Papagno, Valentine, & Baddeley, 1991; Service, 1992; Service & Kohonen, 1995; Speciale, Ellis, & Bywater, 2004). To measure phonological short-term memory, most researchers have used a nonword repetition task requiring participants to immediately recall nonsense words following auditory presentation. Evidence for a link between nonword repetition and vocabulary acquisition supports the proposal that the phonological loop might serve as a language acquisition device (Baddeley, Gathercole, & Papagno, 1998) by allowing individuals to maintain sequences of phonemes or syllables in memory through subvocal articulation. Ellis and Sinclair (1996) have provided evidence that articulatory rehearsal of linguistic sequences promotes their consolidation: Adult language learners who were allowed to practice foreign-language utterances through articulatory rehearsal showed better learning of L2

vocabulary and syntax, as well as greater explicit metalinguistic knowledge of the grammatical regularities, than did learners who were not able to rehearse because of an articulatory suppression procedure. Williams and Lovatt (2003) found phonological short-term memory to be predictive of vocabulary learning, of memory for sequences and combinations of forms in the input, and of the eventual abstraction of grammatical rules based on the form co-occurrence patterns. Similarly, Ellis and Schmidt (1998) showed phonological short-term memory for linguistic utterances to be highly correlated with long-term memory for the same utterances; both of these memory measures predicted the learning of long-distance morphosyntactic dependencies.

Verbal working memory capacity A related proposal is that verbal working memory capacity underlies L2 vocabulary acquisition. Verbal working memory is thought to comprise both storage and processing functions; it is often measured using the reading span task, which requires participants to hold words in mind while reading additional lines of text (Daneman & Carpenter, 1980). Our research has linked verbal working memory capacity to the acquisition of foreign-language vocabulary and grammar—especially those aspects of grammar that are irregular (Kempe & Brooks, 2008; Kempe et al., 2010). Other research has found reading span to be predictive of the comprehension of structurally complex sentences in an L2 (Harrington & Sawyer, 1992; Miyake & Friedman, 1998).

Nonverbal intelligence General intelligence is considered to be one aspect of adult foreign-language learning aptitude (Grigorenko, Sternberg, & Ehrman, 2000). Our research, using the Culture Fair Test of nonverbal intelligence (Cattell & Cattell, 1973), has revealed that nonverbal intelligence predicts Russian grammar learning over and above the effects verbal working memory capacity (Brooks et al., 2006; Kempe & Brooks, 2008; Kempe et al., 2010). The Culture Fair Test requires participants to examine visual–spatial arrays in order to complete patterns; participants are given only a few minutes to complete each of four subtests—hence, the task requires rapid extraction of patterns. Performance on this test tends to be moderately correlated with measures of verbal working memory capacity (Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Unsworth, Brewer, & Spillers, 2009). This suggests that the relationship between verbal working memory capacity and adult L2 learning may be due in part to variance components shared with nonverbal intelligence, such as the ability to sustain focused attention. The fact that the Culture Fair Test predicts additional variance in adult language learning over verbal working memory capacity has been interpreted, in accordance with Garlick and Sejnowski (2006), as indicating a role for pattern

perception and abstraction, in addition to the control of attention, in language learning.

Transfer of prior knowledge of languages A final factor linked to individual differences in L2 learning is prior experience with foreign languages, which provides opportunities to transfer knowledge of specific grammatical features (e.g., gender or case marking) to a new language containing similar structures (Hernandez, Li, & MacWhinney, 2005; Williams, 2005; Williams & Lovatt, 2003). Transfer of knowledge from a previous language to a new one may be akin to priming, such that available structures from a familiar language may facilitate recognition of their counterparts in a foreign language, often without conscious awareness (MacWhinney, 2008; Salamoura & Williams, 2007). Because structural priming is generally thought to involve implicit processing (Bock, Dell, Chang, & Onishi, 2007; Chang, Dell, Bock, & Griffin, 2000; Savage, Lieven, Theakston, & Tomasello, 2006), transfer might occur without conscious awareness of the underlying similarities between the new language and the previously learned language(s). In the present study, we tested this hypothesis by examining whether transfer effects are mediated by conscious awareness.

Method

Participants

A group of 77 students (50 women, 27 men; mean age 23 years, range 17–42) were recruited using flyers at the College of Staten Island, City University of New York. Participants completed six 1-h language-learning sessions conducted in a psychology laboratory and were paid \$10/h as compensation for their time. The sample was ethnically diverse and was representative of the student population of the City University of New York. All of the participants had native or near-native proficiency in English; 16 participants (20.8 %) also reported having native or near-native proficiency in an additional language. The recruitment procedures were established in accordance with guidelines provided by the Institutional Review Board of the City University of New York. None of the participants had previously learned Russian or any Slavic or Baltic language.

Materials

A set of 16 masculine and 16 feminine Russian nouns served as the stimuli. All nouns were bisyllabic, to minimize variation in ease of pronunciation across noun genders (i.e., it is possible to construct a set of monosyllabic masculine nouns, but not of feminine nouns). In each gender category, half of the nouns had diminutive suffixes (*-ik* for masculine and *-ka* for

feminine).¹ These nouns thus shared a higher degree of morphophonological similarity in noun endings than did the remaining simplex nouns. The results of this manipulation are presented elsewhere (Brooks, Kempe, & Donachie, 2011). A complete list of the stimuli used in the training and test sessions is provided in the [Appendix](#). A total of 12 masculine and 12 feminine nouns were used in the training sessions; each trained noun was presented in two of the three contexts for eliciting dative and genitive case marking and adjective–noun gender agreement. For each trained noun, the third context was reserved for the test session in order to assess participants' ability to generalize case-marking and gender agreement patterns (i.e., to productively generate phrases that they had not heard during training). These reserved items are marked in bold in the [Appendix](#). To double the number of generalization trials without increasing the size of the training vocabulary and the length of the training sessions, eight additional nouns were used only in the test session and were presented in all three contexts. These additional generalization items are also marked in bold in the [Appendix](#).

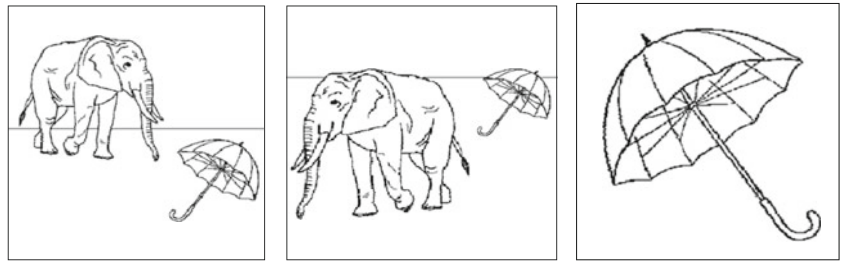
All nouns, when presented in the nominative case, were transparent with respect to the link between morphophonological form and gender. Thus, in the nominative case, all masculine nouns ended in a consonant, whereas all feminine nouns ended in /a/. Each noun was associated with a corresponding line drawing taken from the Snodgrass and Vanderwart (1980) set of standardized pictures. Line drawings of each object were presented in conjunction with a picture of an elephant walking toward the object (for dative case) or away from the object (for genitive case), and as a red or blue object (for adjective–noun gender agreement). In [Fig. 1](#), we present an example set of pictures.

Procedure

Participants were tested individually and completed six sessions of training, followed by a session of testing. Presentation of the task instructions and stimuli was controlled using PsychoScope experimental software (Cohen, MacWhinney, Flatt, & Provost, 1993) run on an Apple computer. During training and testing, all Russian phrases were presented through

¹ Some of the masculine diminutive nouns were productively derived from simplex forms (e.g., the diminutive *stolik* [table] is derived from the simplex form *stol*). Other masculine diminutives, and all of the feminine diminutives, were so-called *frozen* forms for which the corresponding simplex form either does not exist or has a different meaning. Although the distinction between frozen and productive diminutives is important when trying to ascertain the productivity of the derivation (e.g., in child language development), it was not relevant in this study, as our participants were completely naïve to the structure and frequency of the diminutive or of any other derivation in Russian. All that mattered was that the diminutive nouns shared the same endings.

Fig. 1 Example set of pictures for elicited production (respectively) of dative and genitive case-marking inflections and of adjective–noun gender agreement. Note that the last picture would have appeared in either red or blue



headphones along with the corresponding pictures; participants never saw any words or phrases written in Russian and were never provided with any written translations of the Russian. Throughout the experiment, an experimenter sat with the participant and manually advanced the trials so as to allow the participant as much time as needed to make a response. Experimenters did not speak any Russian and did not provide any feedback regarding response accuracy. Because of the complexity of the procedure, involving multiple blocks with different learning tasks, experimenters answered procedural questions after each participant had read the task instructions on the computer screen. The experimenters also provided nonspecific encouragement (e.g., “You are doing just fine”) at the end of each block of trials.

Training session The six training sessions were administered within a span of 14 days and lasted 45–60 min each. Each training session comprised four blocks utilizing different tasks (listen and repeat, noun comprehension, location/color comprehension, and production). Table 1 presents an example trial for each of the tasks (Blocks 1–4).

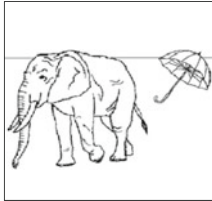
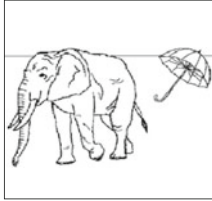
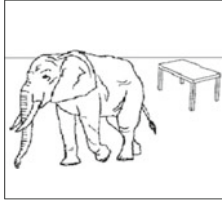
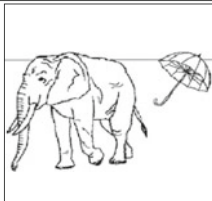
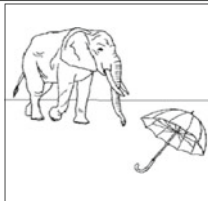

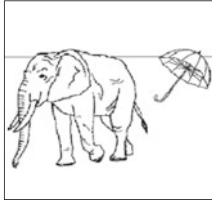
The training tasks were designed to engage learners in different activities while receiving exposure to phrases exemplifying dative and genitive case-marking patterns and adjective–noun gender agreement with nominative nouns. Specifically, Blocks 1–3 (see below) were designed to provide exposure to the materials, whereas Block 4 (production) was designed to probe the learners’ mastery of the system. All participants reached ceiling-level performance on comprehensions trials after two or three sessions; this indicates that participants readily grasped the task instructions and could match the Russian phrases with the corresponding pictures after several blocks of trials. Comprehension trials thus provided an opportunity for participants to experience task mastery while gaining additional exposure to the Russian case-marking and gender agreement patterns. This was crucial for establishing a positive attitude toward the language-learning sessions, given the high level of difficulty of the production trials.

In Block 1 (listen-and-repeat task), participants viewed a series of pictures, one at a time, while listening to short dialogues that described each picture. Each dialogue consisted of a question posed by a male speaker, followed by an answer

spoken by a female speaker. For example, when shown a picture of an elephant walking away from an umbrella, the participant would hear the man ask the question *Otkuda ukhodit slon?* [From where is elephant coming?], followed by the woman’s answer *Ot zontika* [from umbrella]. Participants were instructed to repeat the answer as spoken by the woman. After the participant repeated the phrase, the woman’s answer was presented again, and the participant was asked to repeat the phrase a second time. Participants’ responses were audio-recorded. Block 1 comprised 48 randomized trials (16 involving dative case marking, 16 involving genitive case marking, and 16 involving adjective–noun gender agreement with nominative nouns). Table 2 presents the dialogue questions used throughout the experiment, along with examples of answers for masculine and feminine nouns with dative and genitive case markings and showing adjective–noun gender agreement.

In Block 2 (noun comprehension task), participants were tested on their comprehension of the nouns used in the dialogues. On each trial, two pictures were shown side by side on the computer screen at the same time as the dialogue was presented. For trials involving case marking, both pictures depicted events corresponding to the same case. For example, for genitive case, both pictures showed an elephant walking away from an object (e.g., an elephant walking away from an umbrella and an elephant walking away from a train). For adjective–noun gender agreement, both pictures showed objects in the same color (e.g., a blue umbrella and a blue table). Thus, the two pictures differed only in terms of which objects were shown. After listening to the dialogue, the participant was instructed to select the picture corresponding to the object mentioned in the woman’s answer to the man’s question. For example, the participant would hear the man ask *Čto eto?* [What is this?] and the woman answer *Sinij zontik* [blue umbrella], and the participant would then have to choose between pictures of a blue umbrella and a blue table. Participants responded by pressing the corresponding button on a button box (i.e., left button for the left picture or right button for the right picture). After the participant made a choice, the phrase containing the noun (e.g., *Sinij zontik*) was presented again at the same time that the correct picture (e.g., the blue umbrella) was shown. Block 2 comprised 48 randomized trials (16 dative case marking, 16 genitive case marking, and

Table 1 Example of a trial for each learning task

Task and Dialogue	Picture(s) presented	Instructions
<p>Block 1: Listen and Repeat</p> <p>Male voice: <i>Otkuda ukhodit slon?</i> [From where is elephant coming?]</p> <p>Female voice: <i>ot zontika</i> [from umbrella]</p>		<p>Listen to the dialogue and repeat the woman's answer to the question.</p> <p>(Answer is replayed)</p> <p>Repeat the answer a second time.</p>
<p>Block 2: Noun Comprehension</p> <p>Male voice: <i>Otkuda ukhodit slon?</i> [From where is elephant coming?]</p> <p>Female voice: <i>ot zontika</i> [from umbrella]</p>	 	<p>Listen to the dialogue and select the picture that matches the dialogue.</p> <p>Press left button for left picture, right button for right picture.</p> <p>(After response) Answer is replayed as correct picture is shown.</p>
<p>Block 3: Location (or Color) Comprehension</p> <p>Male voice: <i>Otkuda ukhodit slon?</i> [From where is elephant coming?]</p> <p>Female voice: <i>ot zontika</i> [from umbrella]</p>	 	<p>Listen to the dialogue and select the picture that matches the dialogue.</p> <p>Press left button for left picture, right button for right picture.</p> <p>(After response) Answer is replayed as correct picture is shown.</p>
<p>Block 4: Production</p> <p>Female voice: <i>Eto zontik</i> [This is umbrella]</p> <p>Male voice: <i>Otkuda ukhodit slon?</i> [From where is elephant coming?]</p>	 	<p>Listen to the question and answer it in Russian.</p> <p>(After response) Female voice: <i>ot zontika</i> [from umbrella]</p> <p>Repeat the woman's answer to the question.</p>

16 adjective–noun gender agreement). On each trial, the two pictures depicted vocabulary items of the same gender (i.e., both nouns were feminine or both were masculine).

In Block 3 (location/color comprehension task), the participants heard the same dialogues as in Blocks 1 and 2 and were tested on their comprehension of the Russian locative prepositions (for dative and genitive cases) and on their comprehension of the color words (for adjective–noun gender agreement). Two pictures were shown side by side on the computer screen. Both pictures depicted events involving the same object (e.g., an umbrella). For the case-marking

trials, the picture on the left depicted the elephant moving away from the object (genitive case), and the picture on the right, the elephant moving toward the object (dative case). The participant was instructed to listen carefully to the dialogue to find out whether the elephant was moving toward or away from the object. For example, the participant would hear the man ask *Otkuda ukhodit slon?* [From where is elephant coming?] and the woman answer *Ot zontika* [from umbrella], and the participant would select the correct picture according to his or her understanding of the dialogue. For the gender agreement trials, the picture on the left

Table 2 Examples of dialogues and correct responses for the elicited production of dative and genitive case markings and of adjective–noun gender agreement

	Question (Same for Both Genders)	Gender	Answer
Dative case marking	Kuda idjot slon? [To where is elephant going?]	Masculine	k zontiku [to umbrella] k veeru [to fan]
		Feminine	k sumke [to purse] k sveche [to candle]
Genitive case marking	Otkuda ukhodit slon? [From where is elephant coming?]	Masculine	ot zontika [from umbrella] ot veera [from fan]
		Feminine	ot sumki [from purse] ot svechi [from candle]
Gender agreement	Čto eto? [What is this?]	Masculine	sinij zontik [blue umbrella] krasnyj veer [red fan]
		Feminine	sinjaja sumka [blue purse] krasnaja svecha [red candle]

depicted the object in red, and the picture on the right, the same object in blue. The participant would be instructed to listen to the dialogue to determine whether the woman's answer referred to the red or the blue object. For example, the participant would hear the man ask *Čto eto?* [What is this?] and the woman answer *Sinij zontik* [blue umbrella], and the participant would have to choose between pictures of a red umbrella and a blue umbrella. After the participant made a choice by pressing the corresponding button on a button box (i.e., left button for the left picture or right button for the right picture), the phrase (e.g., *Sinij zontik*) would be presented again at the same time that the correct picture (e.g., the blue umbrella) was shown. Block 3 comprised 48 randomized trials (16 dative case marking, 16 genitive case marking, and 16 adjective–noun gender agreement).

In Block 4 (production task), participants' mastery of the case-marking and gender agreement patterns was probed by requiring them to produce what had been presented as the woman's answers in the previous dialogues. First, to remind the participant of the noun to be used in their answer, at the start of each trial, a picture of an object was presented in black and white, along with a phrase spoken by the female voice to introduce the noun in the nominative case (e.g., *Eto zontik* [This is umbrella]).

To elicit production of the dative case marking, a picture of the elephant moving toward the object would be shown, and the participant would hear the man ask the question *Kuda idjot slon?* [To where is elephant going?]. For dative trials, the correct response required the participant to produce the preposition *k* [toward] along with the case-inflected noun. To elicit the genitive case marking, a picture of the elephant moving away from the object would be shown, and the participant would hear the man ask the question *Otkuda ukhodit slon?* [From where is elephant coming?]. Genitive trials required a response involving the preposition *ot* [from] along with the

case-inflected noun. After the participant produced a response, the female voice provided the correct answer (e.g., *Ot zontika*), and the participant was instructed to repeat her answer.

To elicit production of adjective–noun gender agreement, a picture of a red or a blue object was shown, and the participant heard the question *Čto eto?* [What is this?]. Each gender agreement trial required a response containing the Russian color adjective, with gender agreement, followed by the name of the picture in the nominative case. After the participant responded, the female voice provided the correct response (e.g., *Sinij zontik*), and the participant was instructed to repeat her answer.

Participants' responses were audio-recorded. Block 4 comprised 48 randomized trials (16 dative case marking, 16 genitive case marking, and 16 adjective–noun gender agreement).

Testing session Testing was conducted on the same day as the last training session (i.e., after a short break) and lasted about 15 min. The procedure for testing trials was identical to that of Block 4 of training (production), except that test items (in bold in the [Appendix](#)) were added in order to examine participants' ability to generalize patterns of case-marking inflections and gender agreement beyond the trained items. "Reserved" test items were nouns from the trained vocabulary presented in new contexts, whereas "new-noun" test items were novel vocabulary presented in all three contexts. The final test comprised 96 trials (48 trained and 48 test items) presented in a randomized order. We did not include Blocks 1–3 in the testing phase, because of the length of the session and because we did not want to expose participants to the reserved and new-noun test items prior to the production block. Thus, we did not include any comprehension trials in testing for generalization of the case-marking inflections and adjective–noun gender agreement patterns to new items.

At the end of the block of test trials, we administered a vocabulary test. Black-and-white line drawings of each object from the trained set ($N = 24$) were presented one at a time, and the participant was asked to name each picture. No feedback was given. If the participant could not retrieve the name of the picture, he or she was instructed to say “I don’t know” in order to proceed to the next item.

After completing the vocabulary test, we asked each participant to complete an exit questionnaire that asked them (1) whether they had noticed any patterns in the Russian words and phrases, and to describe these patterns; (2) to describe their strategy for producing the correct color words and how they decided which forms of the color adjectives to produce, or whether they had just guessed; (3) to describe their strategy for answering the questions about the elephant moving toward and away from the objects, and how they decided which forms of the nouns to produce, or whether they had just guessed; (4) whether anything about the new vocabulary words had helped them to use the words in phrases; and (5) to describe anything they had learned about the structure of Russian words and phrases and of Russian grammar. The exit questionnaire was designed to measure the participants’ metalinguistic awareness of the Russian gender categories and case-marking paradigm.

Predictor variables

Four cognitive tasks (nonword repetition, reading span, Culture Fair Intelligence Test, and auditory sequence learning) and foreign-language background were used as predictors of Russian vocabulary retention, grammar learning and generalization, and metalinguistic awareness. The foreign-language background questionnaire was completed at the start of the first session to ensure that the participants did not have any prior knowledge of Russian or of any other Slavic or Baltic language. The remaining tasks were presented in a randomized order over Sessions 1–4, with one task administered prior to each training session.

Nonword repetition A set of 90 nonword stimuli from Gupta (2003, Exp. 2) were used in the nonword task. The stimuli comprised 30 two-syllable, 30 four-syllable, and 30 seven-syllable nonwords, recorded by a female native speaker of English. They were divided into five blocks of 18 nonwords, with each block containing six nonwords of each syllable length. The task was run on a Dell PC running E-Prime 1.1 software (Schneider, Eschman, & Zuccolotto, 2002). Each nonword was presented auditorily, with the order of stimuli randomized within each block. Participants were instructed to repeat each nonword as soon as a fixation cross appeared on the computer screen, 100 ms after the offset of the nonword. Participants’ responses were audio-recorded and were scored as correct or incorrect using a

binary criterion of right or wrong. The mean proportion of correct repetition was used as predictor variable.

Reading span To obtain a measure of verbal working memory capacity, we used the materials and procedure described in Daneman and Carpenter (1980). Participants were asked to read aloud a set of unrelated sentences, with each sentence printed individually on an index card. The sentences were read one at a time. After reading a set of sentences, participants were instructed to recall the final word of each sentence in the set. Sentences were presented in sets of increasing size, starting with a size of two sentences per set, up to a size of five sentences per set. The test included 70 sentences, comprising five sets at each set size (two to five sentences). Two scoring methods have previously been used for reading span tasks. In the first method, the researcher determines the span (2–5) by checking whether the last words have been reproduced correctly for all sentences on a given span level. If participants complete only three out of five items correctly, 0.5 span level is added. The second method simply involves counting the number of correctly reproduced words out of the total of 70. The latter measure has been shown to correlate highly with the former (Shah & Miyake, 1996), but it suffers less from a restriction of range, and was therefore chosen for the present analyses.

Culture Fair Intelligence Test To obtain an estimate of non-verbal intelligence, we administered the Cattell Culture Fair Intelligence Test, Scale 3, Form A (Cattell & Cattell, 1973). Participants were given a booklet with four sets of abstract geometrical multiple-choice pattern completion problems and an answer sheet. Each test set started with several example problems and then progressed with problems of increasing difficulty. Participants were instructed to solve as many problems as they could in the allotted time (ranging from 2.5 to 4 min per problem set). Two of the problem sets (“series” and “matrices”) involved selecting an abstract geometric stimulus (from six alternatives) to complete a series or pattern (matrix). One problem set (“classification”) required the participant to identify which two out of five stimuli were alike in some way (i.e., different from the other three). The last problem set (“conditions/topology”) required the participant to select a stimulus (out of five alternatives) that matched a template with respect to the placement of a dot among geometric forms. This test was analyzed using the provided scoring template to determine the total number of correct items for each participant.

Auditory sequence learning To measure auditory sequence-learning ability, we used the “adjacent dependencies” statistical learning task of Misyak and Christiansen (2007), run on a Dell PC running E-Prime 1.1 software (Schneider et al.,

2002), which presented patterns based on the following production rules:

$S \rightarrow NP VP$

$NP \rightarrow d N$

$NP \rightarrow D A N$

$VP \rightarrow V(NP)$

Adjacent dependencies occurred both within and between phrases generated by this grammar. Regarding phrase-internal dependencies, there were two types of determiners—one of which (d) always occurred prior to a noun (N), and the other (D) always directly preceded an adjective (A) that, in turn, always occurred before a noun (D A N). Between-phrase dependencies resulted from every verb phrase (VP) being consistently preceded by a noun phrase (NP) and optionally followed by another noun phrase. The language was instantiated through ten distinct nonwords (*hep, tam, biv, dupp, jux, lum, meep, sig, zoet, and rauk*) distributed over these categories, with three Ns, three Vs, three As, one d, and one D. The assignment of nonwords to categories was randomized for each participant—that is, which nonwords were d, D, N, A, and V. The task was to listen to the sequences for approximately 30 min (60 sequences, each presented three times). In the test, participants performed 40 two-alternative forced choice trials in which they heard two sequences and had to decide which sequence followed the same rules as the input. The correct sequence in half of the test trials was a familiar sequence from the training set; in the other half, it was a novel sequence generated according to the same rules. In the foils, one or more of the nonwords were reordered to violate the rules. The overall proportion of correct responses (calculated over familiar and novel sequences) was used as the predictor variable.

Foreign-language background Participants completed a foreign-language background questionnaire on which they listed the foreign languages that they had studied and/or used at home. Participants rated their proficiency in each language in the domains of reading, writing, listening, and speaking using a 6-point scale, with 1 indicating *very poor* and 6 *excellent*. Participants listed a total of 28 different languages: Spanish was the most popular language (studied in high school and/or college by 71.4 % of our sample), followed by French (studied by 18.2 % of our sample) and Italian (studied by 10.4 % of our sample).² We calculated

² These languages are taught at the College of Staten Island, CUNY, and at many public and private (Catholic) high schools in New York City. Many undergraduate programs require students to take a foreign language, which precluded access to a sample of students without prior exposure to these languages.

two variables based on the questionnaire responses: (1) the total number of languages studied and (2) self-rated proficiency in languages with grammatical gender systems, averaged over the four domains (reading, writing, listening, speaking). We did not calculate self-rated proficiency in languages with case marking, due to the very low frequencies of exposure to any such languages (i.e., fewer than 4 % of participants had any prior exposure to German, the most popular language with case marking). Due to the diversity of the studied languages, we were unable to utilize a standardized assessment of the participants' competence in each of their languages; self-ratings, nevertheless, have been shown to be predictive of individual differences in adult language learning in prior research (Kempe & Brooks, 2008; Kempe et al., 2010).

Coding of language-learning tasks

Participants' responses in the testing session were transcribed and coded by a native speaker of Russian; these comprised 96 production trials and 24 vocabulary recall trials. Data from the six training sessions were not transcribed nor included in any of the analyses, due to the enormous amount of time that would be required to transcribe and code 22,176 responses (48 production trials \times 6 training sessions \times 77 participants); see Kempe and Brooks (2008) for detailed reporting of the participants' learning trajectories in two similarly structured experiments. Case-marking responses were coded as correct if the participant provided the correct preposition and the correct suffix of the target noun. Gender agreement responses were coded as correct if the participant provided the correct adjective suffix indicating the target gender. For a suffix to be coded as feminine, participants had to produce an extra syllable, relative to the masculine adjective, as well as the /a/ ending. Thus, forms like *krasna* and *sinya* were not coded as feminine.

Twenty-five percent of the data were retranscribed and recoded by the same coder without access to the previous transcription, in order to ascertain coding reliability. We computed kappa coefficients as a measure of the agreement between the two rounds of coding. For the case-marking responses, kappa was .82, which indicates very good agreement. For the gender agreement responses, kappa was .86, which also indicates very good agreement.

Vocabulary recall was scored using the following criteria: If the initial phoneme, as well as the basic shape of the word stem, resembled the original Russian word, the item was classified as correctly recalled, irrespective of the final vowel. That is, any errors in inflecting the noun were not counted as vocabulary errors.

Responses to the exit questionnaire were coded for meta-linguistic awareness of the Russian gender categories and case-marking paradigm; scores ranged from 0 (*no awareness*) to 2 (*awareness of both gender categories and case*)

marking). All of the exit questionnaires were coded twice with very high agreement, kappa = .89. We used the combined score (gender and case-marking awareness) as the measure of metalinguistic awareness because accurate production of case marking requires learners to produce inflections that vary in accordance with noun gender (see Table 2 for examples of the required case-marking responses).

Participants received a score of 0 if they said “I don’t know” or “I guessed” in response to the probe questions and/or described mnemonics for remembering vocabulary while failing to provide any additional information. They received a score of 1 for gender awareness if they mentioned that Russian had gender categories and/or mentioned that nouns ending in *-a* used adjectives ending in *-a*. Participants received a score of 1 for case-marking awareness if they mentioned that the forms of the nouns changed depending on whether the elephant was going toward or away from the object, and/or if they listed two or more different vowels that occurred at the ends of the nouns. None of the participants used the term “case marking” to describe the inflections; instead they used a variety of other terms, such as “endings,” “word forms,” “suffixes,” “inflections,” “noun tenses,” and “noun conjugations.” Participants received a score of 2 if they stated conditional rules for case-marking inflections (e.g., stating that nouns ending in consonants took different suffixes than nouns ending in *-a*) or listed three or four inflections and mentioned that the various inflections were used with different sets of words. Participants also received a score of 2 by receiving credit for both gender awareness and case-marking awareness, according to the criteria described above.

Results

Table 3 shows the means, standard deviations, and ranges for all predictor and outcome variables. The participants exhibited a wide range of learning outcomes, which indicates considerable individual differences in the extent to which they learned the basic features of the Russian gender and case-marking system. For our measure of metalinguistic awareness, many more participants showed awareness of case marking than of gender agreement, despite their greater exposure to other languages with grammatical gender (especially Spanish). The distribution of scores was 23.4 % (score of 0, *no awareness*), 45.5 % (score of 1, *awareness of case marking*), 5.2 % (score of 1, *awareness of gender*), and 26.0 % (score of 2, *awareness of case marking and gender*).

Table 4 shows Pearson correlations computed between all of the predictor variables. While a number of predictor variables showed a trend toward a positive link, none of the correlations remained significant after Bonferroni correction ($\alpha = .003$). Table 5 shows the correlations between the outcome measures using Bonferroni correction ($\alpha = .003$). Note that learning of

Table 3 Proportions of correct responses for production of adjective–noun gender agreement, case-marking inflections, and vocabulary recall, along with mean scores for metalinguistic awareness of gender and case marking (metaL), reading span, and the Culture Fair Intelligence Test (CFI); proportions of correct nonword repetitions (NWRRep) and correct responses in the auditory sequence-learning task (AudSL); mean self-ratings of proficiency in languages with grammatical gender (GenderL2); and total number of languages studied (#L2)

		Mean (SD)	Range
Gender agreement	Old	.74 (.14)	.38–1.00
	New	.60 (.13)	.31–.94
Case marking	Old	.77 (.15)	.22–1.00
	New	.42 (.18)	.03–.91
Vocabulary recall		.74 (.16)	.29–1.00
MetaL		1.03 (0.71)	0–2
Reading span		47.52 (9.05)	26–68
CFI		24.06 (4.53)	13–35
NWRRep		.73 (.10)	.48–.96
AudSL		.57 (.12)	.30–.90
GenderL2		2.48 (1.59)	0–6
#L2		2.62 (0.84)	1–7

case marking was positively correlated with vocabulary recall and metalinguistic awareness; learning of gender agreement was not correlated with the other outcome measures.

To determine which of the predictors had an independent effect on the outcome measures, we performed a simultaneous regression analysis, which determined the contribution of each predictor over and above all of the other predictors. The results are given in Table 6. Collinearity diagnostics indicated that all VIFs were below 1.4. This analysis shows that memory measures were predictive of vocabulary retention and that learning of gender agreement was predicted by proficiency with languages that have grammatical gender. Learning of the case-marked nouns presented during training was predicted by auditory sequence learning and by number of studied languages, whereas generalization of case marking was predicted

Table 4 Pearson correlation coefficients for the predictor variables reading span (ReadSpan), Culture Fair Intelligence Test (CFI), nonword repetition (NWRRep), auditory sequence learning (AudSL), self-rated proficiency in languages with grammatical gender (GenderL2), and total number of languages studied (#L2)

	2	3	4	5	6
1. ReadSpan	.22	.30	.16	.24	.24
2. CFI	1	.29	.29	-.03	-.07
3. NWRRep		1	.25	.23	.10
4. AudSL			1	.18	-.08
5. GenderL2				1	.25
6. #L2					1

None of the correlations remained significant after Bonferroni correction.

Table 5 Pearson correlation coefficients for the outcome variables

	2	3	4	5	6
1. Gender (old)	.43*	.12	.28	.15	.12
2. Gender (new)	1	.00	.20	.07	.01
3. Case (old)		1	.58*	.48*	.53*
4. Case (new)			1	.43*	.54*
5. Vocabulary				1	.49*
6. Metalinguistic awareness					1

Asterisks mark correlations that remained significant after Bonferroni correction ($\alpha = .003$).

by nonverbal intelligence. Finally, metalinguistic awareness was predicted by auditory sequence learning and by nonverbal intelligence.

We added metalinguistic awareness as a further predictor into the last set of regression analyses; the results are given in Table 7. This measure was predictive of vocabulary learning and learning of case marking. Note that the effects of nonverbal intelligence and auditory sequence learning were not significant anymore, once metalinguistic awareness was entered into the equation. This points toward metalinguistic awareness mediating the effects of auditory sequence learning and nonverbal intelligence on learning of case marking and on vocabulary retention. To test mediation explicitly, we employed bootstrapping to estimate the 95 % confidence intervals of the indirect effect, using the procedure suggested by Hayes and Preacher (2012) for multiple predictor variables, along with the associated SPSS macro provided by those authors (MEDIATE, downloaded from www.afhayes.com/spss-sas-and-mplus-macros-and-code.html). Mediation is assumed to

Table 6 Standardized regression coefficients obtained from multiple regressions with gender agreement and case-marking performance in old and new items, vocabulary retention, and metalinguistic awareness of gender and case as criterion variables, as well as reading span (ReadSpan), Culture Fair nonverbal intelligence (CFI), nonword repetition (NWRRep), auditory sequence learning (AudSL), self-rating in languages with grammatical gender (GenderL2), and total number of learned languages (Total#L) as predictor variables

	Gender (Old)	Gender (New)	Case (Old)	Case (New)	Vocabulary	Metalinguistic Awareness
ReadSpan	.13	-.03	.01	.07	.25*	.13
CFI	.02	.19	.13	.27*	.05	.28*
NWRRep	.02	.06	.19	.06	.26*	.10
AudSL	.09	-.04	.24*	.16	.09	.27*
GenderL2	.27*	.27*	-.05	.08	-.08	-.16
Total#L	-.02	.16	.30*	.19	.19	.12
R ² total	.13	.16	.24	.22	.26	.29
Model F	1.8 ⁺	2.2 ⁺	3.6**	3.3**	4.0**	4.7***

⁺ $p < .1$. * $p < .05$. ** $p < .01$. *** $p < .001$

Table 7 Standardized regression coefficients obtained from multiple regressions with gender agreement and case-marking performance in old and new items and vocabulary retention as criterion variables, as well as reading span (ReadSpan), Culture Fair nonverbal intelligence (CFI), nonword repetition (NWRRep), auditory sequence learning (AudSL), self-rating in gender languages (GenderL2), total number of learned languages (Total#L2), and metalinguistic awareness of gender and case (MetaL) as predictor variables

	Gender (Old)	Gender (New)	Case (Old)	Case (New)	Vocabulary
ReadSpan	.12	-.01	-.05	.01	.20 ⁺
CFI	.00	.22	.01	.14	-.06
NWRRep	.01	.07	.15	.02	.22*
AudSL	.07	-.02	.12	.05	-.04
GenderL2	.28*	.26*	.02	.15	.07
Total#L2	-.03	.17	.24*	.14	.12
MetaL	.08	-.10	.43***	.45***	.41***
R ² total	.14	.16	.37	.36	.37
Model F	1.6	1.9 ⁺	5.8***	5.6***	5.8***

⁺ $p < .1$. * $p < .05$. ** $p < .01$. *** $p < .001$

be present when the confidence interval for the indirect effect does not include 0. The obtained confidence intervals for the omnibus indirect effect of all predictors simultaneously, based on 5,000 bootstrap samples, were 0.007–0.057 for case marking (old), 0.010–0.062 for case marking (new), and 0.141–1.470 for vocabulary retention. Inspection of the independent indirect effects of the individual predictor variables showed that, specifically, the effects of nonverbal intelligence and auditory sequence learning on case-marking performance and vocabulary retention were mediated by metalinguistic awareness. Note also that the effect of experience with multiple prior languages (measured as the total number of languages that a participant had been exposed to) on learning case marking in the training items remained significant when metalinguistic awareness was included in the regression model. Moreover, the mediation analysis failed to show an independent indirect effect of prior language experience mediated through metalinguistic awareness (the 95 % confidence intervals encompassed 0). These results strongly suggest that, in adult language learners, nonverbal intelligence and auditory sequence learning ability contribute to explicit awareness of underlying rules, while the transfer of knowledge from one language to another, if present, appears to be implicit and not mediated by metalinguistic awareness.

For vocabulary learning, the effect of nonword repetition was retained in the regression model and was not mediated by metalinguistic awareness, which suggests that phonological short-term memory makes an independent and direct contribution to the ability to retain novel words. For accuracy of gender agreement, including metalinguistic awareness into the

regression model did not change the facilitating effects of knowledge of languages with grammatical gender, and estimating the size of the indirect effect through bootstrapping did not provide evidence for mediation by metalinguistic awareness (all 95 % confidence intervals of the indirect effect encompassed 0). This supports the idea that transfer from previously encountered L2s is predominantly implicit.

Discussion

The adults in our study differed markedly in their success at learning a fairly simple 2 (gender) \times 3 (case) inflectional paradigm of Russian. Our main finding was that the strongest predictor of successful learning of case marking was the extent to which learners became aware of the patterns in the input. Learners who discovered the regularities underlying both gender marking and case marking tended to perform better with case marking; these adults were able to describe the two sets of noun inflections that were used with different sets of words and to generalize the inflections to new nominative nouns; that is, they were able to apply conditional rules that varied the case-marking inflections in accordance with the gender of the noun. In contrast, learning of gender agreement was best predicted by the learners' familiarity with other languages with similar gender systems, like Spanish or Italian, where feminine nouns and adjectives also have *-a* endings. These results allowed us to derive two conclusions: First, learners benefit from becoming aware of the regular patterns that underlie novel systems of inflectional morphology, and second, learners transfer their prior knowledge of inflectional morphology from one foreign language to another when encountering morphological systems that are similar to ones they have learned before. The latter finding confirms results from previous studies on gender learning (Kempe et al., 2010; Williams, 2005; William & Lovatt, 2003) that adult learners readily transfer knowledge of similar structures from familiar languages to new ones.

One question that arises is whether learners transfer abstract knowledge of grammatical gender or "reuse" familiar patterns of marking and agreement. The literature on transfer from first to second languages suggests that specific patterns of gender marking and gender agreement are what are transferred (Foucart & Frenck-Mestre, 2011), as opposed to knowledge of abstract gender categories. In the present study, it is likely that learners also transferred specific patterns of gender marking and gender agreement from an already studied foreign language to the novel language. Specifically, over 80 % of our participants had studied Spanish or Italian; both are languages with the majority of feminine nouns ending in *-a*, as in Russian, and with feminine adjectives displaying assonance with the *-a* endings of feminine nouns. Importantly, the effects

of transfer on gender agreement were not mediated by metalinguistic awareness, which suggests that transfer was implicit, with prior knowledge of feminine *-a* endings on nouns and adjectives (from Spanish or Italian) priming their Russian counterparts during acquisition. Thus, our findings extend research on structural priming in speech production to adult L2 learning: Whereas prior research with children and adults had indicated that priming morphosyntactic structures facilitates subsequent use of the same structures in speech production (e.g., Bock et al., 2007; Chang et al., 2000; Savage et al., 2006), our findings suggest that activation of relevant morphosyntactic structures also can occur from one language to another. This interplay between one's knowledge of different languages supports theoretical frameworks such as the competition model (MacWhinney, 2008), whereby any structure from a known language that can find a match in the newly encountered language will transfer.

Both the Culture Fair Intelligence Test and the auditory sequence-learning task appear to tap into participants' ability to detect patterns of regularity: In the case of the Culture Fair Test, the patterns of regularity pertain to visual-spatial configurations of geometrical figures, and in the case of auditory sequence learning, to the temporal ordering of the elements. Although the correlation between those two predictors fell short of significance ($r = .288$), the trend suggests that the ability to detect patterns of regularity distributed in space and in time might be related, and this idea should be pursued in future studies. It should be noted here that a correlation of similar magnitude ($r = .198$) between the same two predictors was reported in a study of foreign-language speech perception (Marronaro, Kempe, & Brooks, 2010). Reber, Walkenfeld, and Hernstadt (1991), likewise, reported a similar trend ($r = .25$) between a standard artificial-grammar-learning task (with visual stimuli) and a different measure of IQ (WAIS-R), although a significant negative correlation ($r = -.34$) was reported by Robinson (2005) in a Japanese replication of Reber et al.'s study.³ Taken together, these studies suggest that further investigations of the relationship between sequence learning and general intelligence are warranted; such studies will require relatively large sample sizes to have sufficient power to detect what may be a small effect.

Our finding that auditory sequence learning was predictive of learners' awareness of the underlying grammatical regularities was unexpected, given the widely held view that

³ Robinson (2005) also reported a negative correlation between performance on the Block Design subtest of the WAIS intelligence test (i.e., a measure of nonverbal intelligence) and a written grammaticality judgment task used to assess implicit learning of Samoan grammar. These apparent discrepancies with our findings cannot easily be resolved, given the differences in the stimulus modalities used to assess sequence-learning ability (auditory vs. visual), the measures of nonverbal intelligence (Culture Fair vs. Japanese WAIS-R), and the language-learning outcomes (spoken phrase production vs. grammaticality judgments of written sentences).

sequence learning occurs without conscious awareness and is independent of general intelligence (e.g., Reber et al., 1991). This position, however, may need to be reconsidered in light of findings that sequence learning is adversely affected by cognitive load and by other manipulations that draw attention away from monitoring the input stream (Fernandes, Kolinsky, & Ventura, 2010; Jiménez & Méndez, 1999; Remillard, 2009; Toro, Sinnett, & Soto-Faraco, 2005). Our results suggest that pattern detection is a mechanism by which sequential information is brought to mind, and in the case of foreign-language speech production (as opposed to grammaticality judgments), it is necessary for one's knowledge to be explicit in order to generate grammatically accurate output.

Interestingly, nonverbal intelligence, as measured by the Culture Fair Test, had a larger effect on learning Russian morphology than on learning the underlying grammar of the auditory sequence-learning task, as mentioned above in reference to the nonsignificant correlations between measures of nonverbal intelligence and sequence learning. Our implicit language-learning paradigm differs from the auditory sequence-learning task in a number of ways: First, grammatical patterns were linked to meanings, and participants had to learn semantics and grammar simultaneously. Second, as mentioned above, our participants had to overtly reproduce the learned Russian utterances rather than perform grammaticality judgments. Finally, in the auditory sequence-learning task, the underlying rules were likely too complex to be accessible to explicit representation, whereas the Russian case-marking inflections were fairly transparent and relatively easy to describe. When presented with such relatively simple input, learners with better control of attention were more able to extract and complete the patterns, and were more likely to notice the underlying regularities.

Our findings showed that vocabulary learning was affected by two factors: Learners who were better at repeating non-words and, surprisingly, learners who became aware of the underlying morphological patterns were the most successful at retaining the novel words. The effect of phonological short-term memory capacity, measured via the nonword repetition test, provides a direct confirmation that the phonological loop serves as a language acquisition device supporting the learning of new words (Baddeley et al., 1998). With respect to the effect of metalinguistic awareness on vocabulary learning, we have hypothesized elsewhere that, in miniature language-learning paradigms, factors that support morphology learning may free up resources for vocabulary learning (Kempe & Brooks, 2008), which can add to the effects of phonological short-term memory capacity. The present findings support the view that faster learning of morphology benefits vocabulary acquisition: As described in the Method section, our training set of vocabulary comprised 50 % nouns with diminutive suffixes

(i.e., *-ik* for masculine and *-ka* for feminine in the nominative case) and 50 % simplex nouns. In Brooks et al. (2011), we reported analyses from the present data set showing that learners were more successful in generalizing the morphosyntactic patterns to the diminutive nouns and were more accurate in recalling diminutive than in recalling simplex nouns in the vocabulary test. Diminutives apparently helped learners to decompose the inflected nouns into stems and suffixes. This enhanced their discovery of the inflectional patterns, as well as strengthening their representations of the word stems, which benefited vocabulary retention.

It is important to place our study in the context of the many studies that have explored differences between implicit and explicit L2 learning (cf. DeKeyser, 2003, for a review). In most of these studies, participants were exposed to grammatical patterns in a foreign language under conditions that promoted awareness, such as explicit instruction or perceptual enhancement of the regularities in the input, with learning outcomes being compared to those from conditions that did not promote awareness (e.g., de Graaff, 1997; DeKeyser, 1995; Doughty, 1991; Ellis, 1993; Leow, 1998; Robinson, 1996, 1997; Rosa & O'Neill, 1999). These studies showed a general benefit from interventions that made learners explicitly aware of the structural regularities in the input, especially when these regularities could be described by relatively simple categorical rules. This benefit was reduced, however, when learners were left to discover the regularities on their own (Robinson, 1996). Our findings suggest that this reduction in learning success may be attributable to the substantial individual differences in how likely learners are to discover underlying rules. The finding that nonverbal intelligence and sequence-learning ability strongly modulate success in explicit rule discovery suggests that differences in these individual abilities may, in fact, render the outcomes of learning through rule discovery highly variable. This conclusion is supported by evidence from educational research, which suggests that explicit teaching and scaffolding generally yield better learning outcomes than does discovery-based instruction (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Our findings, placed in the context of what we know about the role of explicit awareness, support pedagogical approaches that aid learners in discovering the underlying regularities in L2 input.

Author note This research was supported by grants from *Language Learning—A Journal of Research in Language Studies* and from PSC-CUNY to P.J.B. A preliminary report of our findings was presented at the 50th Annual Meeting of the Psychonomic Society, Boston, Massachusetts. Previous analyses of the data set have been published elsewhere (Brooks et al., 2011) but did not include the analyses of individual differences reported here. We thank Annemarie Donachie, Manuel Garcia, Jessica Lee, Rosemarie Marronaro, Irosha Pathirage, and Urooj Syed for their assistance with data collection and tabulation. We also thank Jennifer Misyak, Morten Christiansen, and Prahlad Gupta for sharing their stimulus materials.

Appendix

Table 8 Experimental stimuli

Noun [Meaning] (Gender, Derivation)	Dative Case Marking	Genitive Case Marking	Gender Agreement
chajnik [kettle] (M, D)	k chajniku	ot chajnika	krasnyj chajnik
domik [house] (M, D)	k domiku	ot domika	krasnyj domik
mjachik [ball] (M, D)	k mjachiku	ot mjachika	krasnyj mjachik
persik [peach] (M, D)	k persiku	ot persika	sinij persik
stolik [table] (M, D)	k stoliku	ot stolika	sinij stolik
zontik [umbrella] (M, D)	k zontiku	ot zontika	sinij zontik
listik [leaf] (M, D)	k listiku	ot listika	krasnyj listik
sharik [balloon] (M, D)	k shariku	ot sharika	sinij sharik
cherep [skull] (M, S)	k cherepu	ot cherepa	krasnyj cherep
orekh [nut] (M, S)	k orekhu	ot orekha	sinij orekh
poezd [train] (M, S)	k poezdu	ot poezda	krasnyj poezd
veer [fan] (M, S)	k veeru	ot veera	krasnyj veer
verbljud [camel] (M, S)	k verbljudu	ot verbljuda	sinij verbljud
zabor [fence] (M, S)	k zaboru	ot zabora	sinij zabor
stakan [glass] (M, S)	k stakanu	ot stakana	krasnyj stakan
topor [axe] (M, S)	k toporu	ot topora	sinij topor
bochka [barrel] (F, D)	k bochke	ot bochki	krasnaja bochka
chashka [cup] (F, D)	k chashke	ot chashki	krasnaja chashka
kepka [cap] (F, D)	k kepke	ot kepki	krasnaja kepka
pushka [cannon] (F, D)	k pushke	ot pushki	sinij pushka
sumka [purse] (F, D)	k sumke	ot sumki	sinij sumka
vilka [fork] (F, D)	k vilke	ot vilki	sinij vilka
banka [jar] (F, D)	k banke	ot banki	krasnaja banka
skripka [violin] (F, D)	k skripke	ot skripki	sinij skripka
kniga [book] (F, S)	k knige	ot knigi	krasnaja kniga
mukha [fly] (F, S)	k mukhe	ot mukhi	krasnaja mukha
noga [leg] (F, S)	k noge	ot nogi	sinij noga
skamja [bench] (F, S)	k skamje	ot skamji	sinij skamja
svecha [candle] (F, S)	k sveche	ot svechi	krasnaja svecha
vishnja [cherry] (F, S)	k vishne	ot vishni	sinij vishnja
shljapa [hat] (F, S)	k shljape	ot shljapy	krasnaja shljapa
tykva [pumpkin] (F, S)	k tykve	ot tykvy	sinij tykva

Items in bold were reserved for testing. M, masculine; F, feminine; D, diminutive; S, simplex

References

- Alfieri, L., Brooks, P. J., Aldrich, N., & Tenenbaum, H. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1–18.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173. doi:10.1037/0033-295X.105.1.158
- Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*, 437–458.
- Brooks, P. J., Kempe, V., & Donachie, A. (2011). Second language learning benefits from similarity in word endings. *Language Learning, 61*, 1142–1172.
- Brooks, P. J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics, 27*, 185–209.
- Cattell, R. B., & Cattell, H. E. P. (1973). *Measuring intelligence with the Culture Fair Tests*. Champaign, IL: Institute for Personality and Ability Testing.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research, 29*, 217–229.
- Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271. doi:10.3758/BF03204507
- Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114*, 356–371.
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24–39. doi:10.1037/0278-7393.31.1.24
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905–912. doi:10.1111/j.1467-9280.2006.01801.x
- Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology, 21*, 561–580.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163–183.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466. doi:10.1016/S0022-5371(80)90312-6
- de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*, 249–276.
- DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 379–410.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Malden, MA: Blackwell.
- Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*, 431–469.
- Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology, 5*, 289–318.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305–352.
- Ellis, N., & Schmidt, R. (1998). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19*, 145–171.
- Ellis, N., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49A*, 234–250.
- Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without

- specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 50, 177–195.
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52, 321–335.
- Fernandes, T., Kolinsky, R., & Ventura, P. (2010). The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues. *Attention, Perception, & Psychophysics*, 72, 1522–1532. doi:10.3758/APP.72.6.1522
- Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14, 379–399.
- Frost, R., Narkiss, A., Afek, L., & Siegelman, N. (2011, November). *What predicts faster literacy acquisition in a second language?* Paper presented at the 52nd Annual Meeting of the Psychonomic Society, St. Louis, MO.
- Garlick, D., & Sejnowski, T. J. (2006). There is more to fluid intelligence than working memory capacity and executive function. *The Behavioral and Brain Sciences*, 29, 134–135. doi:10.1017/S0140525X0630903X
- Gathercole, S., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28, 200–213.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–436.
- Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal*, 84, 390–405.
- Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology*, 56A, 1213–1236. doi:10.1080/02724980343000071
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25–38.
- Hayes, A. F., & Preacher, K. J. (2012). *Indirect and direct effects of a multicategorical causal agent in statistical mediation analysis*. Manuscript submitted for publication.
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, 9, 220–225.
- Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 236–259. doi:10.1037/0278-7393.25.1.236
- Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Macintosh, N. (2010). Implicit learning as an ability. *Cognition*, 116, 321–340.
- Kempe, V., & Brooks, P. J. (2001). The role of diminutives in the acquisition of Russian gender: Can elements of child-directed speech aid in learning morphology? *Language Learning*, 51, 221–256.
- Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning*, 58, 703–746.
- Kempe, V., Brooks, P. J., & Kharkhurin, A. (2010). Cognitive predictors of generalization of Russian grammatical gender categories. *Language Learning*, 60, 127–153.
- Lemhöfer, K., Schriefers, H., & Hanique, I. (2010). Native language effects in learning second-language grammatical gender: A training study. *Acta Psychologica*, 135, 150–158.
- Lemhöfer, K., Schriefers, H., & Indefrey, P. (2011, September). *Cognitive mechanisms of L2 errors and “fossilization”*: An ERP study. Poster presented at the 17th Annual Conference on Architectures and Mechanisms for Language Processing, Paris, France.
- Leow, R. P. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa’s (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, 19, 133–159.
- Lum, J. A. G., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders*, 45, 96–107.
- MacWhinney, B. (2008). A unified model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). Mahwah, NJ: Erlbaum.
- Marronaro, R., Kempe, V., & Brooks, P. J. (2010, March). *Exploring gender differences in tone perception in adult English speakers*. Poster presented at the Annual Meeting of the Eastern Psychological Association, Brooklyn, NY.
- Misyak, J. B., & Christiansen, M. H. (2007). Extending statistical learning farther and further: Long-distance dependencies, and individual differences in statistical learning and language. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 1307–1312). Austin, TX: Cognitive Science Society.
- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, 2, 138–153. doi:10.1111/j.1756-8765.2009.01072.x
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.
- Mueller, J. L. (2006). L2 in a nutshell: The investigation of second language processing in the miniature language model. *Language Learning*, 56, 235–270.
- Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, 30, 331–347.
- Plante, E., Gómez, R., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, 35, 453–462.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219–235. doi:10.1037/0096-3445.118.3.219
- Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 888–896.
- Remillard, G. (2009). Pure perceptual-based sequence learning: A role for visuospatial attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 528–541.
- Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning*, 45, 283–331.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 19, 223–247.
- Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, 27, 235–268.
- Rosa, E., & O’Neill, M. D. (1999). Explicitness, intake, and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511–556.
- Salamoura, A., & Williams, J. N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual mental lexicon. *Applied Psycholinguistics*, 28, 627–660.

- Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2*, 27–49.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158. doi:10.1093/applin/11.2.129
- Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206–226. doi:10.1017/S0267190500002476
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45A*, 21–50.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16*, 155–172.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4–27. doi:10.1037/0096-3445.125.1.4
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215. doi:10.1037/0278-7393.6.2.174
- Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293–321.
- Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning and adolescents with and without specific language impairment. *Language Learning and Development, 3*, 269–293.
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*, B25–B34.
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to the working memory–fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review, 16*, 931–937. doi:10.3758/PBR.16.5.931
- Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269–304.
- Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning, 53*, 67–121.