COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 36 (2012) 286–304

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2011.01200.x

Statistical Learning Is Related to Reading Ability in Children and Adults

Joanne Arciuli, Ian C. Simpson

Faculty of Health Sciences, University of Sydney

Received 18 June 2009; received in revised form 8 December 2010; accepted 14 February 2011

Abstract

There is little empirical evidence showing a direct link between a capacity for statistical learning (SL) and proficiency with natural language. Moreover, discussion of the role of SL in language acquisition has seldom focused on literacy development. Our study addressed these issues by investigating the relationship between SL and reading ability in typically developing children and healthy adults. We tested SL using visually presented stimuli within a triplet learning paradigm and examined reading ability by administering the Wide Range Achievement Test (WRAT-4; Wilkinson & Robertson, 2006). A total of 38 typically developing children (mean age of 9;5 years, range 6;4–12;5) and 37 healthy adults (mean age of 21 years, range 18–34) were assessed. In children, SL was significantly related to reading ability. Importantly, this relationship was independent of grade and also age. The adult data, too, revealed that SL was significantly related to reading ability. A regression analysis of the combined child and adult data revealed that SL accounted for a unique amount of variance in reading ability, after age and attention had been taken into consideration. For the first time, this study provides empirical evidence that a capacity for more effective SL is related to higher reading ability in the general population.

Keywords: SL; Statistical learning; Reading; Reading aloud; Reading development; Language acquisition

Cognitive scientists have long been interested in determining the relative contributions of endowment and learning during language acquisition. Interest in the role of learning has been reinvigorated by demonstrations of the brain's powerful capacity for learning associations; a critical component of which is sensitivity to the statistical probability that items are associated. This sensitivity is known as statistical learning (SL) and is thought to be a form of implicit learning (Perruchet & Pacton, 2006). While some recent research has provided

---

Correspondence should be sent to Joanne Arciuli, Faculty of Health Sciences, University of Sydney, PO Box 170, Lidcombe, Sydney, New South Wales 1825, Australia. E-mail: joanne.arciuli@sydney.edu.au

evidence that a capacity for SL is related to proficiency with natural language (e.g., Evans, Saffran, & Robe-Torres, 2009), there is strikingly little empirical data that directly support the hypothesis. Moreover, to date, discussion of the link between SL and language has focussed largely on oral rather than written language. To our knowledge, there has been no previous investigation of whether a capacity for SL is related to reading ability in typically developing children and healthy adults. Here, we set out to determine whether more effective SL is related to higher reading ability in the general population.

1. Assessment of SL

An individual's capacity for SL can be measured in a number of ways. For instance, it can be assessed by asking a participant to watch a continuous stream of evenly paced, individually presented items on a monitor and, later, surprising the participant with questions about what he or she saw during this *familiarization phase*. Typically, stimulus presentation times are consistent and are set at somewhere between 200 and 800 ms per item. One common design (e.g., see Brady & Oliva, 2008; Fiser & Aslin, 2002; Turk-Browne, Isola, Scholl, & Treat, 2008; Turk-Browne, Jungé, & Scholl, 2005) is to arbitrarily group the individual items together in triplets during the familiarization stream. After several minutes of watching the familiarization stream, the experimenter surprises the participant with a forced-choice task. On each trial of this *test phase* two triplets are presented, one after the other, with a short pause between them (the three component items of each triplet are displayed individually, in the same manner they were presented in the familiarization stream). Stimuli in one of these triplets occurred together during the familiarization phase, while stimuli in the other triplet, a foil, never appeared together during familiarization (i.e., the component items of the two internal pairs within the foil triplet exhibit the lowest possible transitional probability: TP = 0).[1] The participant is asked to judge which of the two triplets occurred during the familiarization stream. Chance performance equates to correct identification of 50% of the embedded triplets during the test phase. Studies that have used this paradigm usually report an overall group effect whereby embedded triplets are correctly identified over foil triplets at a rate significantly greater than chance. However, there is often a range of performance whereby some individuals appear to have learned more of the embedded triplets than other individuals.

It is important to emphasize that this SL effect emerges even though there is no advance warning of either embedded patterns or a subsequent test phase, and no reinforcement is provided at any time. Usually, participants have no conscious sense of familiarity during the test phase even though they are given an explicit forced-choice task (for further discussion see Arciuli & Simpson, 2011; Turk-Browne, Scholl, Chun, & Johnson, 2009). Indeed, similar levels of SL are obtained using other paradigms that do not include explicit questions during a separate test phase. This is because SL is thought to be a form of implicit learning; one that operates on a wide range of stimuli that contain probabilistic associations, be they spatial or temporal (e.g., Fiser & Aslin, 2001, 2002, 2005). Recent neuroscientific evidence has demonstrated that SL begins to operate almost as soon as an individual is exposed to stimuli containing regularities and has confirmed that SL proceeds without awareness or intention to learn (Turk-Browne et al., 2009).

A seminal study showed that 8-month-old infants can learn associations between syllables embedded in pseudo speech after only 2 min of exposure (Saffran, Aslin, & Newport, 1996). A study using visually presented shapes showed SL in infants as young as 2-month-olds (Kirkham, Slemmer, & Johnson, 2002). Thus, SL appears to be a capacity which begins to operate at a very young age. Studies of SL have usually examined the ability to detect associations among adjacent items; however, it has been shown that SL can also operate on certain nonadjacent patterns (e.g., Newport & Aslin, 2004). There is some evidence to suggest that SL may be a domain-general mechanism that operates similarly across linguistic and nonlinguistic stimuli (see Evans et al., 2009), although there is also a literature on modality effects (e.g., Conway & Christiansen, 2005). The issue of whether SL is domain-specific or domain-general continues to be discussed in the literature and additional empirical research is required.

2. The language acquisition debate

The hypothesized link between SL and natural language acquisition is highly topical (e.g., Chater & Manning, 2006; Kuhl, 2004). The claim that SL plays a central role in language acquisition stands in contrast with the persuasive notion that language, and grammar in particular, is too complex and the language learner's environment too impoverished for language acquisition to be subserved by a general learning mechanism. These views have led to the suggestion that language acquisition might be subserved by an innate language acquisition device which is pre-programed with information about linguistic universals. Patterson and Plaut (2009, p 46) cited Chomsky: ''…we take for granted that the organism does not learn to grow arms or to reach puberty… When we turn to the mind and its products, the situation is not qualitatively different from what we find in the case of the body'' (Chomsky, 1980, pp. 2–3). An interesting question is whether some aspects of language, such as grammatical structure, are less ''learnable'' than others—more heavily underpinned by innate knowledge (e.g., Nowak, Komarova, & Niyogi, 2002; Peña, Bonatti, Nespor, & Mehler, 2002; Rowland & Pine, 2000; Seidenberg, MacDonald, & Saffran, 2002). For some, the debate does not center on a strict distinction between innateness versus learning, but rather on the relative contributions of these via ''innately guided learning'' (Gould & Marler, 1987, p. 62; Yang, 2004).

There are three bodies of empirical evidence that inspire the claim that SL plays a role in the acquisition of natural language. First, it has been demonstrated there is rich statistical structure in both oral and written language. Much of this research has only been possible with the advent of powerful computing resources that enable large corpora of language data to be analyzed. For instance, in English, there are multiple probabilistic cues (phonological, orthographic, semantic, and syntactic) that help participants to distinguish whether a word is a noun or a verb (e.g., see Arciuli & Monaghan, 2009 for discussion). Second, the acquisition of artificial languages proceeds more effectively when they contain the types of statistical structure found in natural language (as reviewed by Gómez & Gerken, 2000). Note that Pelucchi et al. (2009) recently demonstrated that infant English learners' were able to track the TPs of syllables contained within fluent infant-directed Italian speech, providing evidence to support the claim that the SL mechanism can operate as effectively on natural language stimuli as it does on artificial language stimuli. Third, some studies have indicated that individuals with language-based impairments show deficiencies in implicit learning (Grunow, Spaulding, Gómez, & Plante, 2006; Tomblin, Mainela-Arnold, & Zhang, 2007).

Surprisingly, very few studies have examined whether a propensity for SL is directly related to proficiency with natural language. There are some data indicating that certain processing capacities present in infancy (thought to be underpinned by SL) are related to later language proficiency in typically developing children. For example, Tsao, Liu, and Kuhl (2004) found a relationship between 6-month-olds' ability to perceive phonemes and their language development at 13, 16, and 24 months. In a retrospective study, Newman, Ratner, Jusczyk, Jusczyk, and Dow (2006) showed a relationship between the ability of infants (up to 12 months of age) to segment speech into individual words and later proficiency with natural language at 24 months and also between 4 and 6 years. The relationship was not mediated by IQ.

A recent study has revealed a direct link between SL and proficiency with natural language. Evans et al. (2009) examined English-speaking children aged 6;5–14;4 years. They reported that children with specific language impairment performed more poorly than normal language controls on tasks of SL. Interestingly, children with language impairment did show evidence of SL, but they required a longer familiarization stream in order to do so. In addition, Evans et al. showed that SL performance was positively correlated with performance on standardized tests of vocabulary in both children with specific language impairment and in normal controls. SL performance was not correlated with IQ.

Indeed, it has long been thought that vocabulary acquisition involves implicit processes: ''It is both a theoretical and an empirical fact that not all vocabulary can or must be learned through formal instruction and that vocabulary words can also be learned through incidental and indirect ways'' (National Institute of Child Health and Human Development, 2000). For the first time, Evans et al. (2009) have provided empirical evidence of a relationship between an individual's capacity for SL and his or her vocabulary growth. Yu (2008) reported on a computational model of vocabulary growth that relied on sensitivity to statistical regularities. The model successfully simulated naturalistic language data in the form of child–caregiver interactions and displayed key phenomena associated with vocabulary development; namely, that word learning does not occur until late in the first year of life even though a capacity for SL is thought to be present at birth or even before, and that there appear to be changes in the rate of vocabulary growth during childhood.

In summary, there is reason to believe that a capacity for SL is directly related to proficiency with natural language, but there are little empirical data to support this hypothesis. The study by Evans et al. demonstrates such a link, but additional data are required in order to strengthen the evidence base. For instance, vocabulary is only one aspect of natural language proficiency. It is interesting to consider whether there may be a link between the capacity for SL and proficiency with written language.

3. Statistical learning and written language

The ability to read aloud and/or spell words correctly may reflect an ability to detect statistical regularities. There is a substantial body of research that supports this view of reading (e.g., Arciuli, Monaghan, & Ševa, 2010; Harm & Seidenberg, 1999, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; Ševa, Monaghan, & Arciuli, 2009). While there is less research on spelling in general (relative to reading), there is also a body of research that supports this view of spelling (e.g., Brown & Loosemore, 1994; Deacon, Conrad, & Pacton, 2008; Kessler, 2009; Treiman & Kessler, 2006).

Learning to read and spell in an alphabetic language means learning the correspondences between arbitrary visual symbols (i.e., letters) and the linguistically meaningful sounds of a language (i.e., phonemes). In a language such as English the mapping between letters and phonemes can be thought of as being probabilistic. For example, the letter ''c'' often maps on to the phoneme /k/. Indeed, children are often explicitly taught to think of the letter ''c'' as corresponding to the phoneme /k/ in words like ''cat'' when they begin formal instruction in reading. However, the letter ''c'' can be associated with different phonemes (as in ''circus''). Even if children are not taught explicitly about the letter ''c'' in these other words they are likely to learn these other mappings implicitly—as a result of increasing exposure to written language. Over time they are likely to detect that most words beginning with the letter ''c'' followed directly by the letter ''i'' have the phoneme /s/ as their initial phoneme. That is, they will become sensitive to contextual cues such as the co-occurrence of letters.

In fact, there are a myriad of different kinds of statistical regularities that someone learning to read and spell English will become sensitive to, but not necessarily consciously aware of. For example, he or she will learn that English words can be spelled with double letters at the end (''inn,'' ''add,'' ''cliff'') but are seldom spelled with double letters at the beginning. Importantly, not all of these statistical regularities concern adjacent units. A beginning reader in English will learn about the significance of the pairing of certain nonadjacent letters (e.g., ''a'' later followed by ''e'' as in ''hate'' rather than ''hat'').

Children might be taught explicitly about some of these regularities when they undertake formal instruction in reading; however, they are not taught about every correspondence and contextual cue in the English language. In fact, some statistical regularities can only be revealed through large corpus analyses, and researchers are still in the process of discovering them. Arciuli and colleagues have examined the probabilistic cues contained within English orthography that help readers to assign lexical stress appropriately when reading aloud words that have more than one syllable (e.g., ''ZEbra'' vs. ''girAFFE''). For example, around 90% of disyllabic English words that end with ''-an'' have first syllable stress, and over 80% of disyllabic English words that end with ''-uct'' have second syllable stress. Adults demonstrate sensitivity to these regularities by assigning first syllable stress and second syllable stress, respectively, when reading nonwords such as ''curban'' and ''feduct'' (Arciuli & Cupples, 2006; see Arciuli & Cupples, 2007; regarding probabilistic cues to stress assignment in word beginnings). Recently, additional corpus analyses, behavioral testing, and computational modeling have demonstrated that sensitivity to probabilistic orthographic cues to lexical stress follows a developmental trajectory in children aged 5–12 years (Arciuli et al., 2010). As children are exposed to an increasing volume and variety of written language in their age-appropriate reading materials, their sensitivity to these kinds of probabilities increases without having to draw children's attention to them explicitly through formal instruction.

Byrne and colleagues have investigated literacy development in twins over a number of years and in different countries. Byrne et al. (2008) have suggested that there is overlap in children's ability to read aloud, spell, and learn novel words and that ''a common factor, best described as a learning parameter, links these aspects of literacy'' (p. 18). Byrne and colleagues did not specify precisely what they meant by the term ''learning''; for instance, it is not clear what they think might be the relative contributions of explicit and implicit learning. It is interesting that they mention a capacity for novel word learning, which can be thought of as vocabulary growth. As mentioned, the recent study by Evans et al. (2009) demonstrated that higher SL is related to vocabulary growth in primary aged children. Gillon (2004) provides an illustration of how vocabulary, itself, can impact upon reading performance. A child who is reading a sentence such as ''The horse likes to eat gr___'' and struggles to properly decode the remainder of the final word can be assisted by vocabulary knowledge (concerning the materials that horses like to eat). If the word ''grass'' appears in the child's vocabulary, he or she is perhaps more likely to achieve a successful reading of the final word.

In summary, SL may be related to proficiency with written language in a number of ways, including (a) by enabling the detection of probabilistic correspondences between letters and phonemes and the detection of regularities that exist among letters; and (b) by boosting a range of linguistic resources (including vocabulary) which can, themselves, impact upon reading ability. It is important to remember that children vary greatly in their literacy ability and that, even after years of formal instruction in literacy, there remains substantial variation in reading and spelling ability among adults. Is it the case that individuals' capacity for SL is at least partly responsible for this variability?

4. Visual implicit learning and reading ability

Some previous studies have attempted to examine the link between visual implicit learning and reading ability. Rüsseler, Gerth, and Münte (2006) compared German adults who had reported a childhood diagnosis of developmental dyslexia (DD) with controls. They concluded that both their DD and control participants were able to implicitly learn regularities. Vicari et al. (2005) assessed Italian children and adolescents who had recently received a clinical diagnosis of DD and concluded that children with DD were impaired on implicit learning tasks when compared with controls. There are questions around the fact that DD was verified differently across these two studies. Moreover, some of the implicit learning tasks used in these previous studies might be considered sub-optimal. Rüsseler et al. used two tasks: a computer-presented serial response time task where the stimuli included a sequence of 12 items, and a manually presented artificial grammar learning paradigm using sequences of between four and seven items. Vicari et al. also used two, albeit different, tasks: a mirror drawing task which required multiple attempts and a computer-presented serial response time task which incorporated a nine-item sequence. Neither of these studies used the embedded triplet SL task. Neither study examined whether variability in the capacity for SL is directly related to variability in reading performance in typically developing children and healthy adults.

5. The current study

The hypothesized link between SL and natural language acquisition is highly topical and yet there is little empirical evidence that can be used to further the debate. Here, we wanted to determine whether typically developing children and healthy adults with a higher capacity for SL show higher levels of reading ability. A small number of previous studies that have examined the link between visual implicit learning and reading ability by comparing those with DD against a control group have produced mixed findings. To our knowledge, no previous study has utilized the embedded triplet test of SL to examine whether *variability* in SL is related to reading performance in the general population. An effective test of SL (one that elicits a significant group effect of learning) usually elicits individual variation in performance. We hypothesized that such variation is related to the ability to accurately decode written language given (a) the many statistical regularities that are present between letters and sounds, and among letters, in English orthography; and (b) recent evidence that other linguistic abilities that are known to impact upon reading, such as vocabulary knowledge, are related to SL. By including both children and adults, we hoped to shed light on whether any relationship between SL and reading ability that might be present during childhood can also be observed in adults.

6. Experiment 1

In Experiment 1 we examined the relationship between SL and proficiency with written language in typically developing primary aged children. We used a standardized test of language proficiency (as in Evans et al., 2009). In line with relevant models of reading we focused on the reading aloud of single words (e.g., Arciuli et al., 2010; Harm & Seidenberg, 1999, 2004; Plaut et al., 1996; Ševa et al., 2009). As in previous studies that have examined the link between implicit learning and reading ability, we chose a learning task with visually presented stimuli (Rüsseler et al., 2006; Vicari et al., 2005). Unlike any previous studies, we used the embedded triplet paradigm to assess SL.

6.1. Method

6.1.1. Participants

A total of 42 typically developing primary aged children from Kindergarten to Year 6 took part (mean age = 9;1; range 5;10–12;5, 20 females). Children were recruited from schools in Sydney, Australia. Those who returned consent forms and met our inclusion criteria were tested. Inclusion criteria: native monolingual speakers of Australian English who were judged to be typically developing by their classroom teacher and had no known learning, language, hearing or speech impairments.

6.1.2. Test materials

Materials consisted of the reading subtest of the Wide Range Achievement Test, fourth edition (WRAT-4; Wilkinson & Robertson, 2006) and a visual test of SL. The WRAT-4 reading subtest measures participants' ability to read aloud individually presented orthographic strings. As the WRAT-4 is designed for use with both young children and competent adult speakers of English, the test items range from very simple (i.e., single letters of the alphabet comprise the first 15 stimuli of the test) to very difficult (uncommon words with irregular pronunciations). Healthy children and adults deemed to be at a sufficiently advanced level are not required to read aloud the single letters, and instead start at the first whole word stimulus. In this instance, they automatically start with a score of 15 points, which is the score they would have achieved had they read aloud correctly all of the single letter stimuli. There are 55 whole word stimuli and participants score one point for each correct response, with the maximum score being 70 (testing continues until subjects make 10 consecutive errors, or until they read all test items aloud).

6.1.2.1. SL test: We used the same SL test described by Arciuli and Simpson (2011) which was modeled on previously published tasks (e.g., Brady & Oliva, 2008; Fiser & Aslin, 2002; Saffran et al., 1996; Turk-Browne et al., 2008). The SL test comprised a familiarization phase in which participants were engaged in a cover task, followed by a surprise forced-choice test phase. Eighteen cartoon-like figures were chosen as stimuli for the experiment. The figures were sourced from the website http://www.clipartconnection.com/en/.[2] Each figure could be loosely described as an ''alien.'' None resembled common real-word animals or people or popular cartoon characters. Six of the aliens were used only for instructional purposes and during practice trials. The remaining 12 aliens were divided into four groups of three (four *base triplets*), and these are referred to as *ABC*, *DEF*, *GHI*, and *JKL* (see Appendix).

The familiarization phase consisted of a continuous stream of aliens, shown one at a time, in the center of the display against a white background (each visible for 400 ms with an interstimulus interval of 200 ms). Aliens were always displayed as part of a base triplet with all of the aliens in one triplet appearing before any aliens from another triplet were displayed. Ordering within each triplet was always maintained such that, for triplet *ABC*, alien A was always displayed prior to alien B, and alien B was always displayed prior to alien C. Each of the four base triplets were selected for inclusion in the familiarization stream on 24 occasions each (giving a total of 96 triplets). For six of these 24 instances, one of the aliens was presented twice in a row. This repetition was included in order to provide a cover task that would ensure participants paid attention during the familiarization phase, with participants required to press a button whenever they saw a repeated alien. These repetitions were counterbalanced among the three aliens within the triplet. So, for example, for base triplet *ABC* there were two occurrences of *AABC*, two occurrences of *ABBC*, two occurrences of *ABCC,* and 18 occurrences of *ABC*. The familiarization stream thus consisted of 312 individual presentations of the aliens, with each of the 12 types of aliens appearing 26 times each. The order of the triplets within the familiarization stream was randomized; although, as per Turk-Browne et al. (2005), this was constrained in two ways. Using T_1 and T_2 as hypothetical triplets, no repeated triplets were allowed (e.g., $\ldots T_1T_1\ldots$) and no repeated pairs of triplets were allowed (e.g., $\ldots T_1T_2T_1T_2\ldots$). These constraints did not take into consideration any repeated aliens and so the sequence *ABCCABC* was still invalid as this was considered to be a repetition of the triplet *ABC*. Due to these constraints, four different pre-randomized familiarization lists were created with each participant viewing one of these four lists. The familiarization phase lasted 3 min and 7 s.

For the test phase four new triplets were created, with each containing one alien from each of three base triplets. Due to the ordering constraints, these new triplets never appeared in the familiarization stream. These four *impossible triplets* are referred to as *AEI*, *DHL*, *GKC*, and *JBF*. Constructing the impossible triplets in this manner meant that the TPs of the internal pairs were zero. This contrasts with the TPs of the base triplets which were close to 1.0.[3] For each test trial, one actual base triplet was displayed along with one impossible triplet. The aliens in each triplet were presented one at a time using the same presentation time and interstimulus interval used in the familiarization phase with a 1,000 ms gap separating the base triplet from the impossible triplet. After all six aliens had been presented a new screen appeared which prompted participants to identify which of the two triplets had appeared previously in the familiarization phase, with no time constraints imposed. Each base triplet was presented with each impossible triplet on four separate occasions with the presentation order counterbalanced. Across the 64 test trials each base triplet was seen 16 times and each impossible triplet was seen 16 times. This insured that if any SL took place during the test phase itself, the opportunities to learn were equal for both types of triplet. Each participant received a different random order for the test trials. Item presentation and data collection in both phases were controlled using E-prime (Schneider, Eschman, & Zuccolotto, 2002).

6.1.3. Procedure

First we administered the reading subtest of the WRAT-4, then the SL test.

6.1.3.1. Familiarization phase of the SL test: Before beginning the familiarization phase the experimenter described the initial phase of the experiment using matching visual information that appeared on the screen. As a cover task, participants were told that they would see a group of aliens lining up to enter a spaceship and were asked to monitor the line looking for two aliens in a row that looked identical. Participants were asked to press the space bar whenever they saw two identical aliens in a row. During the instruction phase, three short practice trials were administered. Participants could not proceed to the familiarization phase until all of the practice trials had been successfully executed. No feedback on performance was given during the actual familiarization stream. As mentioned, the aliens used during the instructions and practice trials were not the same aliens used in the familiarization stream or test phase.

6.1.3.2. Test phase of the SL test: Immediately following the familiarization phase participants were informed that the aliens often lined up together in groups of three, and that they would be assessed to determine whether they had noticed this. Participants were told that for each trial they would see one group of aliens that always lined up together and another group which never lined up together. Participants were to indicate verbally which group they thought had previously lined up together. Two practice trials were presented, and the entire set of instructions could be repeated if the participant was unclear what was required. No feedback was given once the actual test trials commenced.

6.2. Results and discussion

Data from the familiarization phase were inspected to determine the number of repeated aliens successfully identified by each participant. Participants who failed to identify at least 12 of the 24 repeated aliens were excluded, on the grounds that they may not have been attending to the familiarization stream. As a result of this screening four participants were excluded.

The percentage of triplets correctly identified during the test phase ranged from 40.6 to 82.8%, with a mean of 56.3% of base triplets successfully identified (*SD* 11.02%). Chance performance represents correct identification of 50% of the base triplets. A one-sample t-test demonstrated that, overall, the children's performance was significantly better than chance performance, $t(37) = 3.544$, $p < .05$. This indicates that the children were able to learn the statistical regularities that were present in the familiarization stream. Scores on the reading subtest of the WRAT-4 ranged from 24 correct to 56 correct (out of a total of 70), with a mean of 38.1 (*SD* 7.76). Not surprisingly, scores on the WRAT-4 were correlated with age (expressed in months), with older children performing better ($r = .863$, $p < .0005$). While SL performance and age were not significantly correlated, there was a trend for higher SL scores to be associated with older ages ($r = .294$, $p = .073$). Due to the relationship between age and performance on the WRAT-4, and the marginally significant association between SL performance and age, we conducted a correlational analysis between performance on the WRAT-4 and performance on the SL test while partialling out age. Results showed a significant correlation between the WRAT-4 and SL, $r = .327$, $p < .05$. We conducted a second correlational analysis, this time partialling out grade rather than age. Once again there was a significant correlation between the WRAT-4 and SL, $r = .364$, $p < .05$. These results suggest that better performance on the SL task was associated with better performance on the reading aloud task. Importantly, this result does not simply reflect the fact that older children might be expected to perform better on both tasks.

7. Experiment 2

In Experiment 2 we examined the relationship between SL and reading ability in healthy adults using the same tests and procedures as in Experiment 1.

7.1. Method

A total of 37 undergraduates recruited from the University of Sydney took part (mean age 21;0, range 18–34, with 35 females), receiving course credit in return for their participation. All were healthy monolingual speakers of Australian English. The test materials and procedure used for the adults were identical to those used for the children, with one exception. During the test phase the children verbally indicated which triplet they thought had appeared previously in the familiarization phase. In contrast, adult participants were required to press one of two buttons on a response box, corresponding to their selection.

7.2. Results and discussion

Data from the familiarization phase were again inspected to determine the number of repeated aliens successfully identified by each participant. Scores ranged from 75% to 100% of repetitions successfully identified. For this reason, all adult participants were retained for analysis as these scores suggested that participants had been attending to the familiarization stream. Results on the SL test ranged from 35.9% to 92.2%, with a mean of 61.4% of base triplets successfully identified (*SD* 13.09). Overall, this was significantly better than chance performance, $t(36) = 5.278$, $p < .0005$.[4]

Scores on the reading subtest of the WRAT-4 ranged from 54 correct to 66 correct (out of a total of 70), with a mean of 60.2 (*SD* 3.13). Age was not correlated with either SL performance ($r = .045$, $p = .793$) or with WRAT-4 performance ($r = .203$, $p = .229$); however, in line with procedures undertaken in Experiment 1, we conducted a correlational analysis to examine the relationship between performance on the WRAT-4 and performance on the SL test, partialling out age. Once again there was a significant correlation between the WRAT-4 and SL, $r = .338$, $p < .05$. As with the child data, higher scores on the SL task were associated with better performance in the WRAT-4 reading aloud subtask, and this relationship was independent of age.

We also analyzed the combined child and adult data from Experiments 1 and 2 as a single group of 75 participants. Overall, results on the SL test ranged from 35.9% to 92.2% of base triplets correctly identified, with a mean of 58.8%. This was significantly better than chance performance, $t(74) = 6.222$, $p < .0005$. As performance on the SL task is a measure of the amount of learning that took place during the familiarization phase, one possibility is that scores on the SL task reflect how much attention participants paid while they viewed the familiarization stream. More generally, a positive association between SL performance and WRAT-4 scores may indicate that some participants were more engaged, and hence scored well on both tasks, while others were less engaged and subsequently scored relatively lower on both tasks. One way to examine this possibility is to take into consideration how much attention subjects paid while they were attending to the familiarization stream. As the cover tasks required participants to indicate whenever they saw a repeated alien, a crude measure of attention is available in the form of how many of the repeated item they were successfully able to identify. We performed a multiple linear regression using three independent variables: age in years, performance in the familiarization phase, and the amount of SL demonstrated, in order to predict performance on the WRAT-4 reading task.

From Table 1 it can be seen that Age uniquely accounts for just over half of the variance in WRAT-4 scores. The amount of attention paid by participants in the familiarization phase (as measured by the number of repeated aliens successfully identified) was correlated with WRAT-4 scores, and in the regression model it accounts for a significant amount of unique variance in WRAT-4 scores. Importantly, although the effect size is modest, the amount of SL shown by participants is a highly significant predictor of WRAT-4 scores after Age and Familiarization Performance have been taken into account. For every 1% increase in SL, WRAT-4 performance goes up .138 points. Put another way, a 7.2% increase in SL performance would result in a 1 point increase in WRAT-4 performance (with Age and Familiarization Performance held constant).

8. General discussion

Strong evidence is emerging that SL, a form of implicit learning, might contribute to almost every mental activity. Indeed, research on SL is ‘‘growing exponentially’’ (Perruchet & Pacton, 2006, p. 233). The hypothesized link between SL and proficiency with natural language is particularly significant, as language is one of the most complex of all mental activities. The role of SL during language acquisition has been hotly debated in the literature (e.g., see Newport & Aslin, 2004; Peña et al., 2002; Saffran et al., 2008; Seidenberg, 1994, 1997; Seidenberg et al., 2002; Yang, 2004). Yet, empirical evidence required to further develop the arguments in this debate has been lacking.

In the current study, we employed a well-established triplet learning paradigm, which incorporates visually presented sequential stimuli embedded within a familiarization phase followed by a surprise test phase, in order to assess SL. The SL task we used was modeled on

Table 1
Summary of multiple regression analysis for variables predicting subject performance (combined adults and children) on the WRAT-4

Predictors	*r*	r^2%	Raw B	Standard β	*t*-Value	*p*-Value	% Fit Last
Full/final model							
Age (years)	.910	82.8**	1.557	.814	16.007	<.0005	50.6
Familiarization performance	.510	26.0**	13.623	.146	2.947	.004	1.7
VSL performance	.339	11.5*	.138	.134	2.916	.005	1.7

Note. $^*p \leq .005$. $^{**}p \leq .0005$.
Full/final model R^2 = 86.0%, MSE = 23.293, error *df* = 71.

numerous previous studies (e.g., see Brady & Oliva, 2008; Fiser & Aslin, 2002; Turk-Browne et al., 2005, 2008) and was recently reported by Arciuli and Simpson (2011). In Experiments 1 and 2 we obtained significant effects of SL. Overall, participants were able to correctly identify more that 50% of the embedded triplets after viewing a familiarization stream for only 3 min and 7 s. This effect was obtained even though participants were given no advance warning that they were expected to learn anything during the familiarization stream, much less pay attention to triplets. Moreover, the stimuli were unfamiliar ''alien'' characters, the presentation rate was fast, and participants were engaged in a cover task during familiarization. Although we did not systematically collect data on the question of conscious awareness, many participants voluntarily reported that they had no conscious sense of familiarity concerning triplets presented during the test phase. In a follow-up of a subgroup of adult participants that took part in Experiment 2, we found that neither explicit phonological working memory nor intelligence was related to performance on the SL task (see note 4).

As expected, while there was a significant overall effect of learning for children and for adults, there was substantial variation in SL performance in both groups of participants. Similarly, participants exhibited a range of reading performance as measured by the WRAT-4. Separate correlational analyses of the child and the adult data revealed a significant relationship between SL and reading ability, after controlling for age (and, in the case of children, grade). A regression analysis of the combined child and adult data showed that SL was a significant predictor of reading ability after age and attention were taken into consideration. Participants who learned more of the TPs embedded in the familiarization stream of the SL task demonstrated higher reading ability. These findings are in line with a reported link between higher SL and greater vocabulary knowledge (Evans et al., 2009) and evidence that impaired readers have difficulty with visual implicit learning (Vicari et al., 2005). Our data suggest that the relationship between SL and reading ability appears to linger beyond childhood.

Clearly, there are a number of factors at play during the highly complex task of reading. We posit that a capacity for SL may be one of these factors. Our data do not enable us to disentangle the relative contributions of the direct versus indirect mechanisms that may subserve the relationship between SL and reading ability. A propensity for SL may assist reading ability directly by enabling the detection of statistical regularities that exist between letters and phonemes and among letters. There are well-established models of reading/spelling and accompanying behavioral data that emphasize the importance of the probabilistic relationship between letters and phonemes, and among letters, in English orthography (e.g., Arciuli & Cupples, 2006, 2007; Brown & Loosemore, 1994; Deacon et al., 2008; Harm & Seidenberg, 1999, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989; Ševa et al., 2009; Treiman & Kessler, 2006). In accordance with this approach, ''Learning is statistical in that it goes beyond all-or-none patterns to encompass probabilistic patterns. Such learning is implicit: it does not require an explicit understanding…'' (Treiman & Kessler, 2006, p. 643). In addition, a propensity for SL may assist reading ability indirectly by boosting a range of linguistic resources such as vocabulary which may, themselves, impact upon reading. Evans et al. (2009) demonstrated that higher SL is related to increased vocabulary growth in primary aged children and vocabulary growth has been linked to reading ability (see Gillon, 2004, for review).

A strong test of the hypothesis that SL is related to reading ability involves selection of (a) an SL task that has no particular pre-determined relationship with reading processes; and (b) use of a standardized test of reading ability that was not designed with the probabilistic link between letters and sounds, and among letters, in mind. Indeed, this is the course we undertook in the current study. Our results confirmed that individuals with a higher capacity for SL are more accurate when reading aloud. There are a number of additional investigations that might provide converging evidence. For example, one could design a reading test that contains particular probabilities in letter-phoneme correspondences and among letters, and then seek to determine the relationship between performance on this kind of reading task with performance on an SL task. We expect that the use of such a purpose-designed reading test would result in larger effect sizes than we have reported here. It would also be worthwhile to examine further the hypothesis that SL is a domain-general learning capacity, and that it is this general mechanism that is related to reading ability, by running the same set of experiments reported here using an auditory version of the embedded triplet SL task (i.e., using tones rather than aliens). One could also assess spatial rather than temporal SL.

Another avenue for future research is the investigation of variables that could, potentially, mediate the relationship between SL and reading ability. For example, one might wonder whether an association between SL and reading reflects the amount of attention paid by participants during the SL task (effects of attention during SL have been reported in studies such as Baker, Olson, & Behrmann, 2004; Toro, Sinnett, & Soto-Faraco, 2005; Turk-Browne et al., 2005). We have reasons to believe that this is not the case in the current study. As described in our Results section in Experiment 1, we excluded the data from four children who performed poorly during the familiarization phase of our SL task (Footnote 4). For instance, one child aged 7 identified only 10 of the 24 repeated aliens and registered 11 false positives. In addition, our regression analysis indicated that SL was a significant predictor of reading ability even after performance during familiarization, a proxy measure for attention, was taken into consideration. In a follow up of some of the adults we tested in Experiment 2, neither phonological working memory nor intelligence was related to performance on the SL task. It is important to remember that Evans et al. (2009) found that IQ did not mediate the relationship between SL and vocabulary. Similarly, Newman et al. (2006) found that the relationship between infants' ability to segment speech into individual words (thought to rely on SL) and later proficiency with natural language was not mediated by IQ. Admittedly, our data do not allow us to claim definitively that attention, phonological working memory, and IQ have no role to play in mediating the relationship between SL and reading ability. Additional research will help to clarify any possible role for mediating variables.

Finally, in order to determine whether there might be a causal relationship between SL and proficiency with natural language, it would be valuable to conduct a longitudinal study where SL is measured in infancy, prior to the emergence of oral or written language. Following this, both SL and proficiency with oral and written language could be tracked over a number of years. A study of adults has indicated that increasing a participant's exposure to particular linguistic constructions, such as those in some relative clauses, can make them easier to learn (Wells, Christiansen, Race, Acheson, & MacDonald, 2009); thus, potentially, language impairments associated with inefficient SL might be remediated by focusing on the nature and volume of the input to which learners are exposed.

This study contributes to the highly topical debate concerning the link between SL and natural language acquisition. For the first time, our results demonstrate that variability in the capacity for SL is related to reading ability in the general population. Along with the recent study of the relationship between SL and vocabulary reported by Evans et al., our findings encourage further investigation regarding the role of implicit learning during language acquisition.

Notes

1. The transitional probability, also known as the conditional probability, $P(B|A)$ is the probability of event B occurring after event A. It is possible to alter the transitional probabilities of the stimuli in SL tasks (i.e., using $0 < \text{TPs} < 1$). Recent research suggests that similar levels of SL are obtained when transitional probabilities are manipulated. Pelucchi, Hay, and Saffran (2009) assessed 8-month-olds and found similar levels of SL when they compared high probability (TP = 1) and low probability (TP = 0) in Experiment 2 versus high probability (TP = 1) and low probability (TP = 0.33) in Experiment 3. Evans et al. (2009) manipulated the strength of transitional probabilities comparing target stimuli with a range of transitional probabilities (TP range of 0.37–1) to foil stimuli with TP = 0. All were learned by typically developing primary school-aged children.
2. Access to this website is via subscription. Images can be legally downloaded from this website during the subscription period and can be used in print or electronic form indefinitely.
3. The transitional probabilities of the internal pairs in the base triplets were 0.92. A number of previous studies have used the maximum possible contrast between target triplet and foil triplet transitional probabilities, that is, 1.0 and 0.0, respectively, including Brady and Oliva (2008), Experiment 1; Fiser and Aslin (2002); Saffran et al. (1996), Experiment 1; and Turk-Browne et al. (2008), Experiment 1. While the present study was based on these previous studies, the fact that each of the three items in each triplet was occasionally repeated in the familiarization stream in the present study (a feature not present in the previous studies), resulted in the internal TPs of the base triplets dropping from 1.0 to .92.
4. At the request of reviewers we examined whether SL performance might reflect phonological working memory processes and/or IQ. We were not able to recall any child participants, but we were able to recall 11 of our adult participants from Experiment 2 (range 18–34, 10 females). SL scores for this subgroup ranged from 46.9% to 73.4% (mean 60.8, *SD* 8.2). Overall, there was a significant degree of learning, $t(10) = 4.355$, $p < .005$. Recall testing materials were the combined forward and backward digits subtasks of the Weschler Adult Intelligence Scales (WAIS-III; Weschler, 1997), both the memory for digits and the nonword repetition subtasks of the Comprehensive Test

of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999), and the Test of Nonverbal Intelligence third edition (TONI-3; Brown, Sherbenou, & Johnsen, 1997). To determine whether performance on the SL task was related to phonological working memory and/or IQ, we performed correlational analyses looking at the relationship between SL and each of these four measures. None of the correlations were significant (WAIS-III combined scores and SL: $r = -.088$, $p = .798$; nonword repetition and SL: $r = .313$, $p = .348$; memory for digits and SL: $r = .209$, $p = .538$; TONI-3 and SL: $r = .095$, $p = .781$). If participants were somehow able to verbalize all of the stimuli in our SL task (which we think very unlikely given the unfamiliar nature of the stimuli, the fast stimulus presentation time, and the requirements of the cover task), this did not appear to assist performance, at least not in this subgroup. As in the study by Evans et al. (2009), SL performance was not a reflection of IQ, at least not in this subgroup.

Acknowledgement

This research was supported by ARC grant LX0775703 awarded to the first author.

References

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *Quarterly Journal of Experimental Psychology*, *59*(5), 920–948.

Arciuli, J., & Cupples, L. (2007). Would you rather 'embert a cudsert' or 'cudsert an embert'? How spelling patterns at the beginning of English disyllables can cue grammatical category. In A. C. Schalley & D. Khlentzos (Eds.), *Mental states: Language and cognitive structure* (Vol. 2, pp. 213–237). Amsterdam: John Benjamins.

Arciuli, J., & Monaghan, P. (2009). Probabilistic cues to grammatical category in English orthography and their influence during reading. *Scientific Studies of Reading*, *13*(1), 73–93.

Arciuli, J., Monaghan, P., & Ševa, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioural and computational investigations. *Journal of Memory and Language*, *63*, 180–196.

Arciuli, J., & Simpson, I. C. (2011). Visual statistical learning in typically developing children: The role of age, gender and speed of presentation. *Developmental Science*, *14*, 464–473.

Baker, C., Olson, C., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*(7), 460–466.

Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes. *Psychological Science*, *19*(7), 678–685.

Brown, G. D. A., & Loosemore, R. P. W. (1994). Computational approaches to normal and impaired spelling. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and application* (pp. 319–335). Chichester, England: Wiley.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Test of nonverbal intelligence, third edition (TONI-3): A language-free measure of cognitive ability*. Austin, TX: PRO-ED.

Byrne, B., Coventry, W. L., Olson, R. K., Hulslander, J., Wadsworth, S., DeFries, J. C., Corley, R., & Willcutt, E. G. (2008). A behaviour-genetic analysis of orthographic learning, spelling and decoding. *Journal of Research in Reading*, *31*(1), 8–21.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Science*, *10*(7), 335–344.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*, 1–15.

Conway, C., & Christiansen, M. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24–39.

Deacon, S. H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on children's learning about graphotactic and morphological regularities in spelling. *Canadian Psychology*, *49*(2), 118–124.

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.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504.

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), 458–467.

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*(4), 521–537.

Gillon, G. T. (2004). *Phonological awareness: From research to practice*. New York: The Guilford Press.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science*, *4*(5), 178–186.

Gould, J. L., & Marler, P. (1987). Learning by instinct. *Scientific American*, *255*(1), 74–85.

Grunow, H., Spaulding, T. J., Gómez, R. L., & Plante, E. (2006). The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders*, *39*(2), 158–170.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662–720.

Kessler, B. (2009). Statistical learning of conditional orthographic correspondences. *Writing Systems Research*, *1*(1), 19–34.

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.

Kuhl, P. K. (2004). Early language acquisition: Cracking the code. *Nature Reviews Neuroscience*, *5*, 831–843.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: Author.

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

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, *417*, 611–617.

Patterson, K., & Plaut, D. C. (2009). ''Shallow Draughts Intoxicate the Brain'': Lessons from cognitive science for cognitive neuropsychology. *Topics in Cognitive Science*, *1*, 39–58.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674–685.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Science*, *10*(5), 233–238.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Rowland, C. F., & Pine, J. M. (2000). Subject–auxiliary inversion errors and wh-question acquisition: ''What children do know?''. *Journal of Child Language*, *27*, 157–181.

Rüsseler, J., Gerth, I., & Münte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 808–827.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Hauser, M. D., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*(2), 479–500.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools, Inc.

Seidenberg, M. S. (1994). Language and connectionism: The developing interface. *Cognition*, *50*, 385–401.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*(5306), 1599–1603.

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does grammar start where statistics stop? *Science*, *298*(5593), 553–554.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.

Ševa, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics*, *22*(3), 237–249.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*, *3*(4), 269–293.

Toro, J., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), B25–B34.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, *98*(3), 642–652.

Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084.

Turk-Browne, N. B., Isola, P. J., Scholl, B. J., & Treat, T. A. (2008). Multidimensional statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 399–407.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945.

Vicari, S., Finzi, A., Menghini, D., Marotta, L., Baldi, S., & Petrosini, L. (2005). Do children with developmental dyslexia have an implicit learning deficit? *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(10), 1392–1397.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing – CTOPP*. Austin, TX: PRO-ED.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271.

Weschler, D. (1997). *Weschler Adult Intelligence Scales – Third edition (WAIS-III)*. San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test – Fourth edition: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Science*, *8*(10), 451–456.

Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning. *Language Learning and Development*, *4*(1), 32–62.

Appendix: Alien stimuli


Image A


Image B


Image C


Image D


Image E

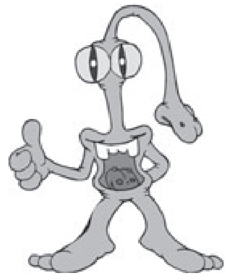
Image F


Image G


Image H


Image I

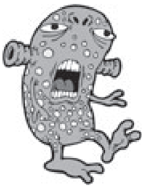
Image J

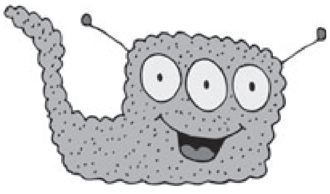
Image K


Image L