Distinctions in the acquisition of vocabulary and grammar: An individual differences approach

Neil Walker [1, 2], Padraic Monaghan [3, 4], Christine Schoetensack [2], and Patrick Rebuschat [2,5]

[1] School of Language and Global Studies, University of Central Lancashire

[2] Department of Linguistics and English Language, Lancaster University

[3] Department of Psychology, Lancaster University

[4] Faculty of Humanities, University of Amsterdam

[5] LEAD Graduate School and Research Network, University of Tübingen

Acknowledgements:

Our research was supported by the Economic and Social Research Council, UK (grant ES/L008955/1), the LEAD Graduate School & Research Network (grant DFG-GSC1028), a project of the Excellence Initiative of the German federal and state governments, and by an Early Career Research Grant awarded by Lancaster University to Patrick Rebuschat.

Abstract

Learning language requires acquiring the grammatical categories of words in the language, but learning those categories requires understanding the role of words in the syntax. In this study, we examined how this chicken and egg problem is resolved by learners of an artificial language comprising nouns, verbs, adjectives and case markers following syntactic rules. We also measured individual differences in declarative and procedural memory processing, which have been linked to vocabulary and grammar learning, respectively. The results showed that grammar and vocabulary can be acquired simultaneously, but with distinctive patterns of acquisition – the syntactic role of verbs and their referents first, then other lexical categories, and finally the syntactic function of case markers. Interdependencies in learning were found for word order and verbs, which related to verbal declarative memory, and also for nouns, adjectives and case markers, which related to procedural memory.

Keywords: language acquisition; grammar; vocabulary; declarative memory; procedural memory; cross-situational learning.

Distinctions in the acquisition of vocabulary and grammar: An individual differences approach

A Syrian refugee claims asylum in Sweden; the child of a Chinese economic migrant starts her first day of school in Canada; a British tourist passes through Turkish customs at the beginning of a fortnight's holiday. In each of these examples the learner may know almost nothing of the local, second language. The early stages of second language learning under such immersion conditions involve a great deal of ambiguity as learners struggle to make sense of the stream of input they hear by detecting word boundaries, decoding the meanings of words, identifying lexical categories and understanding the relations between categories defined by the syntax. Even at this early stage, there is individual variation in the ease with which learners pick up the new language (see Dörnyei, 2014 for an overview). How learning is achieved and how individual differences may affect the learning process have been critical questions in the cognitive sciences (Frost & Monaghan, 2016; Marcus, 1996; McGregor, Sheng & Smith, 2005; Siegelman, Bogaerts, Christiansen, & Frost, 2017).

Recent research has shown that it is possible for children and adults to learn vocabulary within basic categories of words when they are presented across multiple ambiguous learning situations without any feedback, a process known as cross-situational learning (Yu & Smith, 2007). Smith and Yu (2008) showed that 12 to 14-month old infants could learn the meanings of novel nouns by keeping track of cross-trial statistics. Scott and Fisher (2012) further demonstrated that 2.5-year-old toddlers could utilise distributional cues to learn novel verbs. With adult participants, Monaghan and Mattock (2012) found that function words could aid the cross-situational learning of nouns and verbs compared to an artificial language where function words did not co-occur with grammatical categories in the language. Then in a follow-up study, Monaghan, Mattock, Davies and Smith (2015) found that nouns and verbs could be learned simultaneously without syntactic cues. The learning mechanisms underlying cross-situational learning are still an area of debate, with some theories proposing an associative, accumulation of statistical probabilities (Yu & Smith, 2007) and others hypothesis-testing accounts (Medina, Snedeker, Trueswell & Gleitman, 2011). A recent study by Khoe, Perfors and Hendrickson (2019) modelling the two approaches suggests that with more ambiguity in the learning environment, learning is more associative. These associative statistical learning mechanisms may paradoxically be both domain-general yet also constrained by, and therefore distinctly represented within, different modalities (Frost, Armstrong, Siegelman & Christiansen, 2015). While statistical learning has tended to be examined with word learning tasks (e.g. Smith & Yu, 2008), implicit learning has been examined with artificial grammars (e.g. Gómez & Gerken, 1999). Studies such as Monaghan et al. (2015) mentioned above and that reported in the current study draw together the two research traditions (see also Monaghan, Schoetensack and Rebuschat (2019).

These studies looking at the cross-situational learning of nouns and verbs, however, are still a substantial abstraction from the complexity of natural language acquisition. With every new word category or syntactic phrase added, the number of possible referents for any given word increases, making the tracking of statistical probabilities more complex. In a recent study utilising a novel paradigm, Rebuschat, Monaghan and Schoetensack (submitted) demonstrated that it was possible for adults to learn a more complex artificial language under cross-situational learning conditions. The artificial language consisted of a verb-final syntax and contained nouns, verbs, adjectives and case markers denoting the agent and patient of the sentence. Participants saw two dynamic scenes on a computer screen and heard a sentence in the artificial language. Their task was to decide which scene the sentence referred to, and no feedback was given as to whether the participant was right or wrong. Rebuschat et al. observed that verbs and basic word order were learned first, followed by nouns, then adjectives and finally case markers, which is in line with first language acquisition studies into languages which can and do omit subjects (e.g., Korean, Choi & Gopnik, 1995; Mandarin, Tardif, 1996) and adult second language (L2) first exposure studies which demonstrate the increased salience of sentence-initial and sentence-final positions (Fernald, McRoberts & Herrera, 1992; Shoemaker & Rast, 2013). Studies such as these tentatively suggest that this novel cross-situational learning paradigm may be a useful proxy for the early stages of language learning under immersion settings for L2 adults in future research.

Whereas Rebuschat et al. (submitted) demonstrated the viability of a sentence-to-scene cross-situational learning, two key questions about natural language acquisition remain unanswered by their investigation. First, the training and testing of Rebuschat et al. took place in a single session, and the ability of participants to retain both syntactic and vocabulary knowledge over time is not yet known. If it is possible to learn more complex language through cross-situational learning conditions, is that learning durable? To our knowledge, only one study has investigated the long-term learning effects of cross-situational word learning. Vlach and Sandhofer (2014) found that noun learning under cross-situational conditions was still robust one week later. In related research, several studies have demonstrated the durability of statistical learning (Durrant, Taylor, Cairney and Lewis, 2011; Gómez, Bootzin & Nadel, 2006; Kim, Seitz, Feenstra and Shams, 2009); Vuong, Meyer & Christiansen, 2016). Durrant et al. (2011), using Saffran, Aslin and Newport's (1996) paradigm, showed that statistical learning lasts 24 hours and benefits from sleep consolidation while Kim et al. (2009) found similar robustness in a visual statistical learning task. To this end, the current study tests learners' knowledge of vocabulary and syntax immediately after training, but also after a 24-hour delay. Second, the relations between learning syntax and vocabulary are underexplored. It may be that acquisition of vocabulary and syntax are associated with different processes, as reflected by distinct sets of individual differences in learning and memory.

Relations Between Vocabulary and Syntax

The chicken and egg problem of learning syntax and vocabulary has led to proposals either for independence of learning grammar and vocabulary (e.g., Marcus, 1996; Peña, Bonatti, Nespor & Mehler, 2002), or their inter-relatedness (e.g. Bates & Goodman, 1997; Frost & Monaghan, 2016). In other words, are the referents for vocabulary items (nouns, verbs, adjectives) learned in the same way as grammatical items (word order, case markers) or do they depend on different mechanisms? Many previous studies of artificial language learning have trained participants on vocabulary before testing them on grammatical structure (e.g., Friederici, Steinhauer & Pfeifer, 2002; Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter & Wong, 2014; Williams & Kuribara, 2008), and neuropsychology patient studies (Alario & Cohen, 2004), theoretical models (Bock & Levelt, 1994), and memory models (Ullman, 2004) have tended to treat vocabulary and syntax as distinct. An alternative is that grammar and vocabulary instead depend on a single, domain-general learning mechanism (Bates & MacWhinney, 1987; Frost & Monaghan, 2016; MacDonald, Pearlmutter, & Seidenberg, 1994; Rumelhart, Hinton & McClelland, 1986; Seidenberg, 1997). Whether vocabulary and grammar are related to the same, or different, patterns of cognitive abilities enables a test of whether learning is coherent or fragmented into one or several abilities.

Individual Differences in Cross-Situational Learning

As we suggested above, acquisition of vocabulary and syntax may be differentially sensitive to individual differences in memory. In this paper we will consider a number of types of memory, namely phonological short-term memory (PSTM), working memory capacity, declarative memory and procedural memory. Phonological short-term memory (PSTM), the short-term store for auditory information and articulatory rehearsal device as measured by simple span tasks, has been implicated in vocabulary acquisition (Baddeley, Papagno, & Vallar, 1988; Gupta, 2003; Martin & Ellis, 2012; Papagno, Valentine & Baddeley, 1991), and grammar abstraction (Ellis, 2012; Robinson, 1997; Speidel, 1993; Verhagen & Leseman, 2016). Working memory capacity, defined as the ability to not only hold but also simultaneously process items in short-term memory and measured by complex span tasks, has been linked to the noticing of grammatical regularities in language and on-line language processing (Coughlin & Tremblay, 2013; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Mackey, Adams, Stafford & Winke, 2010; Sagarra & Herschensohn, 2010). Sagarra and Herschensohn (2010), for example, found that for low-level L2 learners working memory capacity predicted success on a grammaticality judgment task for gender agreement. A meta-analysis by Linck, Osthus, Koeth, and Bunting (2014) found that working memory capacity was a better predictor of a range of L2 performance measures than simple, storage-only memory such as PSTM. However, a consensus has yet to be reached about the role of working memory capacity under incidental learning conditions, with some studies reporting a positive relationship between working memory capacity and learning outcomes (Robinson, 2005; Soto & Silvanto, 2014) while others showing no effect of working memory capacity (Grey, Williams & Rebuschat, 2015; Hamrick, 2015; Tagarelli, Borges Mota, & Rebuschat, 2015). In a study on implicit and explicit corrective feedback, Li (2013) also found evidence to suggest that L2 proficiency could mediate the impact of working memory capacity, with lower-proficiency learners less reliant on it than higher-proficiency learners.

Declarative memory is the long-term memory system responsible for storing episodic and semantic knowledge. It is a fast-learning, flexible system that can learn both explicitly and implicitly, although it is thought to be the only system responsible for explicit knowledge (see Ullman, 2016, for an overview). In the field of second language acquisition, it has not only been linked to the acquisition of vocabulary but also, due partly to its relative speed compared to procedural memory, the initial stages of grammar learning (Hamrick 2015; Morgan-Short et al.,2014; Ullman, 2004). In contrast, procedural memory is a non-declarative long-term memory system. It is slower to learn yet more robust than declarative memory and is always implicit (Ullman, 2004). It is thought to be involved in pattern recognition and habit formation in general. In language acquisition in particular, it has been hypothesized to be involved in grammar learning, including non-idiosyncratic aspects of vocabulary learning (Ullman, 2004). Studies have also suggested that the slower-learning procedural memory takes over from declarative memory at later stages of the acquisition process (Hamrick, 2015; Morgan-Short et al., 2014). There is evidence that procedural and declarative memory systems are somewhat redundant in that items can be learned using either and often both systems simultaneously and that learning conditions can affect which system takes a lead role (Ullman & Lovelett, 2016). For example, explicit learning conditions may encourage reliance on the declarative memory system while more implicit conditions may force the use of procedural memory systems (Ullman, 2016; Ullman & Lovelett, 2016).

To our knowledge, only two studies have looked specifically at individual differences in memory in cross-situational learning. Schoetensack (2015) found that neither working memory capacity nor PSTM predicted cross-situational learning of nouns and verbs under incidental learning conditions, although she reported that in the instructed condition the relationship between PSTM and learning approached significance. Vlach and DeBrock (2017) found that a paired-associates memory test predicted success on a cross-situational word-learning task in young children. They further demonstrated that visual, auditory and word-binding declarative memory each play a role.

On the other hand, a number of studies have investigated statistical learning ability, which is the capacity to implicitly keep track of statistical information in the input to acquire linguistic information, as an individual difference in its own right (see Siegelman et al., 2017, for an overview). Statistical learning ability is most commonly measured by means of the auditory speech segmentation task of Saffran et al. (1996) or by means of a visual statistical learning task (Fiser & Aslin, 2001; Kirkham, Slemmer & Johnson 2002). It has been found to predict success in a number of aspects of language learning relevant to the current study. Firstly, several studies have shown that statistical learning predicts vocabulary development (Shafto, Conway & Houston 2012; Singh, Reznick & Xuehua, 2012). Shafto et al. (2012) used a visual sequence learning task with 8-month-old infants and found that it predicted success on vocabulary comprehension tests. Meanwhile, Singh et al. (2012) also tested similar-aged infants and found that an auditory word-segmentation task predicted the size of productive vocabulary at 24 months. In addition, statistical learning is linked to sentence processing. For example, Misyak and Christiansen (2012) showed that adults' performance on auditory statistical learning tasks was the main predictor for L1 sentence comprehension. Finally, statistical learning has also been shown to predict success in the acquisition of syntax. Kidd (2012) found that statistical learning ability, as measured by a visual sequence learning task, predicted the learning of syntax in four and five-year-old children and that this was still the case on a delayed test 24 hours later. In a later study, Kidd and Arciuli (2016) showed that variability in six-to-eight-year-olds' comprehension of aspects of syntax was predicted by their performance on a visual sequence learning task. Interestingly, as in several other studies into statistical learning ability as an individual difference (Hunt & Aslin, 2001; Kaufman et al., 2010; Schvaneveldt & Gómez, 1998), Kidd (2012) used the Serial Reaction Time (SRT) task, which is a common measure of procedural memory and is used as such in the current study. Nevertheless, questions remain to what extent there is an overlap between the statistical learning, procedural memory, and indeed implicit learning constructs (see Kóbor et al., 2018; Monaghan, Rebuschat & Schoetensack, 2019).

The Current Study

The current study focuses on whether learners' ability to acquire an artificial language under cross-situational learning conditions is durable over time, and whether it is affected by individual differences in four memory systems that have been associated with language learning: phonological short-term memory, working memory capacity, declarative memory and procedural memory. In doing so, it may shed light on the nature of the underlying mechanisms of cross-situational learning and also add to the growing body of research that suggests it may be possible to explain language learning under immersion conditions in adults through a general-purpose cross-situational statistical learning mechanism. In addition, investigating whether learning is durable is important from a methodological perspective, with possible implications for study design. Finally, the current study may also help refine theories of how memory systems interact with language learning (e.g. Ullman, 2004, Ullman & Lovelett, 2016). Understanding what makes some adults better learners than others is essential to any theory of language learning.

We therefore report an experiment below in which language learners were exposed to an artificial language under cross-situational learning conditions. We manipulated the cross-situational scenes that they were presented with, in order to test them on their learning of nouns, verbs, adjectives, case markers and word order both during exposure and on a delayed test 24-hours later. Participants were also tested on five cognitive individual difference measures.

We predicted that, similar to Rebuschat et al. (submitted), verbs and word order would be learned first, followed by nouns, adjectives and case markers, although given the short duration of the learning paradigm, it was possible that the latter may not be learned at all (e.g., DeKeyser, 2005). Based on the findings of studies investigating cross-situational learning (Vlach & Sandhofer, 2014) and statistical learning (Durrant et al., 2011, Gómez et al., 2006, Kim et al., 2009 and Vuong et al., 2016), we predicted that learning would be durable after 24 hours in all aspects of the artificial language. Regarding the coherence or distinctiveness of the learning of vocabulary and grammar, we had two competing hypotheses. If there is a clear distinction between grammar and vocabulary, we expected learning of word order and case markers to be related, and learning of nouns, verbs and adjectives also inter-related. This would mean that two underlying components may effectively describe learning of the different tasks in the language. Alternatively, if grammar and vocabulary share the same learning mechanisms, as is postulated in single-system models, we expected to see inter-relations between all aspects of the language. Finally, with regards to individual differences, while a majority of studies into working memory capacity and PSTM and language learning suggest we could make firm predictions for their role in both vocabulary and grammar learning (see e.g. Baddeley, Papagno, & Vallar, 1988; Mackey et al., 2002; Sagarra & Herschensohn, 2010), two factors may diminish their influence. Firstly, it remains a possibility that other cognitive measures (specifically procedural memory and declarative memory) may mediate the effects of PSTM and working memory capacity and better predict grammar learning (Brooks & Kempe, 2013). If this were the case, we would not find such a strong link. Secondly, if working memory does not play such a strong role in incidental learning conditions as it does under explicit conditions, then that too may reduce their influence in this study. Based on Ullman's (2004) declarative/procedural model, we expected declarative memory to predict success during these early stages of acquisition. However, we left open the possibility that procedural memory might be a predictor for the more grammatical aspects, particularly under the more incidental learning conditions of cross-situational learning.

Method

Participants

Sixty-four native speakers of English (47 women, 17 men) were randomly assigned to two exposure conditions (*massed* vs *distributed*, each n = 32). These conditions varied in terms of whether there were three 20-minute pauses between blocks of exposure on an artificial language learning task.[1] Participants were students or graduates of the University of Central Lancashire or Lancaster University, both in the North West of England. The mean age was 26.0 years (*SD* = 7.1). None of the participants had previously studied any verb-final languages. Four participants reported being bilingual in Spanish, Punjabi, Thai and Guajarati, in each case together with English. Thirty-three participants reported at least one other language to an intermediate-level of proficiency or above in Spanish (13), French (11), German (5), Polish (1), Italian (2), Portuguese

[1] These conditions did not exert any effect on performance and therefore will not be considered further.

(1), Chinese (2) and Malay (1). Twenty-two had previous experience of studying a case-marked language. Fifteen participants were studying or had studied a linguistics or language-related degree. Participants in the massed group received 20 GBP and participants in the distributed group received 28 GBP. The difference was due to the extra time involved in the distributed condition. Power analyses of the effect sizes found in Rebuschat et al. (submitted) using G*Power demonstrated that 64 participants were sufficient to achieve .99 power for demonstrating learning of syntax, nouns, and verbs, power of .56 for adjectives, and .05 power for case marker words (which are learned only under certain conditions). Despite the low power for case markers, we decided to include them in the current study for completeness.

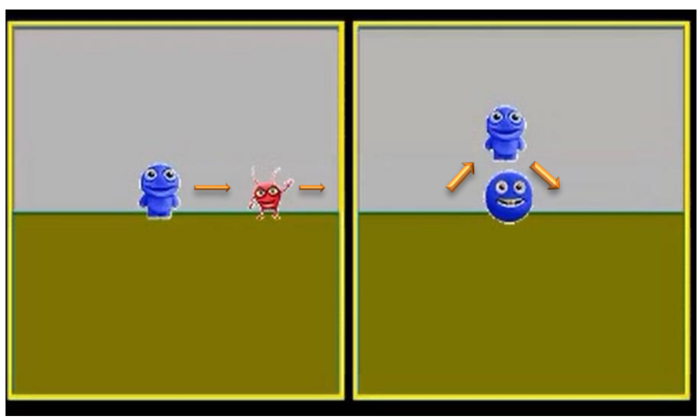


Figure 1. Screenshot of the cross-situational learning exposure task. Participants see two dynamic scenes and hear a sentence (e.g., “Hagal chilad tha garshal sumbad noo thislin”). Their task is to decide which scene the sentence refers to. The arrows indicate the direction of movement of the aliens and are for reference only, not appearing on the screen.

Materials

The artificial language developed by Rebuschat et al. (submitted) was used in this experiment (see also Walker, Schoetensack, Monaghan, & Rebuschat, 2017). The lexicon consisted of 16 pseudowords, taken from Monaghan and Mattock (2012; see Appendix A). Fourteen bisyllabic pseudowords were content words: Eight nouns, four verbs, and two adjectives. Two monosyllabic pseudowords served as function words that reliably indicated whether the preceding noun referred to the subject or the object of the sentence. The words were recorded by a female native speaker of British English who was instructed to produce the words in a monotone.

In terms of syntax, the artificial language was based on Japanese. Sentences could either be SOV or OSV, i.e. verbs had to be placed in final position but the order of subject and object noun phrases (NPs) was free. NPs had to contain a noun as its head and a post-nominal case marker that indicated whether the preceding noun was the agent or the patient of the action. Adjectives were optional and only occurred in half the NPs. When adjectives were present, they occurred pre-nominally. The syntactic patterns used in the experiment can be found in Appendix B.

Eight alien cartoon characters served as referents for the language (Appendix C). The aliens could either appear in red or blue and were depicted performing one of four actions (hiding, jumping, lifting, pushing) in dynamic scenes generated by E-Prime (version 2.0). Figure 1 shows a sample screen shot, containing the target scene and a distractor scene. Each noun referred to one alien, the adjectives referred to the colours of aliens, and the verbs referred to the actions. Six different versions of word-referent mappings were randomly generated to control for preferences in associating certain sounds to objects, motions or colours. Since adjectives were optional, sentences were between five and seven pseudowords in length.

Individual Difference Measures

Phonological short-term memory. PSTM capacity was measured by means of Gathercole and Baddeley's (1996) nonword repetition test (NRT). This task required participants to listen to pseudowords of different lengths and to repeat each word exactly as they heard it. Answers were scored as described in Gathercole (1995), i.e. responses were either considered as correct (1 point) or incorrect (0 points), depending on whether or not all phonemes had been reproduced correctly. Cronbach's alpha is .80.

Working memory capacity. The storage and processing function of working memory was measured by means of Unsworth, Heitz, Schrock and Engle's (2005) Automated Operation Span Task (Aospan). During this computerized test, participants were presented with letters interspersed with maths problems, which they were required to solve while keeping the letters in memory at the same time. All letters were randomly selected for each subject from an array of twelve options (F, H, J, K, L, N, P, Q, R, S T and Y). The math problems always followed the pattern (x multiplied/divided by y) +/- z = ? (e.g. (16 / 2) + 3 = 11) and required the participant to choose whether the answer was true or false. The Aospan score was the sum of the number of letters that a participant was able to recall across blocks of increasing difficulty. Following Unsworth et. al., an accuracy criterion of a minimum of 85% for the maths problems was set, which resulted in 10 participants being excluded. Unsworth et al. report a Cronbach's alpha of .78.

Declarative memory. Visual declarative memory capacity was assessed by means of the Continuous Visual Memory Test (CVMT) (Trahan & Larrabee, 1988). During the task, participants saw black-and-white drawings of complex figures, presented in succession, and were required to indicate for each of them if they had seen it before (if it was an "old" picture) or not (if it was "new"). D-prime scores were generated from the responses. The CVMT has a split-half reliability of .80. Verbal declarative memory capacity was assessed by means of the MLAT-V, a paired-associates test from the Modern Language Aptitude Test (Carroll & Sapon, 1959) that requires participants to rapidly learn 24 pseudo-Kurdish words and their English translations. Scores were calculated with one point for every correctly chosen item, for a maximum score of 24. Carroll and Sapon reported a split-half reliability of between .92 and .97.

Procedural memory. Procedural memory was measured through the Serial Reaction Time (SRT) task (Nissen & Bullemer, 1987, adapted by Lum, Gelgic and Conti-Ramsden, 2010). In this task, participants saw a smiley face appear in one of four positions (top, right, bottom, left) on a computer screen. Their task was to press the corresponding button (top, right, bottom or left) on a game-pad controller as quickly and accurately as possible. Unbeknownst to participants, the location of the target was determined by a 10-item sequence (bottom, top, right, left, right, top, bottom, right, top, left). Participants were trained on the 10-item sequence for five blocks. On the sixth block, a pseudorandom sequence is introduced. In order to compute a procedural memory score, the median RT score of block five was subtracted from that for block six, with a higher, positive score indicating procedural learning. If participants had implicitly started to learn the 10-item sequence, then RTs would slowly decrease from block 1 to 5 (structured sequences) and then increase when completing block 6 (random sequences). If, on the other hand, no procedural learning took place, then median RTs for the sequenced block and the random block would be identical (Siegert, Taylor, Weatherall & Abernethy 2006).

Procedure

Participants were trained and tested on the artificial language over two days. On the first day, participants completed twelve blocks of exposure to the cross-situational learning task, four blocks of which also tested vocabulary and word order. Blocks 1, 2, 4, 5, 7, 8, 10 and 11 were pure exposure blocks with 16 trials each. In blocks 3, 6, 9 and 12, intermingled with the 16 exposure sentences were 24 vocabulary test trials (eight each of noun and case marker test trials and four each of verb and adjective test trials). These four blocks then also included 16 trials of a grammaticality judgment task (GJT) to test participants' acquisition of the basic word order of the sentence. See figure 2. No sentence-scene matches were repeated throughout the experiment. Twenty-four hours later, participants returned to the lab to complete a fifth (delayed) test of vocabulary and word order. In this delayed test, there were 40 vocabulary test trials, consisting of eight each of noun, verb and adjective test trials, 16 case marker test trials and 16 word order test trials. Presentation order of trials within each block was randomized but all participants completed blocks in the same sequence.

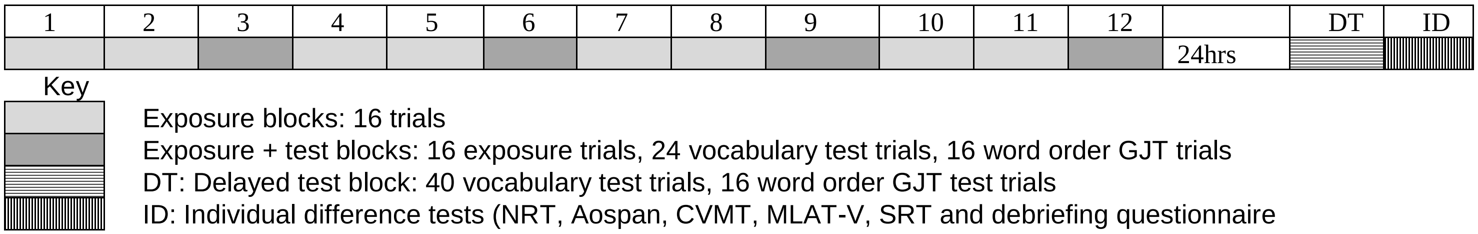


Figure 2. The research design: Participants were exposed to 192 sentences over 12 exposure blocks. Within every third exposure block were intermingled 24 vocabulary test trials and 16 word order GJT trials. The delayed test was administered after 24 hours, together with five individual difference tests and a debriefing questionnaire.

Day 2 comprised the final block of vocabulary and word order test trials, five cognitive tests designed to measure individual differences in memory systems and a debriefing questionnaire (see Appendix D). The five individual difference tests were administered in a randomly-assigned order. Exposure and testing on day 1 took between 90 and 120 minutes, and the delayed test, individual difference measures and debriefing questionnaire on day 2 took around 60 minutes.

Exposure trials. For exposure, there were 12 cross-situational learning blocks with 16 exposure trials each. In each trial, participants were instructed to observe two dynamic scenes on the screen and listen to an artificial language sentence played over headphones. Their task was to decide, as quickly and accurately as possible, which scene the sentence referred to (see Figure 1). Participants received no feedback regarding the accuracy of their choice. Within each block, each alien and action occurred an equal number of times; half the utterances in each block were SOV, the other half OSV. The locations of target and distractor scenes (left or right side of screen) were counterbalanced. In the distractor scene, no actions were the same as in the target scene, and the aliens and their colours were randomly selected.

Vocabulary test trials. In order to make it less likely that participants would know they were being tested, test trials for each lexical category were intermingled in every third cross-situational learning block using the same cross-situational learning task as the exposure trials, but with one exception. The participant saw the two scenes, heard the sentence in the alien language and was required to select the scene to which the sentence referred. However, the target and distractor scenes were exactly the same apart from one piece of information: For testing nouns, the target and distractor scenes were identical except for one of the aliens; for testing verbs, only the scenes' actions differed; for testing adjectives, one of the colours of an alien was changed; and for testing marker words, the two scenes depicted the same aliens performing the same actions but with opposite agent-patient roles. No feedback was provided on response accuracy.

Word order test trials. The acquisition of word order was tested by means of a grammaticality judgment task at the end of every third block of exposure. In this task, participants were told that they would see only one scene and hear a sentence spoken by another alien from a very different planet who was also learning the new language. Their task was to listen carefully and decide if the sentence sounded "good" or "bad" in relation to the artificial language. The pseudowords always matched the scene but only half the sentences were grammatical, following the licensed SOV or OSV word order. The other half contained syntactic violations (*VSO, *VOS, *OVS, *SVO). That is, the task tested sensitivity to correct sequencing of phrases (noun phrase, noun phrase, verb) rather than sequences within phrases (e.g., adjective, noun, marker word, within the noun phase). None of the grammaticality judgment task sentences occurred during exposure trials. Again, no feedback was provided.

Results

Performance on Exposure Trials

In order to ascertain when learning had taken place during the exposure blocks, a one-sample t-test was conducted to compare the mean scores for each block to a chance score of .5. Table 1 gives a summary of the findings for the one-sample t-test and the descriptive statistics. Participants performed significantly above chance from block two (M = .57, SD = .18) onwards, 95% CI [.53 to .62], $t(63) = 3.27$, $p = .002$. Thus, 32 trials of exposure (without feedback) were enough to lead to above-chance performance in the cross-situational learning task.

Table 1

Summary of Descriptive Statistics and One-Sample T-Tests on Mean Scores for Each Block of Exposure

Exposure block	*M*	*SD*	95% CI *Lower*	95% CI *Upper*	*T*	*df*	*Sig. (2-tailed)*
Exposure block 1	.53	.15	.49	.57	1.498	63	.14
Exposure block 2	.57	.18	.53	.62	3.265	63	.002
Exposure block 3	.70	.18	.66	.74	8.883	63	<.001
Exposure block 4	.76	.19	.71	.80	10.812	63	<.001
Exposure block 5	.75	.21	.70	.81	9.697	63	<.001
Exposure block 6	.79	.20	.74	.84	11.900	63	<.001
Exposure block 7	.79	.22	.74	.85	10.622	63	<.001
Exposure block 8	.82	.22	.76	.88	11.528	63	<.001
Exposure block 9	.84	.19	.80	.89	14.170	63	<.001
Exposure block 10	.83	.23	.78	.89	11.803	63	<.001
Exposure block 11	.85	.21	.80	.90	13.556	63	<.001
Exposure block 12	.85	.20	.80	.90	14.078	63	<.001

Performance on Test Trials

As the grammaticality judgment test was a two-way forced choice between grammatical and ungrammatical sentences, we first conducted a signal detection analysis as an extra check that accuracy scores were a reflection of discrimination, rather than response bias. In order to do this, d' (prime) and C (bias) scores were calculated for each of the word order tests 1-4. They were then entered into repeated measures ANOVAs. Results, with Greenhouse-Geisser correction, showed that d' was significant, $F(2.270, 143)= 14.585$, $p < 0.001$ but C was not, $F(2.168, 136.5)= 1.856$, $p = 0.157$. This indicates that the accuracy scores are a true reflection of discrimination and will therefore be used for the further analyses. Employing accuracy rather than d' ensures that the task is more closely comparable to the accuracy scores of the vocabulary tests.

. We next looked at the effects of learning syntactic word order, nouns, verbs, adjectives and case markers over the two days by conducting repeated measures ANOVAs with reverse Helmert contrasts on accuracy for learning each aspect of the language over tests 1 to 5. Reverse Helmert contrasts indicate the point at which there is significant step-change in learning from previous blocks to the subsequent block. For a summary of the reverse Helmert contrasts see appendix E. Greenhouse-Geisser correction was applied when the test of Sphericity was significant. The results are shown in Figure 3 and Table 2. Word order, nouns and verbs all improved significantly over the five tests.

For word order, participants' test scores improved significantly between each test, including a significant improvement from day 1 to day 2, $F (1, 63) = 4.136$, $p = .046$, $\eta_p^2 = .062$. A one-sample t-test shows that word-order scores were significantly above chance from test block 1 onwards (See Appendix F for one-sample t-tests), (*M* = .76, *SD* = .19), 95% CI [.72 to .81], $t(63) = 10.9$, $p < .001$.

Nouns also improved significantly between each of the first four tests on day 1 but then saw a significant drop in scores from day 1 to day 2, $F(1, 63) = 5.650$, $p = .021$, $\eta_p^2 = .082$. One-sample t-tests show that results were significantly above chance from noun test 2 (*M* = .60, *SD* = .19, 95% CI [.55 to .65], $t(63) = 4.16$, $p < .001$) onwards, including for noun test 5 (*M* = .68, *SD* = .25, 95% CI [.47 to .83], $t(63) = 5.69$, $p < .001$).

Verbs showed no significant improvement between the first four tests, although a one-sample t-test shows that this is because participants learned verbs early and were already significantly above chance by verb test 1, (*M* = .70, *SD* = .25, 95% CI [.63 to .76], $t(63) = 6.16$, $p < .001$). Interestingly, there was a significant improvement in verb scores from day 1 to day 2, $F(1, 63) = 19.028$, $p < .001$, $\eta_p^2 = .23$.

For adjectives, there was no significant improvement over the five tests, although scores for adjective test 4 were a significant improvement on the mean scores of the previous three adjective tests, $F(1, 63) = 8.401$, $p = .005$, $\eta_p^2 = .118$. No significant change in accuracy from day 1 to day 2 occurred, with one-sample t-tests confirming that both adjective test 4, (*M* = .64, *SD* = .27), 95% CI [.58 to .71], $t(63) = 4.22$, $p < .001$) and test 5, (*M* = .55, *SD* = .22), 95% CI [.53 to .65], $t(63) = 2.92$, $p = .005$) were both significantly above chance.

For case markers, no significant step change in improvement over time occurred. Test 5, was, however, just significantly above chance, (*M* = .54, *SD* = .16), 95% CI [.50 to .58], $t(63) = 2.06$, $p = .043$) . In the ANOVA, there were no significant main effects or interactions with massed/distributed exposure condition, with all $F < 2.3$ and all $p > .13$.

Table 2. *Summary of Repeated Measures ANOVA over Tests 1 to 5 Showing Effect for Test Block*

Effect	*Df*	*F*	*P*	η_p^2
Word order	2.812, 177.170	7.20	<.001	.10
Noun	4, 252	12.4	<.001	.17
Verb	4, 252	5.27	<.001	.077
Adjective	4, 252	2.30	.059	.035
Case marker	4, 252	.86	.59	.013

There were no significant differences in test results for linguistics or language students, those who had at least one other language to intermediate level, or for knowledge of case marked languages. For a summary of the results see Appendix G.

The debriefing questionnaire offered insights into the extent to which learning was verbalizable, as well as increased detail as to what exactly participants learned from the study. For word-order, 47 participants (73%) could identify the basic syntax of noun phrase, noun phrase, verb, although only 17 participants (26.6%) correctly identified the exact word order of the noun phrase (adjective, noun, marker). Regarding case markers, only 14 participants (21.9%) could correctly identify that they referred to the agent and patient of the sentence. In terms of which lexical item participants noticed first, 38 (59.4%) reported that they noticed verbs first, with only 15 (23.4%) reporting noticing nouns first.

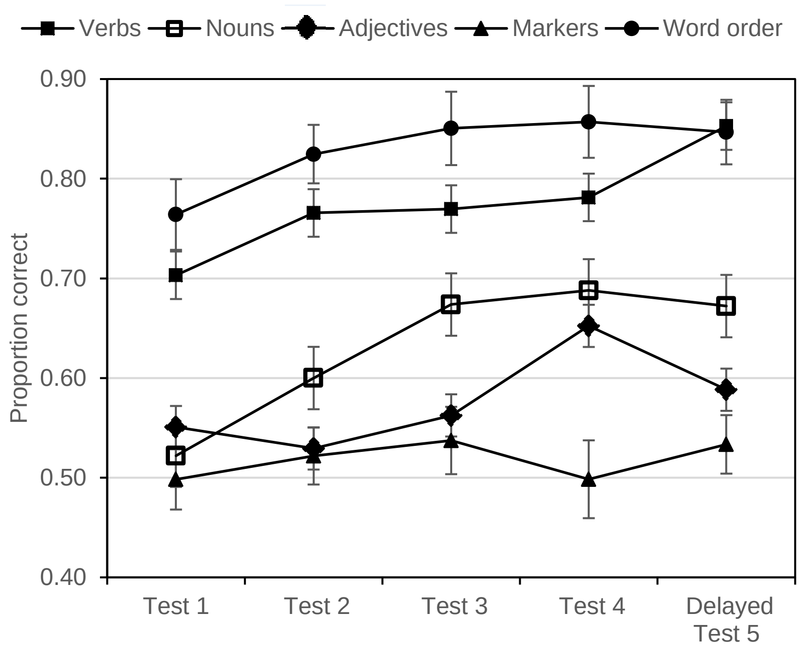


Figure 3. Proportion of correct trials across the five tests. Tests 1 to 4 were completed on day 1, with test 5 administered after a 24-hr delay. The error bars represent standard errors of the mean.

Determining Relations Between Learning Different Information Types

In order to determine which underlying components drove performance in the task – whether learning was independent or interdependent for different types of information – we conducted an exploratory principal components analysis for word order, nouns, verbs, adjectives, and case markers on test performance for tests 1 to 4 combined and for the final, delayed test. For both tests 1 to 4 combined and test 5, there were two components with eigenvalues greater than 1, and the loadings of the individual tests on these components, with varimax rotation, showed a simple solution (i.e., each test loaded > 0.4 on only one component). Due to this simple structure, we did not separately relate the learning of types of information to the individual difference measures. The components and their loadings are shown in Table 4 and Table 5.

For tests 1 to 4 combined, the first component related to learning nouns, case markers and adjectives, and the second component related to learning word order and verbs. This indicated that performance across the five information types was effectively explained by two aspects of the data: The first relates to learning the vocabulary items of nouns and adjectives and how the marker words affected the role of the adjective-noun phrases. The second indicated a close relation between learning the identities of verbs and learning that the word order of sentences was verb-final. We return to this point in the Discussion.

For the delayed test 5, a similar picture emerged. The first component related to learning nouns, adjectives, and marker words, and the second component related to learning word order and verbs.

Table 4

Loadings of the Five Tests on the Two Principal Components for Tests 1-4 Combined

Tests 1-4	First component	Second component
Noun	**.626**	.145
Adjective	**.763**	.212
Case marker	**.692**	-.210
Verb	-.034	**.848**
Word order	.140	**.690**

Table 5

Loadings of the Five Delayed Tests on the Two Principal Components for Test 5

Test 5	First component	Second component
Noun	**.778**	.104
Adjective	**.769**	.034
Case marker	**.604**	.081
Verb	.322	**.718**
Word order	-.090	**.873**

Individual Difference Measures

Scores for measures of the five individual cognitive measures were converted into z-scores. (see Table 6)

Table 6

Descriptive Statistics for ID Measures

ID measure	*M*	*SD*	*SE*	Range
NWT	21.4	2.68	0.34	12
Aospan	38.0	19.4	2.44	75
CVMT	1.80	0.55	0.07	2.94
MLAT-V	16.8	4.89	0.61	19
SRT	54.2	35.6	4.45	185

Note: NRT = Non-word repetition Test; Aospan = Automated Operation Span task; CVMT = Continuous Visual Memory Test; MLAT-V = Modern Languages Aptitude Test, Part V; SRT = Serial Reaction Time task.

We conducted a series of stepwise linear regressions to determine the relations between the individual difference measures and the two components of learning the language derived from the principal component analysis. All five ID measures were included in the same level. Again, we distinguished performance on day 1 (so tests 1-4 combined) and performance on day 2 (test 5) as the dependent variables in separate regressions, and the five ID measures (NRT, Aospan, CVMT, MLAT-V and SRT) as the independent variables. No predicting variable was found for test 5: word order, verbs and case markers. Table 7 illustrates the predicting variables for each of the principal components for tests 1-4 and test 5.

Table 7

Summary of Step-Wise Linear Regression with Principal Component Analysis Variable Scores for Lexical and Word Order Tests 1-4 Combined and Test 5 as the Dependent Variables and the Five ID Measures (NRT, Aospan, CVMT, MLAT-V and SRT) as the Independent Variables

Component		*b*	*SE b*	*B*
Test 1-4: component 2 (word	Constant	.01	.12	
order and verbs)	MLAT-V;	.25	.16	.25**
Test 1-4: component 1 (nouns,	Constant	-.011	.12	
adjectives and markers)	SRT	.35	0.12	0.35***
Test 5: component 1 (nouns,	Constant	.009	.12	
adjectives and markers)	MLAT-V	.36	.12	.36**

Discussion

In this study, we replicated and extended Rebuschat et al. (submitted) by investigating whether, and in which order, adult learners could acquire the syntax and vocabulary of a novel language via cross-situational learning, without feedback and without any explicit instruction about the structure of the language or its vocabulary. We then further explored this paradigm by providing a delayed post-test after 24 hours to determine whether any acquired knowledge had been maintained. Furthermore, we investigated how learning of syntax and of vocabulary cohered and which cognitive individual differences (particularly working memory capacity, PSTM, declarative and procedural memory) affected learning of different aspects of this artificial language.

Learning Under Cross-Situational Conditions

Our results indicated that adult learners can rapidly acquire both vocabulary and certain aspects of syntax of the language simultaneously, in line with Rebuschat et al. (submitted). While this study did not investigate the underlying mechanisms which facilitated the learning, these findings are consistent with the use of a combination of statistical learning mechanisms (Yu & Smith, 2007), syntactic and semantic bootstrapping (Abend et al., 2017; Gleitman, 1990) and a propose-but-verify procedure (Trueswell, Medina, Hafri & Gleitman, 2013), with initial learning occurring implicitly (consistent with associative learning mechanisms) and more explicit, top-down knowledge then interacting and working in unison with that unconscious learning (consistent with strategic approaches to learning). Replicating the findings from Rebuschat et al. (submitted), verbs and basic word order were learned to a level above chance first, followed by nouns[2], then adjectives and finally case markers (though while case markers did reach a significant level above

[2] We do not know if participants learned proper or common nouns. However, we know that they learned the labels for the aliens and that these are nouns.

chance for test 5, the lack of main effects on the repeated measures ANOVA suggests we ought to conclude that learning did not occur). The fact that verbs were learned to a level above chance before nouns may be accounted for by the increased salience of the sentence-final verbs (Fernald, McRoberts & Herrera, 1992; Shoemaker & Rast, 2013), consistent with recency effects in sequence processing that Freudenthal, Pine, Aguado-Orea and Gobet (2007) demonstrate can explain a variety of morphological and grammatical effects in language learning and Jones and Rowland (2017) show can explain phonotactic effects on word learning. However, it is possible that the verb-final position and its immovability provided an extra recency advantage that is not available in language learning in general and L1 acquisition in particular by increasing the likelihood of verbs and their position in the sentence reaching the level of awareness and thus providing opportunities for more metacognitive strategy use. Another possible contributor to the order of acquisition effects observed in this study is that the artificial language included eight nouns but only four verbs, potentially making the latter easier to learn. Yet, despite these caveats, the similarity to the order of acquisition in verb-dominant languages (Choi & Gopnik, 1995; Tardif, 1996) raises the question of whether and to what extent the findings of this particular study can be generalised to first language acquisition. While all our participants were adults who already possessed syntax, vocabulary and often metalinguistic knowledge from their first languages, we contend that the underlying mechanisms of cross-situational learning are similar for both first and subsequent languages under immersion conditions, albeit with differing levels of pre-existing knowledge and cognitive development affecting the relative roles of explicit and implicit learning.

The Durability of Cross-Situational Learning

We built upon the previous study to investigate whether learning under cross-situational learning conditions is durable over time. Importantly, the results showed that the learning effects can be retained overnight, and performance actually improved with tests for verbs, word order and case markers, albeit for the latter non-significantly. This is an important methodological observation as the majority of studies in cross-situational learning do not have a delayed post-test, which means that it is unclear whether the learning is robust. By including a 24-hr delayed post-test, we show that learning is indeed robust and that this applies to words and syntax. It is recommended, therefore, that future studies into cross-situational learning include delayed post-tests to show that learning is robust and to catch any learning effects brought on through consolidation.

Coherence of Syntax and Vocabulary

Regarding the coherence of learning of syntax and vocabulary, we found that acquisition of word order and verb learning were interdependent. As to why word order and verbs cohered, this could at least be partially explained by the nature and simplicity of the word order test. The grammaticality judgment task could be completed successfully by noticing that in the ungrammatical sentences (*OVS, *VSO, *SVO, *VOS) the last word of the sentence was mono-syllabic (the post-nominal case marker) rather than bi-syllabic (all the content words but particularly verbs). In the debriefing questionnaire, a large majority of participants reported that the first thing they explicitly noticed was that the final word was an action. Questionnaires are only able to capture verbalizable, explicit knowledge, so it is impossible to tell from this instrument what was happening at a subconscious, implicit level. However, taken together with the results showing a coherence of learning of word order and verbs, and consistent with the findings of Khoe et al. (2019) that ambiguous learning situations rely more on associative learning rather than hypothesis-testing strategies, this tentatively suggests that participants first learned one referent for the final position, and therefore its word category (captured in the grammaticality judgment results), before going on to learn the referents for the other actions. In addition, we found that nouns, adjectives and case markers were also interdependent but acquired somewhat independently of verbs and word order. These lexical categories comprise the noun phrase, and it is therefore perhaps unsurprising that success with adjectives and case markers corresponds to success with nouns, on which their meanings and syntactic roles depend. In sum, these results support a view of language in which the syntactic knowledge associated with case markers begins to develop only after the syntactic roles and semantic meanings of a core vocabulary of content words has been learned (Bannard, Lieven & Tomasello, 2009).

Individual Differences in Cross-Situational Learning

The final objective of this study was to investigate the role of short and long-term memory systems (PSTM, working memory capacity, declarative memory and procedural memory) in the acquisition of this artificial language under cross-situational learning conditions. Neither working memory capacity, as measured by the Aospan, nor PSTM, as measured by NRT, predicted success on the lexical test scores. This mirrors recent findings into incidental learning conditions (Hamrick, 2015; Tagarelli et al., 2015). It is possible that working memory capacity is utilised mainly when language is learned under explicit conditions.

For tests 1 to 4 combined, regressions of the principal component factor scores revealed that component 1, which included nouns, adjectives and case markers, was predicted by SRT, a measure used to assess procedural memory, while component 2, which included verbs and word order, was predicted by MLAT-V, a measure used to assess declarative memory. On the surface these results do not tally with proposals that associate vocabulary learning primarily with declarative and grammar learning with procedural memory (Ullman, 2004), nor with declarative memory being responsible for all the early stages of learning including grammar (Hamrick, 2015; Morgan-Short et al., 2014; Ullman, 2016). However, one possible interpretation of these results involves the amount of attentional resources dedicated to each lexical category at different times during the study. As mentioned earlier, Ullman and Lovelett (2016) state that altering the learning conditions may force more of a reliance on either declarative or procedural memory. While no deliberate manipulation occurred in this study, it is possible that this happened as a by-product of the nature of the exposure task and lexical tests. In order to complete the exposure blocks on the first day, participants may have explicitly and strategically focused on verbs (and associated word order), and so success with these was predicted by declarative memory. Nouns, adjectives and case markers, on the other hand, received less explicit attention and so were learned more incidentally. This might explain why success in these lexical categories, therefore, was predicted by procedural memory. On the second day, with verbs and word order reaching ceiling effects, participants could shift their attention to nouns, adjectives and case markers. As a result, success with these categories was now predicted by declarative memory. While we acknowledge that there is a possibility that this is a consequence of the construction of this particular artificial language, we believe that it is more analogous to the shifting of attention at different stages of real language learning. A further study in which attention is directed towards certain lexical categories might shed light on this aspect of the declarative/procedural model.

An alternative explanation for these results is that acquiring the lexical categories which comprised the noun phrase required more pattern learning than for verbs and for completion of the grammaticality judgment task, which tested knowledge of basic word order. Adjectives were optional in our artificial language; nouns could therefore adopt a number of different positions in the sentence, and case markers further influenced the position of nouns and adjectives. Thus, in this interpretation success with the components of the noun phrase was predicted by procedural memory. Performance on verb tests and the grammaticality judgment task, on the other hand, only required attention to the last word in the sentence, and so their syntactic roles in the sentence were perhaps less important than their semantic meanings. Hence, success with this component was predicted by declarative memory. Ullman and Lovelett (2016) do indeed assert that procedural memory is responsible for the more rule-governed aspects of vocabulary acquisition. Therefore, it is possible that in the early stages of the acquisition process under cross-situational learning conditions, procedural memory plays an important part in determining the syntactic roles of vocabulary (Evans, Saffran, Robe-Torres, 2009). Turning to the 24-hr delayed test 5, component 1, which contained adjectives, nouns and markers, was predicted by MLAT-V, the measure of declarative memory. Follow-up analyses from debriefing questionnaires on the awareness of rule knowledge of word order and the function of case markers suggest that declarative memory may play a role on delayed tests when learners become aware of the rules, that is when rules become explicit and verbalizable. Component 2, which contained verbs and word order, did not have a predictor. This may have been affected by a ceiling effect on scores for verbs and the grammaticality judgment task. While both these explanations are plausible, neither is fully satisfactory and nor do either account for the early stage dominance of declarative memory (Ullman, 2016). Further research is needed before the relationships between declarative memory, procedural memory and early vs. late stages, attention vs. less attention, and arbitrary vs. rule-governed items are fully understood.

One issue with the research tools used in the current study is what was actually being tested with the SRT task, statistical learning or procedural memory. In order to best tap into procedural memory as a long-term memory system, one solution would be to include a delayed post-test on the SRT. This has been done with several recent studies (Desmottes, Meulemans, & Maillart, 2016; Desmottes, Meulemans, Patinec & Maillart, 2017; Hedenius et al., 2011) into specific language impairment (SLI), which has been hypothesized to be due to a deficit in procedural memory (Ullman & Pierpont, 2005). It is possible, therefore, that individual differences for a delayed post-test for SRT would be better predictors for performance on delayed tests for nouns, adjectives and case markers.

Conclusion

Our study confirms that it is possible for adults to learn syntax and vocabulary simultaneously under cross-situational learning conditions and that the order of acquisition follows verb-dominant language acquisition, but also that the learning effect persists after 24 hours. The durability aspect is important from a methodological perspective as not only does this show that knowledge is robust but also that without a delayed test, learning can be in some cases underreported. Moreover, the patterns of results we found for this verb-final language in our experimental paradigm did not neatly correspond with a distinction between grammar and vocabulary learning (e.g., Ullman, 2004), with word order being related to verb acquisition, and case marking being related to noun and adjective learning. Complex interactions between lexical categories and grammar do not appear to lend themselves to a clear distinction in acquisition of these sources of linguistic knowledge, nor in the individual differences that predicted them. Future studies into language learning in adults under cross-situational learning conditions should continue to investigate individual differences and how they cohere, as it is a central question to both language learning theory and pedagogical applications such as computer assisted language learning design (Meurers, De Kuthy, Nuxoll, Rudzewitz, & Ziai, 2019).

A limitation to this study is the number of participants (n=64) compared to the number of analyses that were carried out. Caution should therefore be exercised before generalising these findings. The extent to which the effects we observe here are due to participants' first language exposure, or are general across language learning, remains a moot point. However, this paradigm offers further opportunities to investigate cross-linguistic differences in early-stage acquisition of vocabulary and grammar (see Fedzechkina, Newport & Jaeger, 2016 for an overview), and the potential support and interference of different aspects of learning across related and unrelated languages.

References

Abend, O., Kwiatkowski, T., Smith, N. J., Goldwater, S., & Steedman, M. (2017). Bootstrapping language acquisition. *Cognition*, *164*, *116-143. https://doi.org/10.1016/j.cognition.2017.02.009*

Alario, F., & Cohen, L. (2004). Closed-class words in sentence production: Evidence from a modality-specific dissociation. *Cognitive Neuropsychology,, 21*(8), 787-819. https://doi.org/10.1080/02643290342000410

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, *27*(5), 586-595.

Bannard, C., Lieven, E., & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences, 106*(41), 17284-17289. https://doi.org/10.1016/0749-596x(88)90028-9

Bates, E., & Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes, 12*, 507–584. https://doi.org/10.1080/016909697386628

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In MacWhinney, B. (Ed.), *Mechanisms of language acquisition* (pp. 157-193). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In Gernsbacher, M (Ed.*), Handbook of Psycholinguistic*s (pp. 945-984). San Diego, CA: Academic Press.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition, 41*(2), 281-296. https://doi.org/10.3758/s13421-012-0262-9

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test*. New York: Psychological Corporation.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22*(03), 497-529. https://doi.org/10.1017/s0305000900009934

Coughlin, C. E., & Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics, 34*(3), 615-646. https://doi.org/10.1017/s0142716411000890

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning,* 55(S1), 1-25. https://doi.org/10.1111/j.0023-8333.2005.00294.x

Desmottes, L., Meulemans, T., & Maillart, C. (2016). Later learning stages in procedural memory are impaired in children with specific language impairment. *Research in Developmental Disabilities, 48,* 53-68. https://doi.org/10.1016/j.ridd.2015.10.010

Desmottes, L., Meulemans, T., Patinec, M. A., & Maillart, C. (2017). Distributed training enhances implicit sequence acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 60*(9), 2636-2647. https://doi.org/10.1044/2017_jslhr-l-16-0146

Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*, Routledge. https://doi.org/10.4324/9781410613349

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia, 49*(5), 1322-1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics*, *32*, 17-44. https://doi.org/10.1017/s0267190512000025

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*(2), 321-335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Fedzechkina, M. A. R. Y. I. A., Newport, E. L., & Jaeger, T. F. (2016). Miniature artificial language learning as a complement to typological data. *The usage-based study of language learning and multilingualism.*

Fernald, A., McRoberts, G., & Herrera, C. (1992). Prosodic features and early word recognition. In *8th International Conference on Infant Studies*, Miami, FL.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458. https://doi.org/10.1037/0278-7393.28.3.458

Freudenthal, D., Pine, J. M., Aguado-Orea, J., & Gobet, F. (2007). Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science*, *31*(2), 311-341. https://doi.org/10.1207/s15516709cog0000_47

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529-534. https://doi.org/10.1073/pnas.012611199

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in cognitive sciences*, *19*(3), 117-125. https://doi.org/10.1016/j.tics.2014.12.010

Frost, R. L., & Monaghan, P. (2016). Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. *Cognition, 147*, 70-74. https://doi.org/10.1016/j.cognition.2015.11.010

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*(1), 83-94. https://doi.org/10.3758/bf03210559

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. London: Psychological Corporation.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3-55. https://doi.org/10.1207/s15327817la0101_2

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps Promote Abstraction in Language-Learning Infants. *Psychological Science*, 17, 670-674. https://doi.org/10.1111/j.1467-9280.2006.01764.x

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109-135. https://doi.org/10.1016/s0010-0277(99)00003-7

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences, 38*, 44-53. https://doi.org/10.1016/j.lindif.2015.01.019

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(7), 1213-1236. https://doi.org/10.1080/02724980343000071

Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9-15. https://doi.org/10.1016/j.lindif.2015.10.003

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D. & Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with specific language impairment. *Research in Developmental Disabilities*, *32*(6), 2362-2375. https://doi.org/10.1016/j.ridd.2011.07.026

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, *130*(4), 658. https://doi.org/10.1037/0096-3445.130.4.658

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology*, 98, 1-21. https://doi.org/10.31234/osf.io/wdbpc

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321-340. https://doi.org/10.1016/j.cognition.2010.05.011

Khoe, Y. H., Perfors, A., & Hendrickson, A. T. (2019). Modeling individual performance in cross-situational word learning. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*. https://doi.org/10.31234/osf.io/4rtw9

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171. https://doi.org/10.1037/a0025405

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184-193. https://doi.org/10.1111/cdev.12461

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit*? Neuroscience Letters*, *461*(2), 145-149. https://doi.org/10.1016/j.neulet.2009.06.030

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42. https://doi.org/10.1016/s0010-0277(02)00004-5

Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. *Biological psychology*, *135*, 180-193. https://doi.org/10.1016/j.biopsycho.2018.04.001

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *The Modern Language Journal*, *97*(3), 634-654. https://doi.org/10.1111/j.1540-4781.2013.12030.x

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861-883. https://doi.org/10.3758/s13423-013-0565-2

Lum, J. A., 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*(1), 96-107. https://doi.org/10.3109/13682820902752285

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676. https://doi.org/10.1037/0033-295x.101.4.676

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–210). Amsterdam: Benjamins. https://doi.org/10.1075/lllt.2.12mac

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*(3), 501-533. https://doi.org/10.1111/j.1467-9922.2010.00565.x

Marcus, G. (1996). Why do children say "breaked"? *Current Directions in Psychological Science*, 5, 81–85. https://doi.org/10.1111/1467-8721.ep10772799

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379-413. https://doi.org/10.1017/s0272263112000125

McGregor, K. K., Sheng, L. I., & Smith, B. (2005). The precocious two-year-old: status of the lexicon and links to the grammar. *Journal of Child Language*, *32*(3), 563-585. https://doi.org/10.1017/s0305000905006926

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014-9019. https://doi.org/10.1073/pnas.1105040108

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics*, *39*, 161-188. https://doi.org/10.1017/s0267190519000126

Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. A. Hirsch-Pasek & R. M. Golinkoff, (Eds.) .*Action Meets Word: How Children Learn Verbs* (pp. 31-63). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.003.0002

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302-331. https://doi.org/10.1111/j.1467-9922.2010.00626.x

Monaghan, P., & Mattock, K. (2012). Integrating constraints for learning word–referent mappings. *Cognition, 123*(1), 133-143. https://doi.org/10.1016/j.cognition.2011.12.010

Monaghan, P., Mattock, K., Davies, R. A., & Smith, A. C. (2015). Gavagai Is as Gavagai Does: Learning Nouns and Verbs From Cross-Situational Statistics. *Cognitive science, 39*(5), 1099-1112. https://doi.org/10.1111/cogs.12186

Monaghan, P., Schoetensack, C., & Rebuschat, P. (2019). A single paradigm for implicit and statistical learning. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12439

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, *17*(1), 56-72. https://doi.org/10.1017/s1366728912000715

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1-32. https://doi.org/10.1016/0010-0285(87)90002-8

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*(3), 331-347. https://doi.org/10.1016/0749-596x(91)90040-q

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604-607. https://doi.org/10.1126/science.1072901

Rebuschat, P., Monaghan, P., & Schoetensack, C. (submitted). Learning vocabulary and grammar simultaneously from cross-situational statistics.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning, 47*(1), 45-99. https://doi.org/10.1111/0023-8333.21997002

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*(2), 235-268. https://doi.org/10.1017/s0272263105050126

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart & J. L. McClelland & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 110–146). Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 1926-1928. https://doi.org/10.1126/science.274.5294.1926

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*(8), 2022-2039. https://doi.org/10.1016/j.lingua.2010.02.004

Schoetensack, C. (2015). The role of working memory in vocabulary acquisition: A cross-situational word learning experiment. (Unpublished master's thesis), Lancaster University, Lancaster, United Kingdom.

Schvaneveldt, R. W., & Gómez, R. L. (1998). Attention and probabilistic sequence learning. *Psychological Research*, *61*(3), 175-190. https://doi.org/10.1007/s004260050023

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*(2), 163-180. https://doi.org/10.1016/j.cognition.2011.10.010

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275*(5306), 1599-1603. https://doi.org/10.1126/science.275.5306.1599

Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, *17*(3), 247-271. https://doi.org/10.1111/j.1532-7078.2011.00085.x

Shoemaker, E., & Rast, R. (2013). Extracting words from the speech stream at first exposure. *Second Language Research*, *29*(2), 165-183. https://doi.org/10.1177/0267658313479360

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Phil. Trans. R. Soc. B*, *372*(1711), 20160059. https://doi.org/10.1098/rstb.2016.0059

Siegert, R. J., Taylor, K. D., Weatherall, M., & Abernethy, D. A. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology*, *20*(4), 490-495. https://doi.org/10.1037/0894-4105.20.4.490

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

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558-1568. https://doi.org/10.1016/j.cognition.2007.06.010

Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in Cognitive Sciences*, *18*(10), 520-525. https://doi.org/10.1016/j.tics.2014.06.005

Speidel, G. E. (1993). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language*, *13*(37), 69-91. https://doi.org/10.1177/014272379301303705

Tagarelli, K. M., Borges Mota, M.& Rebuschat, P . (2015). Working memory, learning conditions, and the acquisition of L2 syntax. In W. Zhisheng, M. Borges Mota, & A. McNeill (Eds.) *Working memory in second language acquisition and processing: Theory, research and commentary* (pp. 224 – 247). Buffalo, NY: Multilingual Matters. https://doi.org/10.21832/9781783093595-017

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*(3), 492. https://doi.org/10.1037/0012-1649.32.3.492

Trahan, D. E., & Larrabee, G. J. (1988). *Continuous visual memory test: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive psychology*, *66*(1), 126-156. https://doi.org/10.1016/j.cogpsych.2012.10.001

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1), 231-270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2016). The declarative/procedural model: a neurobiological model of language learning, knowledge and use. *The Neurobiology of Language*, 953-968. https://doi.org/10.1016/b978-0-12-407794-2.00076-6

Ullman, M. T., & Lovelett, J. T. (2016). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, *34*(1), 39-65. https://doi.org/10.1177/0267658316675195

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*(3), 399-433. https://doi.org/10.1016/s0010-9452(08)70276-4

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498-505. https://doi.org/10.3758/bf03192720

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, *141*, 65-82. https://doi.org/10.1016/j.jecp.2015.06.015

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230. https://doi.org/10.1016/j.jml.2016.10.001

Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science*, *38*(4), 757-774. https://doi.org/10.1111/cogs.12092

Vuong, L.C., Meyer, A.S. & Christiansen, M.H. (2016). Concurrent learning of adjacent and nonadjacent dependencies. *Language Learning*, 66, 8-30. https://doi.org/10.1111/lang.12137

Walker, N., Schoetensack, C., Monaghan, P., & Rebuschat, P. (2017). Simultaneous acquisition of vocabulary and grammar in an artificial language learning task. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 1307-1312). Austin, TX: Cognitive Science Society.

Williams, J. N., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua*, 118(4), 522-553. https://doi.org/10.1016/j.lingua.2007.03.003

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414-420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Appendix A

Pseudoword lexicon (adapted from Author, xxxx):

Bisyllabic words used content words (nouns, verbs, and adjectives): *barget*, *bimdah*, *chelad*, *dingep*, *fisslin*, *goorshell*, *haagle*, *jeelow*, *limeber*, *makkot*, *nellby*, *pakrid*, *rakken*, *sumbark*

Monosyllabic words used as function words (case markers): *tha*, *noo*

Appendix B

Grammatical sentence patterns used in the cross-situational learning exposure task.

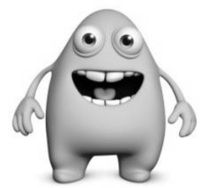
Word order	**Syntactic pattern**			**Number per 16-trial block**
	First phrase	**Second phrase**	**Third phrase**	
SOV	NP_{subj} (Adj N $Case_{nom}$)	NP_{obj} (Adj N $Case_{acc}$)	VP (V)	2
	NP_{subj} (Adj N $Case_{nom}$)	NP_{obj} (N $Case_{acc}$)	VP (V)	3
	NP_{subj} (N $Case_{nom}$)	NP_{obj} (Adj N $Case_{acc}$)	VP (V)	1
	NP_{subj} (N $Case_{nom}$)	NP_{obj} (N $Case_{acc}$)	VP (V)	2
OSV	NP_{obj} (Adj N $Case_{acc}$)	NP_{subj} (Adj N $Case_{nom}$)	VP (V)	2
	NP_{obj} (Adj N $Case_{acc}$)	NP_{subj} (N $Case_{nom}$)	VP (V)	3
	NP_{obj} (N $Case_{acc}$)	NP_{subj} (Adj N $Case_{nom}$)	VP (V)	1
	NP_{obj} (N $Case_{acc}$)	NP_{subj} (N $Case_{nom}$)	VP (V)	2

Note: NP_{subj} = Subject Noun Phrase; NP_{obj} = Object Noun Phrase; VP = Verb Phrase; N = Noun; V = Verb; Adj = Adjective; $Case_{subj}$ = Case marker for subject; $Case_{obj}$ = Case marker for subject.

Appendix C

Alien characters used in the experiment.

Aliens 1 to 8 were used in exposure and testing trials and appeared in red or blue.

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8. 

Appendix D

Debriefing Questionnaire

1. During the different trials of this study, you saw two scenes and heard one sentence. Your task was to choose which scene the sentence referred to. How did you decide which scene the sentence described? Did you just guess throughout the experiment, or did you follow any particular strategies? If so, what strategies did you follow?

2. Do you think the way you made decision on the scenes changed throughout the experiment?

3. Did you feel you learned the names of the aliens?

4. Did you feel you learned the names of the colours?

5. Did you feel you learned the names of the actions?

6. Did you notice what type of word always preceded “tha”? If so, please write down the tope of word below. Also, please tell us when you noticed during the experiment (e.g. before the break etc)? If you did not notice anything, please write down your best guess for what type of word precedes “tha”.

7. Did you notice what type of word always preceded “noo”? If so, please write down the tope of word below. Also, please tell us when you noticed during the experiment (e.g. before the break etc)? If you did not notice anything, please write down your best guess for what type of word precedes “noo”.

8. Do you think that “tha” had a particular grammatical function?

9. Do you think that “noo” had a particular grammatical function?

10. Did you notice any particular patterns or rules in the language while performing the task?

11. What do you think the aim of this study was?

Appendix E

Summary of reverse Helmert contrasts for the repeated measures ANOVA for nouns, verbs, adjectives, case markers and word order

	Nouns			Verbs			Adjectives			Case markers			Word Order		
	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared
Test 2 vs test 1	7.10	.010	.10	3.07	.084	.057	0.27	.61	.004	0.96	.33	.015	8.60	.005	.12
Test 3 vs previous	17.0	<.001	.21	1.06	.31	.027	0.38	.54	.006	0.65	.42	.010	7.49	.008	.11
Test 4 vs previous	20.4	<.001	.25	3.93	.052	.059	8.4	.005	.12	0.31	.58	.005	7.61	.008	.11
Test 5 vs previous	5.65	.021	.082	19.0	<.001	.23	0.29	.60	.005	2.07	.16	.032	4.14	.046	.062

Appendix F

Descriptive statistics and summary of one-sample t-tests on mean scores for each test block.

			95% CI	95% CI			
	M	*SD*	Lower	Lower	*t*	*df*	Sig. (2-tailed)
Noun test 1	.53	.18	.48	.57	1.14	63	.26
Noun test 2	.60	.19	.55	.65	4.16	63	<.001
Noun test 3	.68	.24	.62	.74	5.91	63	<.001
Noun test 4	.69	.23	.64	.75	6.82	63	<.001
Noun test 5	.68	.25	.62	.74	5.69	63	<.001
Verb test 1	.70	.25	.63	.76	6.16	63	<.001
Verb test 2	.76	.26	.69	.82	7.95	63	<.001
Verb test 3	.76	.28	.69	.83	7.49	63	<.001
Verb test 4	.79	.29	.72	.86	8.09	63	<.001
Verb test 5	.85	.22	.79	.90	12.8	63	<.001
Adjective test 1	.55	.26	.49	.61	1.58	63	.12
Adjective test 2	.53	.24	.47	.59	0.93	63	.36

Adjective test 3	.56	.26	.50	.63	1.95	63	.055
Adjective test 4	.64	.27	.58	.71	4.22	63	<.001
Adjective test 5	.59	.24	.53	.65	2.92	63	.005
Case marker test 1	.49	.17	.45	.54	-0.26	63	.80
Case marker test 2	.53	.16	.49	.56	1.32	63	.19
Case marker test 3	.53	.18	.49	.58	1.39	63	.17
Case marker test 4	.50	.22	.45	.56	0.086	63	.93
Case marker test 5	.54	.16	.50	.58	2.06	63	.043
Word order test 1	.76	.19	.72	.81	10.9	63	<.001
Word order test 2	.82	.20	.77	.87	13.0	63	<.001
Word order test 3	.84	.22	.79	.90	12.5	63	<.001
Word order test 4	.85	.22	.79	.90	12.6	63	<.001
Word order test 5	.85	.22	.79	.90	12.7	63	<.001

Appendix G

Between-subject effects of repeated measures ANOVA for differences between linguists vs. non-linguistics, no languages to intermediate-level or above vs. at least one other language to intermediate-level or above, no previous experience of case-marked language vs. previous experience of case-marked language, degree vs no degree.

	Nouns			Verbs			Adjectives			Case markers			Word Order		
	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared	F	Sig	Partial Eta Squared
Linguist	0.35	.56	.014	0.38	.54	.016	1.83	.19	.071	0.49	.49	.009	1.57	.22	.062
Degree	0.44	.52	.018	0.30	.59	.013	0.86	.36	.034	0.05	.83	.001	0.02	.90	.001
Extra languages	0.41	.53	.017	0.51	.073	.13	0.95	.34	.038	0.40	.53	.008	2.23	.15	.085
Case-marked language	3.43	.069	.052	0.13	.72	.002	0.31	.58	.005	0.03	.86	.001	0.75	.39	.012