Implicit Statistical Learning Is Directly Associated With the Acquisition of Syntax

Evan Kidd
The University of Manchester

This article reports on an individual differences study that investigated the role of implicit statistical learning in the acquisition of syntax in children. One hundred children ages 4 years 5 months through 6 years 11 months completed a test of implicit statistical learning, a test of explicit declarative learning, and standardized tests of verbal and nonverbal ability. They also completed a syntactic priming task, which provided a dynamic index of children's facility to detect and respond to changes in the input frequency of linguistic structure. The results showed that implicit statistical learning ability was directly associated with the long-term maintenance of the primed structure. The results constitute the first empirical demonstration of a direct association between implicit statistical learning and syntactic acquisition in children.

Keywords: implicit statistical learning, language acquisition, syntax, priming

Work on infant speech perception has revealed an early domain-general ability to detect statistical regularities (e.g., Saffran, 2003; Saffran, Aslin, & Newport, 1996). This facility for implicit statistical learning is assumed by many theories of language acquisition to lay the foundations for the acquisition of syntax (e.g., Bannard, Lieven, & Tomasello, 2009; Bates & MacWhinney, 1982; Chang, Dell, & Bock, 2006; Chang, Lieven, & Tomasello, 2008; Kuhl, 2004; Yang, 2004).[1] Although there have been many empirical demonstrations showing that children can detect statistical regularities in both the visual and auditory modalities (see Gómez & Gerken, 1999, 2000; Romberg & Saffran, 2010; Saffran, Werker, & Werner, 2006), there has been no empirical demonstration showing that this ability is implicated in the acquisition of syntax in natural languages. The current article reports on an individual differences study that directly tested the role of implicit statistical learning in the acquisition of syntax.

Statistics in Language Acquisition

In natural languages, both word tokens and syntactic structures differ in their frequency of occurrence. It is uncontroversial to argue that lexical frequency affects vocabulary acquisition; however, the role of statistics and statistical learning in the acquisition of grammar has been more controversial. This controversy stems from the *argument from the poverty of stimulus* (Chomsky, 1980; Gold, 1967), which argues that the input is too impoverished for children to acquire a grammar through the kind of induction assumed to drive statistical learning. The nature of this debate has remained largely unchanged since the beginning of the modern study of language acquisition (e.g., Lidz, Waxman, & Freedman, 2003; for a reply, see Akhtar, Callanan, Pullum, & Scholz, 2004; MacWhinney, 2004; and for commentaries, see Pullum & Scholz, 2002). Although this debate continues, an increasing number of published studies have shown frequency effects in the acquisition of syntax.

Data from naturalistic studies have shown that input frequency is a major determinant of the acquisition of a variety of syntactic structures (e.g., Diessel, 2004; Goldberg, Casenhiser, & Sethuraman, 2004, 2005; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowland, 2007; Rowland & Pine, 2000; Theakston, Lieven, Pine, & Rowland, 2002, 2004). Other studies have shown that frequency estimates from corpora predict children's linguistic behavior in experiments. For instance, Matthews, Lieven, Theakston, and Tomasello (2005) showed that verb frequency predicts children's tendency to accept or correct ungrammatical word orders. Bannard and Matthews (2008) reported that 2- and 3-year-old children were better able to repeat high-frequency four-word sequences (e.g., *sit in your chair*) than four-word sequences that differed only in their final word (e.g., *sit in your truck*), a change that significantly altered the frequency of the string and, hence, the predictability of the final word in the sequence (see also Matthews & Bannard, 2010).

These frequency effects scale up to more complex constructions. Huttenlocher, Vasilyeva, Cymerman, and Levine (2002) reported on a longitudinal study that showed that the complexity of input that kindergarten children receive from both their parents and their teachers predicted their subsequent syntactic knowledge (see also Huttenlocher et al., 2010). Kidd, Lieven, and Tomasello (2006,

[1] Following Conway et al. (2010), the terms *statistical* and *implicit* learning are used interchangeably throughout this article, but most often the compound *implicit statistical learning* is used.

2010) have shown that 3- to 6-year-old English-speaking children's knowledge of finite sentential complement clauses (e.g., *I think she's wearing a lovely dress*) is predicted by the frequency with which the main verb occurs in a complement construction relative to its use in other syntactic environments. Similarly, Kidd, Brandt, Lieven, and Tomasello (2007) and Brandt, Kidd, Lieven, and Tomasello (2009) have shown that distributional frequency information predicts English- and German-speaking children's production and comprehension of relative clauses. Finally, Casenhiser and Goldberg (2005) showed that skewed frequency distributions, where one exemplar verb occurred many more times than did other verbs, facilitated 5-year-old children's learning of a novel construction (for a longer review, see Lieven, 2010).

Therefore, there is a wealth of evidence to show that (a) infants and children are capable of implicit statistical learning in multiple modalities and (b) children are sensitive to frequency distributions in natural languages. That is to say, it is clear that children are adept at extracting statistical regularities from their input, and frequency effects in children's linguistic behavior suggest that this skill is important for language acquisition. However, as yet there has been no direct empirical demonstration that implicit statistical learning is implicated in the acquisition of grammar, and little is known about the mechanism and the processes that track frequency distributions and, as such, might support the language-learning process. I review some relevant research next.

I begin with the working hypothesis that language learning involves both the implicit and explicit learning processes. Using broad brushstrokes, these processes map onto putatively syntactic and lexical processes, respectively, although the boundary between syntax and the lexicon is almost certainly fuzzy (Bates & Goodman, 1997). Current work on language acquisition that assumes the dual action of these processes has concentrated on the role of implicit statistical and explicit (or declarative) learning in typically developing children and children with specific language impairment (SLI). Implicit statistical learning is defined broadly here as the largely or wholly unconscious process of inducing structure from input following exposure to repeated exemplars (e.g., Perruchet & Pacton, 2006).

Tomblin, Mainela-Arnold, and Zhang (2007) showed that adolescents with a diagnosis of SLI were slower to learn an implicit pattern than a control group without language impairment. They used a Serial Reaction Time (SRT) task to test implicit learning. The task is a test of visual statistical learning: A visual stimulus occurs in a repeating pattern in one of four spatial locations on a computer screen, and the participant's task is to press a button corresponding to the location of the stimulus as quickly as possible. Implicit learning is observed if response times (RTs) decrease across multiple presentations of the pattern. Because the pattern is typically too long a sequence to memorize explicitly (e.g., 10), any learning is argued to be implicit (but see Jamieson & Mewhort, 2009). Tomblin et al. interpreted their findings to suggest that the grammatical deficits observed in SLI are due to a deficit in implicit learning. These findings were supported by research reported by J. L. Evans, Saffran, and Robe-Torres (2009), who showed that younger children with SLI performed poorly compared with controls on statistical learning tests modeled on those used with infants. Furthermore, these authors reported a significant correlation between statistical learning and vocabulary in matched typically developing controls following 21-min of exposure to the

artificial language, suggesting that performance on statistical learning tasks are sensitive to individual differences in developmental populations.[2] Subsequent research has suggested that children with SLI also have compromised explicit learning abilities. Lum, Gelgic, and Conti-Ramsden (2010) tested children with SLI and matched controls on a version of the SRT task but also tested them on a measure of explicit (or declarative) learning. They found that children with SLI were impaired in both tasks. This raises the possibility that the grammatical deficits seen in SLI are the result of impairments in both implicit and explicit learning.

The only published research to have directly investigated the relationship between implicit and explicit learning and natural language acquisition is a study by Kidd and Kirjavainen (2011), who investigated the acquisition of Finnish past tense morphology in Finnish-speaking children ages 4 years 0 months to 6 years 6 months.[3] Past tense morphology has served as a test case in the language sciences in the debate regarding the extent to which language is best characterized as rule driven (e.g., Pinker, 1999) or, following connectionist approaches, the product of associative learning (e.g., Bybee, 1995; Rumelhart & McClelland, 1986). Compared with English, Finnish morphology is very complex: Any one lexeme can have in excess of 100 potential surface forms, raising the possibility that only a powerful implicit learning mechanism could ever learn the language. In spite of this fact, Kidd and Kirjavainen found that implicit statistical learning, as measured by performance on an SRT task, did not predict the children's performance on a test of past tense elicitation of both real and novel verbs. Instead, they showed that performance on an explicit learning task predicted the children's vocabulary knowledge, which in turn predicted the children's morphological knowledge. That is, explicit learning was indirectly associated with morphological knowledge through its association with vocabulary. This was argued to be consistent with connectionist-style single route approaches to the language acquisition of the past tense.[4]

Thus, there is still no direct and specific empirical demonstration that implicit statistical learning is implicated in the acquisition of syntax. Although Kidd and Kirjavainen (2011) did not find an association between implicit learning and the acquisition of past tense morphology, they suggested that such an association was not likely because morphological processes are better characterized as lexical rather than syntactic. Instead, they predicted that implicit

---

[2] The same association was observed only in children with SLI following 42 min of exposure.

[3] Kaufman et al. (2010) reported significant associations between implicit learning in 17-year-olds and foreign language achievement scores from 1 to 2 years earlier. Although this is an impressive result, it is unclear what exactly this association means, given that foreign language achievement is assessed in many different ways (e.g., oral and written exams, listening comprehension, written pieces).

[4] This does not entail that implicit statistical learning is not implicated in vocabulary acquisition. On the contrary, work on infant statistical learning suggests that this skill drives segmentation, thus enabling infants to identify words in the speech stream (see Romberg & Saffran, 2010). Other research suggests that attention to cross-situational statistics aids word learning (e.g., Smith & Yu, 2008; Yu, 2008; Yu & Smith, 2007). The past tense results seem to suggest that the composition of complex morphological forms is dependent on lexical knowledge, consistent with usage-based approaches to language (e.g., Dąbrowska, 2008).

learning is likely to support the kind of sequencing required to analyze linguistic structure at the supralexical level, that is, when analyzing sequences of words. Several results in the adult literature provide preliminary support for this assertion. Conway, Bauernschmidt, Huang, and Pisoni (2010) reported on three studies that investigated the relationship between statistical implicit learning and language processing in adults. In each experiment, they reported a significant positive correlation between accuracy on an implicit learning task and a task that tested participants' ability to predict upcoming words in a sentence under degraded input. The correlations remained significant even after static measures of language, nonverbal intelligence, working memory, and executive control were partialed out. These correlations were intra- and intermodal; visual implicit learning predicted auditory sentence processing (Experiments 1 and 3), and auditory implicit learning predicted audiovisual sentence processing (Experiment 2). Conway et al. interpreted their results to suggest that implicit statistical learning supports the acquisition of knowledge about the predictability of items in a sequence (see also Conway & Pisoni, 2008; Conway, Pisoni, & Kronenberger, 2009). Such knowledge is of great importance in language comprehension, where the integration of incoming speech into the current analysis is eased by high predictability. Consistent with this interpretation, Misyak and Christiansen (2007, in press) showed that implicit learning of nonadjacent dependencies in an Artificial Grammar Learning (AGL) task predicted adults' processing of subject- and object-relative clause constructions (see also Levy, 2008; Misyak, Christiansen, & Tomblin, 2010; for models of production, see Chang et al., 2006; Jaeger, 2010). In a more recent study, Conway, Pisoni, Anaya, Karpicke, and Henning (2011) reported a significant correlation between implicit statistical learning and grammatical knowledge as measured by a standardized test of language in hearing impaired children.

Syntactic Priming

In the current study, I tested children's ability to acquire grammatical knowledge using the syntactic priming technique. In this method, an experimenter or interlocutor provides a description of a picture using a target structure (e.g., a passive). The child is then required to describe a different picture. If they use the same structure that was used by the experimenter, they are said to have been primed. The technique is useful because it is a controlled laboratory-based method of exposing participants to variations in input. Priming has been demonstrated in children in both production and comprehension using a variety of structures (e.g., Bencini & Valian, 2008; Huttenlocher, Vasilyeva, & Shimpi, 2004; Kidd, in press; Savage, Lieven, Theakston, & Tomasello, 2003, 2006; Thothathiri & Snedeker, 2008).

Just as language learning can be conceptualized as involving both implicit and explicit learning processes, syntactic priming has been categorized in the same manner. Evidence supporting a role for explicit learning processes in structural priming comes from the fact that priming has often been observed to be short-lived (e.g., Branigan, Pickering, & Cleland, 1999), suggesting activation yet rapid decay, and from the fact that priming effects are increased when there is lexical overlap between a prime and target sentence (the *lexical boost*; Cleland & Pickering, 2003; Hartsuiker,

Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008; Kaschak & Borreggine, 2008).

The role of implicit learning in priming derives from the argument that exposure to a syntactic structure increases the likelihood that the same structure will be reused because the selection of syntactic structure has been altered as a result of prior experience. Ferreira and Bock (2006) stated that an explanation of syntactic priming as implicit learning "is a reflection of a longer term process of learning how syntactic constructions in a speaker's language map onto the features of meaning that they express" (p. 1013). This explanation has some empirical support. Bock and Griffin (2000) showed that priming effects last over long time lags when priming through production, and Bock, Dell, Chang, and Onishi (2007) showed the same effect when priming through comprehension (see also Nitschke, Kidd, & Serratrice, 2010). Perhaps even more convincingly, developmental data from Savage et al. (2006) demonstrated that priming can last up to a month after exposure in 4-year-old children (see also Vasilyeva, Huttenlocher, & Waterfall, 2006). Finally, Ferreira, Bock, Wilson, and Cohen (2008) showed that people with anterograde amnesia, whose explicit learning ability is severely compromised but whose capacity for implicit learning is left relatively intact, were primed to the same extent as normally functioning adult controls but performed significantly worse on a recognition component of the task that aimed to test participants' explicit (declarative) memory for syntax.

The Current Study

The literature on language acquisition currently shows that (a) infants are capable of detecting statistical regularities over toy grammars and (b) children are sensitive to distributional properties of natural languages. Recent work in the adult literature has shown that implicit statistical learning predicts performance on language processing tasks, but similar work in the developmental literature that has concentrated on morphology has not reported the same relationship (though see Conway et al., 2011). The current study extends this work by investigating the dual role of implicit and explicit learning in syntactic acquisition. One hundred children ages 4 to 6 years were tested on measures of implicit (i.e., statistical) and explicit learning and on a syntactic priming task. A syntactic priming task was used because it provides a dynamic assessment of linguistic performance that indexes learning in response to changes in the input. As such, the study aimed to identify associations between implicit statistical learning, explicit learning, and changes in linguistic behavior in response to changes in input frequency (i.e., syntactic priming). Following Conway et al. (2010) and Misyak and Christiansen (2007, in press), it was hypothesized that implicit statistical learning would be directly associated with children's tendency to be primed in the syntactic priming task. Specifically, following the syntactic priming literature, I predicted that implicit statistical learning would be directly associated with long-term priming effects. In contrast, it was hypothesized that the association between explicit learning and priming would be weak or nonexistent, because the conditions under which the children were primed in the syntactic priming task were not consistent with explicit learning explanations for priming (i.e., there was no lexical overlap). That is, I primed children at the level of syntax and did not expect a major contribution from explicit, lexical processes.

Method

Participants

One hundred children ages 4 years 5 months to 6 years 11 months ($M = 5$ years 7 months) were recruited from primary schools drawn from a medium-sized city in Northern England, United Kingdom. Children of this age range were recruited because their implicit and explicit learning systems are still developing (Arciuli & Simpson, 2011; Lum, Kidd, Davis, & Conti-Ramsden, 2010), thereby maximizing the power of the individual differences analysis. The children's socioeconomic status was mixed yet representative of the region, with a majority being working-to-middle class. There were 44 boys and 56 girls. The children spoke English as a first language and had no cognitive or linguistic impairment. Because the children were native English speakers, ethnicity was not recorded.

Materials

The children were tested individually on a battery of assessments, including (a) a test of syntactic priming, (b) a test of implicit learning, (c) a test of explicit learning, (d) a standardized test of verbal ability, and (e) a standardized test of nonverbal ability. The presentation of the tests was pseudorandomized to avoid order effects. The children were tested over two sessions lasting approximately 30 to 40 minutes each, which were 1 week apart (± 2 days). Each test is briefly described next.

Test of implicit statistical learning. Implicit learning was tested using a version of the SRT task (Nissen & Bullemer, 1987), adapted for use with young children by Lum, Kidd, et al. (2010). In this task, participants implicitly learn a repeating 10-sequence pattern. The sequence consists of a single visual stimulus that moves between four spatial locations on a computer screen. The only instruction provided to participants is to press one of four buttons on a response panel that matches the location of the visual stimulus. For example, if the stimulus appears in Location 1, the participant is required to press Button 1 on the response panel. If the stimulus appears in Location 2, the participant is required to press Button 2, and so on. The repeating sequence was presented to participants over four blocks; in the final block, the visual stimulus appeared in a random order. Participants' RTs were the primary dependent variable of interest. In healthy children (e.g., Lum, Kidd, et al., 2010; Thomas & Nelson, 2001) and adults (e.g., Daselaar, Rombouts, Veltman, Raaijmakers, & Jonker, 2003), RTs typically decrease as they progress through the blocks with the repeating sequence and increase on the random block. This increase is often referred to as a *rebound effect* and is used to demonstrate whether implicit learning has occurred, because decreases in RTs observed over preceding blocks may represent practice effects. SRT tasks are arguably the best way to measure implicit learning. Unlike Artificial Grammar Learning tasks, where participants are explicitly told to memorize strings, learning in the SRT task is incidental, minimizing the role of explicit learning (Destrebecqz & Cleeremans, 2001; Kaufman et al., 2010). Further evidence supporting the assertion that the SRT task measures implicit learning comes from studies involving patients with basal ganglia pathology. Research has shown that patients with Huntington's disease (Kim et al., 2004; Knopman & Nissen, 1991)

and Parkinson's disease (Siegert, Taylor, Weatherall, & Abernethy, 2006), who have compromised implicit learning abilities, demonstrate a smaller rebound effect than controls.

In the current study, children completed the SRT task using a Gravis Gamepad Pro, which was connected to a Dell Latitude C620 laptop computer. The Gravis Gamepad Pro consists of four buttons arranged in the shape of a diamond, which children operated using their right thumb. Lum, Kidd, et al. (2010) showed that presenting the task in this way helped maintain children's interest in the task, because it is presented to children in the context of a computer game, serving to pique the children's interest in the task across the five blocks of trials. The visual stimulus consisted of a well-known cartoon character, which appeared in one of four spatial locations presented on the computer monitor with a black background. The spatial locations on the computer monitor were marked by four boxes with white boarders. The arrangement of these boxes was identical to the arrangement of the buttons on the response pad (i.e., a diamond configuration; see Figure 1). During testing, children sat approximately 40 cm away from the computer screen. The white boxes subtended $6.4° \times 6.4°$ of visual angle.

During testing, children were told that the cartoon character would appear in one of four places and that their task was to press the buttons that matched the character's location. Ten practice trials were presented to ensure the children understood the task. All children obtained an accuracy level of at least 90% on the practice trials. The test trials were then presented. The test trials consisted of five blocks of 60 trials. Unbeknown to the children, on Blocks 1 through 4, the appearance of the visual stimulus followed a 10-item sequence [4, 2, 3, 1, 3, 2, 4, 3, 2, 1]. The presentation of the stimulus in the fifth block was presented in a pseudorandom order, with the following two constraints. First, in this block, the visual stimulus appeared in each spatial location the same number of times as in each of the preceding blocks. Second, the probability of observing pairs of items within the repeating sequence was the same. Introducing this constraint meant I was able to control for the possibility that differences in performance between the repeating and pseudorandom blocks reflected the fact that children had learned only paired associations between picture transitions. The transitional probabilities between locations are shown in Table 1.

For the data analysis, I computed the children's median RT for each of the five blocks. Only RTs for correct responses were used. Thus, each child's performance on the SRT task was summarized with five data points reflecting the median RT of each block (four learning blocks and one random). Following previous research (e.g., Knopman & Nissen, 1991), implicit learning was indexed by
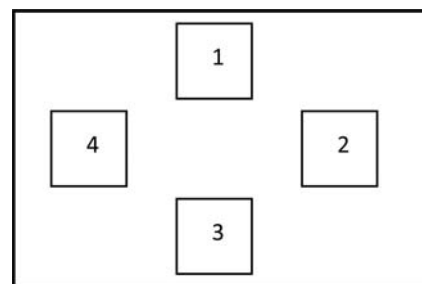

Figure 1. Spatial locations of the repeating stimulus in the Serial Reaction Time task.

Table 1

Transitional Probabilities Between Spatial Locations in the Serial Reaction Time Task

Location	Transitional probability			
	1	2	3	4
1	0	0	0.5	0.5
2	0.33	0	0.33	0.33
3	0.33	0.67	0	0
4	0	0.5	0.5	0

finding the difference in RTs between the patterned blocks and the random block; a significant elevation in RT from the patterned blocks to the random block suggests that the children learned the implicit pattern. Following past research, I call this index the *rebound effect*. For the individual differences analysis, the data were z normalized to control for individual differences in motor speed across the sample (see the Results section).

After completing the SRT task, the children were presented with a recall task that assessed their awareness of the pattern. In past research using SRT tasks, adults have typically been asked whether they detected a pattern during the test trials. Those participants who respond in the affirmative are asked to generate the sequence. Adults who are able to consciously recall the pattern are then excluded from the data analysis. Adhering strictly to this protocol with children was considered problematic given that children may provide a prosocial response; that is, they may indicate that they did identify a pattern to please the experimenter. In the recall task, children were not asked whether they recognized a pattern. Instead, they were informed that there was a pattern and were asked to recall it. Explicit knowledge of the pattern was assessed using a single trial. Children were seated in front of the computer screen. The cartoon character then appeared in Position 4, and the test administrator asked the child to indicate where the character would appear next. The children indicated their responses by pointing to one of the three boxes on the screen. The test administrator encouraged the children to provide 10 responses; that is, the child was encouraged to generate the entire repeating sequence. Credit was given for any two or more sequential positions recalled throughout the child-generated 10-location sequence, although credit was not given for repeated identical sequences (e.g., 4, 3, 4, 3, . . . was only counted as 2). Consistent with past research with children in this age range (Lum, Kidd, et al., 2010; Thomas & Nelson, 2001), none of the children were able to recall the complete pattern. The average number of correct sequential locations recalled was 3.23 ($SD = 1.16$), the mode was 2, and the range was 2–6.[5] The number of correct sequential positions that the children recalled was not related to the size of their rebound effect ($r = .027$, $p = .793$).

Measure of explicit learning. The Word Pairs subtest from the Children's Memory Scales (Cohen, 1997) was used to measure explicit (declarative) learning. Research has shown that performance on learning word pairs is impaired following pathology associated with the left medial temporal lobe (Jones-Gotman, 1992), the major neural structure thought to subserve declarative (i.e., explicit) memory (Squire, Stark, & Clark, 2004).

In this subtest, children are asked to learn a single list consisting of 10 word pairs (e.g., *nurse–fire*). Children are given three trials to learn the list. At the start of each trial, the list of word pairs is presented orally. Following this initial presentation, children are presented with the first word of the pair (e.g., *nurse*) and asked to recall the second (e.g., *fire*). This procedure is followed for the second and third presentation of the word pairs using a different presentation order. At the conclusion of the third trial, there is another recall task in which participants are asked to recall both words in the pair. The Word Pairs test therefore produces two indices that measure declarative memory. Children's performance on this task was described by summing the total number of correct responses over the three trials. Their raw scores were used in the analyses.

Test of verbal ability. Children's vocabulary was tested with the British Picture Vocabulary Scale (2nd ed.; Dunn, Dunn, Whetton, & Burley, 1997). This scale is a published standardized test that measures receptive vocabulary in children. In this test, children are orally presented with a word. Children are asked to identify the picture that matches the word from an array of four. Children's raw scores were used in the analyses.

Test of nonverbal IQ. The Raven's Colored Progressive Matrices (Raven, Raven, & Court 1998) was used to assess nonverbal reasoning. In this test, children are presented with a series of stimulus pictures of abstract patterns. In each picture there is a piece missing; the child's task is to choose the missing piece that matches the pattern in the stimulus picture from an array of six possible alternatives. The decision to include a measure of nonverbal reasoning was motivated by the concern that potential associations found between the measure implicit statistical learning and syntactic acquisition might reflect an association with general intelligence. Children's raw scores were used in the analyses.

Syntactic priming task. The structural priming task was designed to prime the English full *be* passive construction (e.g., *the guitar was played by the man*). The passive construction was chosen because (a) it has been the most successfully primed construction in acquisition research, and (b) it is very low in frequency in spoken language in its full *be* form and seems to be mastered only after formal reading instruction ($<1\%$ in either its full or truncated form in spoken English; Bencini & Valian, 2008). This maximized the chances of observing a priming effect, because children are unlikely to produce passives spontaneously, and increased the chance that a proportional increase in the use of passives could be attributed to learning.

Forty-two pictures depicting transitive scenes that could be described with either an active or a passive construction were used. Twelve of these were prime pictures, and 30 were test pictures that were rotated throughout the different testing phases of the task. The pictures all depicted scenes that contained different actions, such that the children were less likely to use a verb from a prime sentence in their descriptions of the pictures. The prime sentences were based on those used in previous studies (Bencini & Valian, 2008; Huttenlocher et al., 2004; Savage et al., 2003; see Appendix). They contained a mixture of animate and inanimate nouns. The majority of target pictures contained an animate agent performing an

---

[5] This number is likely to be slightly inflated, because it includes 14 children whose responses went totally in an anticlockwise direction, which credited them with five consecutive correct locations. Only one child recalled six consecutive correct locations.

action on an inanimate entity (e.g., *A girl eating a cookie*). This follows the tendency in natural languages for agents to be animate and themes to be inanimate (Fox & Thompson, 1990).

Because this was an individual differences study, the syntactic priming task was within subjects. The task consisted of three blocks: (a) baseline, (b) test, and (c) posttest. During the baseline block, the children were simply asked to describe a picture without guidance from the experimenter. Children described six pictures in this block. The next block was the test block, in which children were primed. In this block, the child was told that this time the experimenter would describe one picture and that they would have to describe the following picture. The children were asked to repeat the experimenter's prime sentence. The experimenter always described the primes using a full *be* passive containing a *by* phrase. The presentation of the prime sentences was pseudorandomized across the 10 lists. Following the test block, the children described a further six pictures without being primed. This posttest phase tested whether any priming effects observed persevered after priming had ceased, thus testing for the long-term priming effects.[6] This was a crucial component of the design, because, following the syntactic priming literature (see Pickering & Ferreira, 2008), I explicitly predicted that implicit learning would be directly associated with long-term priming effects.

The pictures that the children described were fully rotated through the three blocks of testing in 10 experimental lists. Each picture was rotated through each experimental block; thus, each picture had a baseline score, a test score, and a posttest score, ensuring that any priming effects were not simply due to individual quirks of any subset of pictures.

Scoring

Following Bencini and Valian (2008), multiple priming measures were calculated. Whereas they reported three measures (lax, strict, and adult), I report only two (lax and strict), because Bencini and Valian reported no substantial differences between the strict and adult coding schemes. The lax and strict coding schemes are described next.

Lax coding scheme. The children's descriptions of the target pictures were coded as (a) active, (b) passive, or (c) other. Active sentences required an overt subject, a transitive verb optionally preceded by a form of *be* or *have*, and an optional direct object. Passives required the patient in the subject position, the main verb optionally preceded by *be* or *get*, and the main verb optionally followed by the agent or instrument of the action within an adjunct headed by the prepositions *by* or *with*. The priming effect was calculated by dividing the total number of passive responses by the number of trials in each phase; that is, the priming effect was an index of the proportion of passive constructions produced as a function of the total number of opportunities the children were given in each phase of the experiment.

Strict coding scheme. The strict coding scheme was the same as the lax scheme, except that to be coded as a passive, the children's utterances needed to contain a postverbal adjunct that encoded the agent or instrument of the action headed by the prepositions *by* or *with*.

Results

A preliminary analysis showed that task order had no influence on performance in any of the measures. In the first analysis, I establish the presence of a significant priming effect. Table 2 shows proportion of passives produced during each experimental phase for both coding schemes.

Table 2 shows that on both the lax and strict coding schemes the proportion of passives increased from baseline during the test and posttest phases. The priming data were analyzed with the Generalized Estimating Equations function in SPSS. These equations are an extension of generalized linear models that are useful for analyzing binary data (Zeger & Liang, 1986). The children's response on each trial ($+/-$ passive) was the dependent variable, test phase and items were entered as fixed factors, and participants was entered as a random factor. For both coding schemes, the model was significant (lax: Wald $\chi^2 = 39.4$, $df = 2$, $p < .001$; strict: Wald $\chi^2 = 35.54$, $df = 2$, $p < .001$), showing that the proportion of passives differed across the three phases. Table 3 lists the results for subsequent models that were run to compare each phase to the others.

Table 3 shows that, for both coding schemes, the number of passives produced during the test and posttest phases were significantly greater than the number produced in the baseline phase. Significantly more passives were produced in the test phase than the posttest phase on the strict coding scheme; this difference was marginal for the lax coding scheme. These analyses show that on both coding schemes, the children produced more passives during the test phase than at baseline, thus indicating a priming effect at the group level. This effect was maintained into the posttest phase when the children were no longer primed.

Table 4 reports the descriptive statistics for children's performance on the test of verbal and nonverbal ability and the test of explicit learning. These variables were normally distributed. Table 5 reports on the children's accuracy on each block of the implicit statistical learning task. The accuracy scores were analyzed using a five-way univariate analysis of variance (ANOVA) with block accuracy as the dependent measure. The main effect of block was significant, $F(3.63, 359.79) = 3.14$, $p = .018$, partial $\eta^2 = .031$.[7] Post hoc Bonferroni pairwise comparisons revealed no significant differences between any of the blocks. Figure 2 reports the children's average RTs for each block on the statistical implicit learning task.

Figure 2 shows that children's mean RTs decreased from Blocks 1 through 3, remained at the same level between Blocks 3 and 4, and increased at the random block. A five-way univariate ANOVA with mean block RT as the dependent variable was significant, $F(2.65, 262.01) = 6.21$, $p = .001$, partial $\eta^2 = .059$. Post hoc Bonferroni pairwise comparisons showed that RTs in both Blocks 3 and 4 were significantly faster than RTs in the random block ($ps < .001$). The difference between Block 2 and the random block was marginally significant ($p = .062$). This pattern of responding,

---

[6] By *long-term,* I mean the use of the target structure beyond the target picture immediately following priming; in this case, *long-term* indicates the use of a passive when there is a lag of at least one experimental item following priming, which is the typical index of long- or longer term persistence (see Bock & Griffin, 2000).

[7] The statistic is Greenhouse–Geisser corrected.

Table 2
Means (and Standard Deviations) for Each Experimental Block of the Syntactic Priming Task

Coding scheme	Baseline	Test	Posttest
Lax	.03 (.07)	.13 (.014)	.095 (.16)
Strict	.013 (.05)	.097 (.13)	.06 (.14)

where RTs decrease over the patterned blocks but increase at the random block, indicates that the children implicitly learned the repeating pattern at the group level. This is consistent with past research that has shown that implicit learning on SRT tasks is observable in children as young as 4 years (Lum, Kidd, et al., 2010; Thomas & Nelson, 2001).

Table 6 reports the bivariate correlations between age, vocabulary, nonverbal intelligence, and the children's performance on the implicit statistical and explicit learning. The measure of implicit statistical learning used was the rebound effect, that is, each child's mean RT for the random block minus their mean RT for Block 4. Because the SRT task produces RTs and I was computing individual differences analyses, I treated the data further to reduce variability. Following Thomas et al. (2004), the children's RTs were z normalized to control for variability in motor speed across the sample. The RT for each child's correct response was converted to z scores, using their mean RT. This ensured that children could be compared without the added problem of differences in motor speed influencing the results. Then I computed the rebound effect on the basis of the means of the z scores for Block 4 and the rebound block (i.e., rebound = random – Block 4). Positive scores were indicative of greater procedural learning.

I now report on the role of implicit statistical and explicit learning in priming. As can be seen in Table 2, the number of passives that children used in both the test and posttest phases was low, and despite the large sample size, the priming scores were not normally distributed. As such, the next set of analyses investigated whether children's performance on the implicit and explicit learning tasks predicted their tendency to be primed in general.[8] Children were deemed to have been primed if they produced passives over and above the rate at which they produced them in the baseline phase. For the lax coding scheme, this was 61% of the sample ($N = 61$) for the test phase and 33% ($N = 33$) of the

Table 3
Summary of Pairwise Comparisons Between the Three Phases of the Syntactic Priming Task for the Lax and Strict Coding Schemes

Coding scheme	B	$SE\ (B)$	e^B	p
Lax				
Base versus test	1.67	0.28	5.31	<.001
Base versus posttest	1.35	0.32	3.86	<.001
Test versus posttest	0.33	0.17	1.39	.051
Strict				
Base versus test	2.1	0.35	8.17	<.001
Base versus posttest	1.58	0.39	4.85	<.001
Test versus posttest	0.48	0.22	1.62	.032

Note. B = log-odds coefficient; e^B = odds coefficient.

Table 4
Descriptive Statistics for Measures of Vocabulary, Nonverbal IQ, and Explicit Learning

Measure	M	SD	Range
Vocabulary[a]	103.9	14.8	70–137
Nonverbal IQ[a]	93.7	15.4	70–130
Explicit learning	10.5	6.6	0–27

[a] Standardized score ($M = 100$, $SD = 15$).

sample in the posttest phase. For the strict coding scheme, this was 48% ($N = 48$) of the sample for the test phase and 19% ($N = 19$) in the posttest phase. These numbers are consistent with the individual differences reported in syntactic priming in language acquisition (e.g., Bencini & Valian, 2008; Huttenlocher et al., 2004; Savage et al., 2003, 2006). I then compared the children who were primed with the children where were not primed on their scores for each learning task during both phases for both the lax and strict coding scheme. Four binary hierarchical logistic regressions were conducted, one for each phase of each coding scheme. Age (in months), vocabulary score (British Picture Vocabulary Scale), and nonverbal ability (Raven's Colored Progressive Matrices) were entered in the first block to control for any influence they may have had on priming; the measures of explicit (Word Pairs subtest) and implicit learning (rebound effect) were entered in the second block.

For the lax coding scheme in the test phase, the model was not significant after the first and second blocks (Block 1: $\chi^2 = 5.96$, $df = 3$, $p = .114$; Block 2: $\chi^2 = 6.96$, $df = 5$, $p = .224$). For the strict coding scheme in the test phase, the model was marginally significant at Block 1 ($\chi^2 = 7.5$, $df = 3$, $p = .058$). Nonverbal IQ was the only significant predictor ($B = .111$, $SE(B) = .056$, $e^B = 1.11$, $p = .046$); those children who were primed had significantly higher nonverbal ability scores. The inclusion of explicit and implicit learning did not significantly account for any additional variance in the model ($\chi^2 = 1.4$, $df = 2$, $p = .497$). The overall solution was not significant ($\chi^2 = 8.9$, $df = 5$, $p = .11$), and the contribution of nonverbal IQ became marginal. Table 7 details the solution of the final model.

The analyses for the posttest phase were conducted next. Neither model for either coding scheme was significant following the inclusion of the control variables (lax: $\chi^2 = 2.08$, $df = 3$, $p = .57$; strict: $\chi^2 = 1.57$, $df = 3$, $p = .67$). However, there was a significant change in variance explained following the inclusion of explicit and implicit learning ability (lax: $\chi^2 = 10.04$, $df = 3$, $p = .007$; strict: $\chi^2 = 10.89$, $df = 2$, $p = .004$). Both final models were significant (lax: $\chi^2 = 12.12$, $df = 5$, $p = .033$; strict: $\chi^2 = 12.47$, $df = 5$, $p = .029$); in both cases, the children's rebound effect significantly predicted group membership. Table 8 details the solution of the final model.

Figures 3 and 4 show both groups' (primed vs. not primed) performance on the explicit and implicit learning tasks for each test phase.

---

[8] Among the children who were primed, in both the test and posttest phases for both coding schemes the modal score was one passive. As such, dividing the children into those who were primed and who were not primed captures a key difference in the data set.

Table 5
Accuracy Data for Serial Reaction Time Task

Block	M	SD	Range
1	.92	.08	.63–1.0
2	.90	.072	.68–.98
3	.91	.073	.65–.98
4	.90	.076	.68–.98
Random	.90	.066	.68–.98

Table 6
Bivariate Correlations Between Age (in Months), Vocabulary, Nonverbal IQ, and Implicit and Explicit Learning

Variable	Implicit learning	Explicit learning	Nonverbal IQ	Vocabulary
Age	.129	.157	.285**	.180
Vocabulary	.183	.377**	.443**	
Nonverbal IQ	.195	.136		
Explicit learning	.086			

** $p < .01$.

Discussion

My hypothesis was supported: Performance on the implicit statistical learning task predicted the maintenance of the syntactic priming effect into the posttest phase of testing, where no additional primes where administered. In contrast, explicit learning did not predict priming. This is the first demonstration that children's performance on a measure of implicit statistical learning is associated with the detection of changes in the frequency of syntactic structure in their input. The result is consistent with comparable results conducted with adults (Conway et al., 2010; Misyak & Christiansen, 2007, in press; Misyak et al., 2010) and with suggestive results that show a statistical learning deficit in children with SLI (J. L. Evans et al., 2009; Lum, Gelgec, & Conti-Ramsden, 2010; Tomblin et al., 2007).

These data raise a number of issues. The first concerns the precise explanation of the association between implicit statistical learning and syntactic priming. At the procedural level, there is an important commonality between the two tasks: Each involves the repeated presentation of a complex stimulus. In the case of the implicit learning task, the pattern is too long to be easily memorized in its entirety. It is unclear whether success on this task involves only the computation of probabilistic transitions between the four spatial locations or whether learning also involves some degree of exemplar/chunk learning (see Jamieson & Mewhort, 2009; Perruchet & Pacton, 2006). However, these issues need not concern us just yet; the point is that learning of the pattern is implicit, gradual, and requires repeated presentation. The same argument could be made for the acquisition of syntax. Children do not have explicit knowledge of grammar: Acquisition is gradual,
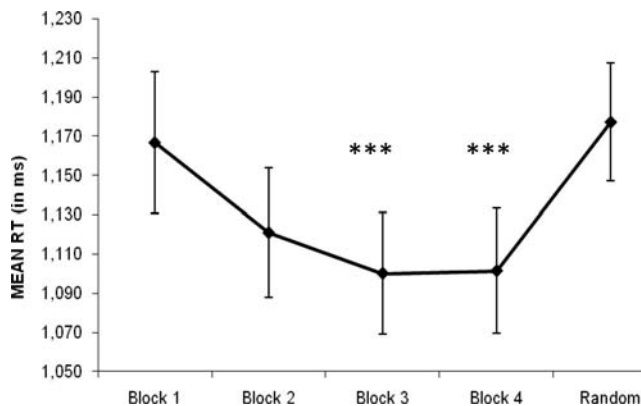

Figure 2. Mean response time (RT) for each block. *** $p < .001$ difference from the random block. Error bars represent standard error.

particularly for noncanonical forms, such as the passive, where children will hear hundreds if not thousands of tokens of the structure before they master its use. The syntactic priming task used in the current study repeatedly presented children with the English full *be* passive, a structure with which the children are unlikely to have had much experience. Exposure to 12 exemplars was enough for many of the children to raise their use of passives above baseline, and their use of the structure in the posttest phase was associated with better implicit statistical learning ability. The persistent use of the target structure in the absence of priming during the posttest phase is indicative of learning. I do not claim here that the children learned the passive from scratch during the priming task. Instead, following Ferreira and Bock (2006), I interpret this learning effect to mean "the longer-term process of learning how syntactic constructions in a speaker's language map onto the features of meaning that they express" (p. 1013). That is, even if the children came to the experiment with some knowledge of the passive, those who were primed into the posttest period left with a stronger mapping between the passive construction and transitive events, a mapping that was supported by an implicit statistical learning mechanism.

Conway et al. (2010) interpreted the relationship they observed between implicit statistical learning and language processing to reflect the acquisition of knowledge about the *predictability* of items in a sequence. That is, they suggest that "superior implicit learning abilities result in more detailed and robust representations of the word order probabilities in spoken language" (Conway et al., 2010, p. 365). This proposal has merit. In particular, it provides a unifying conceptualization of the broad range of research that investigates statistical learning in language acquisition. For instance, statistical learning is indexed in infant artificial grammar learning studies through head preference procedures, where children either prefer to listen to, or show habituation to, predictable segments of speech on which they have been trained. Studies that have reported frequency effects in older children's linguistic behavior have shown that high-frequency and therefore highly predictable sequences of language are easier to process than low-frequency sequences (e.g., Bannard & Matthews, 2008; Kidd et al., 2006, 2007; Matthews & Bannard, 2010).

Prediction also captures both the implicit learning task and the syntactic priming task used in the current study. The degree to which children learn the implicit pattern in the implicit learning task is indexed by their rebound effect. Because the random block does not have the same pattern as the patterned blocks, those children who learned the pattern well should show slower RTs in the random block. That is, because the predicted location of the

Table 7
Logistic Regression Results From the Test Phase for the Lax and Strict Coding Schemes

Predictor variable	B	Wald	e^B	CI$_{95}$ e^B	p	R^2
Lax coding						
Months	−0.014	0.278	0.986	[0.935, 1.04]	.598	.091
Vocabulary	0.007	0.149	1.00	[0.973, 1.04]	.699	
Nonverbal IQ	0.10	3.24	1.11	[0.991, 1.23]	.072	
Word Pairs test	0.001	0.001	1.01	[0.943, 1.07]	.981	
Rebound	0.76	0.985	2.14	[0.476, 9.60]	.321	
Constant	−0.683	0.14	0.505		.708	
Strict coding						
Months	−0.008	0.078	0.992	[0.938, 1.05]	.78	.130
Vocabulary	0.01	0.335	1.01	[0.976, 1.05]	.563	
Nonverbal IQ	0.109	3.66	1.12	[0.997, 1.25]	.056	
Word Pairs test	0.003	0.006	1.00	[0.933, 1.08]	.939	
Rebound	0.966	1.35	2.63	[0.515, 13.42]	.245	
Constant	−1.78	0.839	0.169		.360	

Note. B = log-odds coefficient; e^B = odds coefficient.

visual stimulus will be incorrect in the random block, those with good implicit learning ability will show the greatest interference and therefore the largest rebound effect. Prediction is also an important component of the mechanism thought to underlie syntactic priming. Chang et al. (2006) described a connectionist model that acquires syntax through an implicit learning mechanism that maps input onto event semantic roles. The implicit learning mechanism is a simple recurrent network (Elman, 1990), which predicts an upcoming word on the basis of a memory representation for the previous word it has processed or produced. The model explains priming through the network's use of error-based learning. The network is presented with a prime sentence word by word and incrementally attempts to parse the sentence. Any error is fed back into the network through back-propagation of error (Rumelhart & McClelland, 1986), after which connection weights are changed to better reflect the linguistic system the network is learning. The changes in connection weights increase the chance that the target structure will be used in the future. As such, the model effectively explains priming (and hence language acquisition) as implicit

learning that is driven by incorrect predictions, the outcome of which are used as feedback to recalibrate the linguistic system.

The suggestion here is that an implicit statistical learning mechanism is important for the acquisition of syntax and that acquisition can at least partially be categorized as identifying and recording statistical relations between words. Thompson and Newport (2007) have argued that, just as infants compute transitional probabilities over syllables to segment speech, children might compute transitional probabilities over word classes (i.e., syntactic categories) to induce phrase structure. The identification and/or induction of syntactic categories in acquisition is a nontrivial process that has been debated elsewhere (e.g., Naigles, 2002; Tomasello, 2003), although distributional information in addition to other probabilistic cues are likely to play a role (Mintz, Newport, & Bever, 2002; Monaghan, Chater, & Christiansen, 2005; Monaghan, Christiansen, & Chater, 2007). There is inherent ambiguity in identifying category information; for example, many lexemes can have more than one syntactic function (e.g., *run*). The potential learnability problem that syntactic category ambiguity might cause could be

Table 8
Logistic Regression Results From the Posttest Phase for the Lax and Strict Coding Schemes

Predictor variable	B	Wald	e^B	CI$_{95}$ e^B	p	R^2
Lax coding						
Months	−0.02	0.46	0.981	[0.927, 1.04]	.497	.163
Vocabulary	0.019	1.05	1.02	[0.983, 1.06]	.305	
Nonverbal IQ	−0.012	0.047	0.983	[0.886, 1.1]	.828	
Word Pairs test	−0.02	0.21	0.983	[0.911, 1.06]	.649	
Rebound	2.96	8.07	19.2	[2.5, 147.58]	.005**	
Constant	−0.762	0.153	0.467		.695	
Strict coding						
Months	−0.022	0.340	0.978	[0.909, 1.05]	.56	.225
Vocabulary	−0.016	0.316	0.984	[0.931, 1.04]	.574	
Nonverbal IQ	−0.076	0.812	0.926	[0.784, 1.09]	.368	
Word Pairs test	−0.069	1.47	0.933	[0.834, 1.04]	.226	
Rebound	3.69	7.75	40.02	[2.98, 537.4]	.005**	
Constant	2.06	0.517	7.87		.472	

Note. B = log-odds coefficient; e^B = odds coefficient.
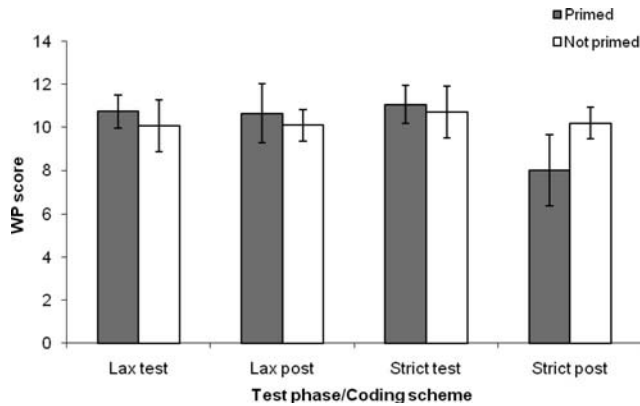** $p < .01$.

Figure 3. Mean explicit learning scores (and standard errors) in children who were and who were not primed across the test and posttest phases for both coding schemes. Error bars represent standard error.

circumvented by an initial focus on probabilistic relationships between specific words rather than word classes. For instance, Bannard et al. (2009) reported on a computational model that showed children's naturalistic speech at 2 years was best categorized by an item-specific grammar that did not contain any abstract category knowledge (e.g., Lieven, Behrens, Speares, & Tomasello, 2003). In contrast, performance at 3 years significantly improved following the addition of the abstract categories of noun and verb. Item-specific stored sequences appear to play a role even in adult language (Arnon & Snider, 2010), suggesting the possibility that statistical learning in syntax operates over multiple types of information (for a statistical learning algorithm that does not use any abstract category information, see Chang et al., 2008).

An important issue arising from these data concerns the extent to which statistical learning in language acquisition is supported by a stimulus-specific or domain-general learning mechanism. The current article showed an association between implicit statistical learning in the visual modality and syntactic acquisition, suggesting that a domain general mechanism makes at least a partial contribution. Conway et al. (2010) also reported cross-modal effects. However, Conway and Pisoni (2008) argued that it is logically possible that multiple domain-relevant implicit learning subsystems exist, which all have similar computational principles. Indeed, Conway and Christiansen (2006) showed that participants exposed to two different artificial grammars in different modalities (visual and auditory) learned the two systems just as well as a group that learned one grammar alone. However, performance was compromised when two grammars were presented within the same modality, suggesting the operation of parallel modality-specific learning mechanisms. If such parallel mechanisms exist, then the data from the current study suggest that either (a) there is a degree of commonality between the modality-specific mechanisms such that statistical learning as observed in one modality (e.g., visual) is also evident in another (e.g., auditory), or (b) an individual's ability in one domain is highly correlated with ability in other domains. This is another area that requires more research.

In the wider literature on implicit learning there is a general assumption that the skill supports learning across a range of systems and tasks. As such, it is entirely likely that the association between implicit statistical learning and syntactic priming that I

observed could be predicted for any learning task that involves the experience-driven tuning that occurs between processing elements following repeated exposure to stimuli (see Gupta & Cohen, 2002). In fact, Kaufman et al. (2010) have recently shown that teenagers' (ages 16–17) performance on an SRT task was associated with performance on a variety of cognitive and personality variables, such as verbal analogical reasoning, processing speed, foreign language learning, intuition, openness to experience, and impulsivity. In contrast, SRT task performance was not related to measures of working memory and explicit learning, nor was it strongly associated with general intelligence (*g*). They interpreted their results to support dual-process theories of cognition, where different cognitive skills are differentially supported by putatively implicit and explicit processes (e.g., J. S. B. T. Evans & Frankish, 2009). From the perspective of language acquisition research, an approach arguing that the acquisition of syntax is supported by an implicit statistical learning mechanism that supports skill learning in other domains is a radical departure from traditional linguistic nativist approaches to acquisition (e.g., Guasti, 2004; Snyder, 2007; though see Yang, 2004). This approach follows Chomsky (1980) in assuming that language is a domain-specific skill that is unrelated to other cognitive processes. However, the statistical learning approach is wholly consistent with domain-general approaches that argue that human language has a phylogenetic and ontogenetic basis in general cognition (e.g., Bates & MacWhinney, 1982; Tomasello, 2003, 2008).

At the same time, I must point out that these data do not speak to the issue of innateness of linguistic knowledge. My simple claim is that implicit statistical learning is implicated in learning grammatical usage patterns of the input language, which, to date, has not been shown. More studies are needed to determine the scope and limitations of implicit statistical learning in acquisition throughout childhood. Longitudinal research that explored implicit statistical learning at different levels of linguistic analysis (e.g., phonology, syntactic) at developmentally significant ages would be particularly illuminating. Additionally, it would be worthwhile to investigate the role of implicit learning in children's ability to learn novel syntactic constructions, for which there is a rich history in language acquisition research (e.g., Abbot-Smith, Lieven, &
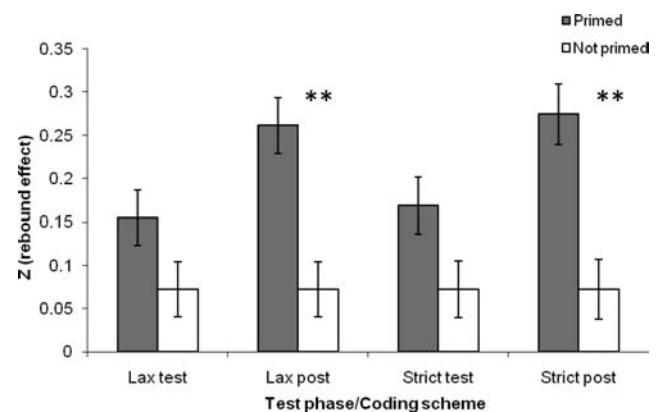

Figure 4. Mean implicit learning scores (i.e., rebound effect and standard errors) in children who were and who were not primed across the test and posttest phases for both coding schemes. Error bars represent standard error. ** *p* < .01. Significance test refers to logistic regression result.

Tomasello, 2001; Akhtar, 1999; Chang, Kobayahsi, & Amano, 2009; Casenhiser & Goldberg, 2005; Matthews et al., 2005).

Conclusion

It has long been suggested that language acquisition involves the detection of statistical regularities from the input. Although there is a rich literature showing that even young infants are sensitive to statistical variation, the current study is the first to show that children's performance on an independent test of statistical learning is directly associated with changes in children's syntactic behavior in response to changes in input frequencies.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development, 16,* 679–692. doi:10.1016/S0885-2014(01)00054-5

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26,* 339–356. doi:10.1017/S030500099900375X

Akhtar, N., Callanan, M., Pullum, G., & Scholz, B. (2004). Learning antecedents for anaphoric *one. Cognition, 93,* 141–145. doi:10.1016/j.cognition.2003.12.002

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science, 14,* 464–473. doi:10.1111/j.1467-7687.2009.00937.x

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62,* 67–82. doi:10.1016/j.jml.2009.09.005

Bannard, C., Lieven, E., & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *PNAS: Proceedings of the National Academy of Sciences, 106,* 17284–17289. doi:10.1073/pnas.0905638106

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19,* 241–248. doi:10.1111/j.1467-9280.2008.02075.x

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

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). New York, NY: Cambridge University Press.

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language, 59,* 97–113. doi:10.1016/j.jml.2007.12.007

Bock, J. K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104,* 437–458. doi:10.1016/j.cognition.2006.07.003

Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129,* 177–192. doi:10.1037/0096-3445.129.2.177

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German- and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics, 20,* 539–570. doi:10.1515/COGL.2009.024

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Structural priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review, 6,* 635–640. doi:10.3758/BF03212972

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes, 10,* 425–455. doi:10.1080/01690969508407111

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science, 8,* 500–508. doi:10.1111/j.1467-7687.2005.00441.x

Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological Review, 113,* 234–272. doi:10.1037/0033-295X.113.2.234

Chang, F., Kobayashi, T., & Amano, S. (2009). Social factors in the acquisition of a new word order. *First Language, 29,* 427–445. doi:10.1177/0142723709105316

Chang, F., Lieven, E., & Tomasello, M. (2008). Automatic evaluation of syntactic learners in typologically-different languages. *Cognitive Systems Research, 9,* 198–213. doi:10.1016/j.cogsys.2007.10.002

Chomsky, N. (1980). *Rules and representations.* New York, NY: Columbia University Press.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language, 49,* 214–230. doi:10.1016/S0749-596X(03)00060-3

Cohen, M. J. (1997). *Children's memory scales.* San Antonio, TX: Psychological Corporation.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114,* 356–371. doi:10.1016/j.cognition.2009.10.009

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17,* 905–912. doi:10.1111/j.1467-9280.2006.01801.x

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences, 1145,* 113–131. doi:10.1196/annals.1416.009

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., & Henning, S. C. (2011). Implicit sequence learning in deaf children with cochlear implants. *Developmental Science, 14,* 69–82. doi:10.1111/j.1467-7687.2010.00960.x

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science, 18,* 275–279. doi:10.1111/j.1467-8721.2009.01651.x

Dąbrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language, 58,* 931–951. doi:10.1016/j.jml.2007.11.005

Daselaar, S. M., Rombouts, S. A., Veltman, D. J., Raaijmakers, J. G., & Jonker, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of Aging, 24,* 1013–1019. doi:10.1016/S0197-4580(03)00030-7

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8,* 343–350. doi:10.3758/BF03196171

Diessel, H. (2004). *The acquisition of complex sentences.* Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511486531

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd ed.). London, United Kingdom: NFER-Nelson.

Elman, J. (1990). Finding structure in time. *Cognitive Science, 14,* 179–211. doi:10.1207/s15516709cog1402_1

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 52,* 321–335. doi:10.1044/1092-4388(2009/07-0189)

Evans, J. S. B. T., & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond.* New York, NY: Oxford University Press.

Ferreira, V. S., & Bock, J. K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21,* 1011–1029. doi:10.1080/01690960600824609

Ferreira, V. S., Bock, J. K., Wilson, M. P., & Cohen, N. J. (2008). Memory for syntax despite amnesia. *Psychological Science, 19,* 940–946. doi:10.1111/j.1467-9280.2008.02180.x

Fox, B. A., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language, 66,* 297–316. doi:10.2307/414888

Gold, E. M. (1967). Language identification in the limit. *Information and Control, 10,* 447–474. doi:10.1016/S0019-9958(67)91165-5

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics, 15,* 289–316. doi:10.1515/cogl.2004.011

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2005). The role of prediction in construction-learning. *Journal of Child Language, 32,* 407–426. doi:10.1017/S0305000904006798

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70,* 109–135. doi:10.1016/S0010-0277(99)00003-7

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4,* 178–186. doi:10.1016/S1364-6613(00)01467-4

Guasti, M. T. (2004). *Language acquisition: The growth of grammar.* Cambridge, MA: MIT Press.

Gupta, P., & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review, 109,* 401–448. doi:10.1037/0033-295X.109.2.401

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language, 58,* 214–238. doi:10.1016/j.jml.2007.07.003

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45,* 337–374. doi:10.1016/S0010-0285(02)00500-5

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language, 50,* 182–195. doi:10.1016/j.jml.2003.09.003

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61,* 343–365. doi:10.1016/j.cogpsych.2010.08.002

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61,* 23–62. doi:10.1016/j.cogpsych.2010.02.002

Jamieson, R., & Mewhort, D. J. K. (2009). Applying an exemplar model to the serial reaction-time task: Anticipating from experience. *Quarterly Journal of Experimental Psychology, 62,* 1757–1783. doi:10.1080/17470210802557637

Jones-Gotman, M. (1992). Presurgical neuropsychological evaluation for localization and lateralization of seizure focus. In H. Luders (Ed.), *Epilepsy surgery* (pp. 469–485). New York, NY: Raven Press.

Kaschak, M. P., & Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language, 58,* 862–878. doi:10.1016/j.jml.2006.12.002

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

Kidd, E. (in press). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics.* doi: 10.1017/S0142716411000415

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives

made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes, 22,* 860–897. doi:10.1080/01690960601155284

Kidd, E., & Kirjavainen, M. (2011). Investigating the contribution of procedural and declarative memory to the acquisition of past tense morphology: Evidence from Finnish. *Language and Cognitive Processes, 26,* 794–829. doi:10.1080/01690965.2010.493735

Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in children's acquisition and processing of sentential complements. *Cognitive Development, 21,* 93–107. doi:10.1016/j.cogdev.2006.01.006

Kidd, E., Lieven, E., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences, 32,* 132–142. doi:10.1016/j.langsci.2009.05.002

Kim, J. S., Reading, S. A. J., Brashers-Krug, T., Calhoun, V. D., Ross, C. A., & Pearlson, G. D. (2004). Functional MRI study of a serial reaction time task in Huntington's disease. *Psychiatry Research: Neuroimaging, 131,* 23–30. doi:10.1016/j.pscychresns.2004.03.002

Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia, 29,* 245–254. doi:10.1016/0028-3932(91)90085-M

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831–843. doi:10.1038/nrn1533

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177. doi:10.1016/j.cognition.2007.05.006

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18-months. *Cognition, 89,* B65–B73. doi:10.1016/S0010-0277(03)00116-1

Lieven, E. (2010). Input and first language acquisition. *Lingua, 120,* 2546–2556. doi:10.1016/j.lingua.2010.06.005

Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage based approach. *Journal of Child Language, 30,* 333–370.

Lum, J., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language and Communication Disorders, 45,* 96–107. doi:10.3109/13682820902752285

Lum, J., Kidd, E., Davis, S., & Conti-Ramsden, G. (2010). Longitudinal study of declarative and procedural memory in primary school-aged children. *Australian Journal of Psychology, 62,* 139–148. doi:10.1080/00049530903150547

MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language, 31,* 883–914. doi:10.1017/S0305000904006336

Matthews, D., & Bannard, C. (2010). Children's production of unfamiliar word sequences is predicted by positional variability and latent classes in a large sample of child directed speech. *Cognitive Science, 34,* 465–488. doi:10.1111/j.1551-6709.2009.01091.x

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20,* 121–136. doi:10.1016/j.cogdev.2004.08.001

Mintz, T. H., Newport, E. L. & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science, 26,* 393–442.

Misyak, J. B., & Christiansen, M. H. (2007). Extending statistical learning farther and further: Long-distance dependencies and individual differences in statistical learning and language. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual meeting of the Cognitive Science Society* (pp. 1307–1312). Austin, TX: Cognitive Science Society.

Misyak, J. B., & Christiansen, M. H. (in press). Statistical learning and

language: An individual differences study. *Language Learning.* doi: 10.1111/j.1467–9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1.* doi:10.3389/fpsyg.2010.00031

Monaghan, P., Chater, N., & Christiansen, M. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition, 96,* 143–182. doi:10.1016/j.cognition.2004.09.001

Monaghan, P., Christiansen, M., & Chater, N. (2007). The phonological–distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology, 55,* 259–305. doi:10.1016/j.cogpsych.2006.12.001

Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition, 86,* 157–199. doi:10.1016/S0010-0277(02)00177-4

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

Nitschke, S., Kidd, E., & Serratrice, L. (2010). First language transfer and long-term structural priming in comprehension. *Language and Cognitive Processes, 25,* 94–114. doi:10.1080/01690960902872793

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238. doi:10.1016/j.tics.2006.03.006

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134,* 427–459. doi:10.1037/0033-2909.134.3.427

Pinker, S. (1999). *Words and rules: The ingredients of language.* New York, NY: Science Masters.

Pullum, G., & Scholz, B. (2002). Empirical assessment of stimulus poverty arguments. *Linguistic Review, 19,* 9–50. doi:10.1515/tlir.19.1-2.9

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for the Raven's Progressive Matrices and Vocabulary Scales.* Oxford, United Kingdom: Oxford Psychologist Press.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews in Cognitive Science, 1,* 906–914. doi:10.1002/wcs.78

Rowland, C. F. (2007). Explaining errors in children's questions. *Cognition, 104,* 106–134. doi:10.1016/j.cognition.2006.05.011

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. doi:10.1017/S0305000999004055

Rumelhart, D., & McClelland, J. (Eds.). (1986). *Parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12,* 110–114. doi:10.1111/1467-8721.01243

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928. doi:10.1126/science.274.5294.1926

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In D. Kuhn & R. Seigler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (6th ed., Vol. 2, pp. 58–108). Hoboken, NJ: Wiley.

Savage, C., Lieven, E. V. M., Theakston, A. L., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6,* 557–567. doi:10.1111/1467-7687.00312

Savage, C., Lieven, E. V. M., Theakston, A. L., & Tomasello, M. (2006).

Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2,* 27–49. doi:10.1207/s15473341lld0201_2

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,* 490–495. doi:10.1037/0894-4105.20.4.490

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

Snyder, W. (2007). *Child language: The parametric approach.* Oxford, United Kingdom: Oxford University Press.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82,* 171–177. doi:10.1016/j.nlm.2004.06.005

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2002). 'Going, going, gone': The acquisition of the verb 'go'. *Journal of Child Language, 29,* 783–811. doi:10.1017/S030500090200538X

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language, 31,* 61–99. doi:10.1017/S0305000903005956

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience, 16,* 1339–1351. doi:10.1162/0898929042304688

Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool- and school-age children. *Journal of Experimental Child Psychology, 79,* 364–387. doi:10.1006/jecp.2000.2613

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development, 3,* 1–42.

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year old children. *Journal of Memory and Language, 58,* 188–213. doi:10.1016/j.jml.2007.06.012

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication.* Cambridge, MA: MIT Press.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3,* 269–293. doi:10.1080/15475440701377477

Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effect of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology, 42,* 164–174. doi:10.1037/0012-1649.42.1.164

Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences, 8,* 451–456. doi:10.1016/j.tics.2004.08.006

Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning. *Language Learning and Development, 4,* 32–62. doi:10.1080/15475440701739353

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

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42,* 121–130. doi:10.2307/2531248

(*Appendix follows*)

Appendix

Prime Sentences

1. The toys were carried by the wagon.

2. The food was cooked by the man.

3. The target was shot by the arrow.

4. The feather was blown by the girl.

5. The biscuit was eaten by the girl.

6. The guitar was played by the man.

7. The door was opened by the boy.

8. The treasure was buried by the pirate.

9. The flower was watered by the rain.

10. The cake was cut by the knife.

11. The window was broken by the ball.

12. The fish was caught by the net.