Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 52 (2005) 444–467
Journal of Memory and Language
www.elsevier.com/locate/jml

Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension

Barbara J. Luka [a,*], Lawrence W. Barsalou [b]

[a] *Department of Psychology, Bard College, P.O. Box 5000, Annandale-on-Hundson, NY 12504-5000, USA*
[b] *Emory University, USA*

Received 8 September 2004; revision received 19 January 2005

Abstract

In five experiments, participants first read grammatical sentences of English and later rated identical, structurally similar, or novel sentences for grammatical acceptability. The experimental method was modeled after "mere exposure" and artificial grammar learning paradigms in which preference ratings are enhanced by prior experience with the material. Participants rated sentences as more grammatical if they had read them earlier. Increased grammaticality ratings were also observed for sentences that shared syntactic structure, but not content words, with those read earlier. A single prior exposure to a similar sentence was sufficient to induce this structural facilitation effect, although more exposures enhanced the effect. We interpret the results with respect to frequency sensitive models of parsing and to syntactic priming observed in language production, and we consider the available evidence for shared representations or mechanisms for language production and comprehension.

Keywords: Structural priming; Structural facilitation; Grammaticality judgment; Evaluative categorization; Sentence processing; Mere exposure

Speakers tend to reuse or repeat sentence structures that they have recently read, heard, or spoken themselves (e.g., Bock, 1986; Branigan, Pickering, Liversedge, Stewart, & Urbach, 1995; Potter & Lombardi, 1998). For instance, a participant who has just heard the sentence "The secretary was taking her boss a cake." (a double-object dative structure) will be more likely to describe an appropriate picture as "The girl handed the man a paintbrush." (also a double-object dative) than as "The girl handed the paintbrush to the man." (prepositional object; Bock, 1989). The typical measure of syntactic priming is the likelihood for a participant to produce one type of structure compared to an alternative structure, although syntactic priming has also been demonstrated using measures of latency to production (Smith & Wheeldon, 2001; Corley & Scheepers, 2002). Syntactic priming for spoken and written sentence production has been demonstrated for at least the following types of constructions in English: active versus passive, prepositional versus double object dative, and optional complementizer production (e.g., Bock, 1989; Ferreira, 2003; Pickering & Branigan, 1998; Pickering, Branigan, & McLean, 2002). Studies have also found syntactic

* Corresponding author. Fax: 1 845 758 9654.
E-mail address: Luka@bard.edu (B.J. Luka).

0749-596X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2005.01.013

priming in German and in Dutch (Hartsuiker & Kolk, 1998; Hartsuiker, Kolk, & Huiskamp, 1999; Hartsuiker and Westenberg, 2000; Scheepers, 2003). Syntactic priming has been attributed to processes specific to language production, such as positional processing or linearization during grammatical encoding (Hartsuiker et al., 1999), although procedural accounts have been proposed that are not restricted to mechanisms of language production (Bock & Loebell, 1990).

The degree to which comprehension and production rely on a common grammatical or syntactic representation is unclear. The fact that spontaneous repetition does occur in natural discourse between interlocutors would argue that listening and speaking share a level of common representation; however, conversational exchanges also tend to share lexical, morphological, and situational elements that make it difficult to attribute observed repetitions to strictly structural factors (Bock & Loebell, 1990; Branigan, Pickering, & Cleland, 2000). Nonetheless, comprehension-to-production priming has been demonstrated under controlled experimental conditions using a structured conversational exchange as part of a card description task (Branigan et al., 2000; Cleland & Pickering, 2003). On the assumption that sentence parsing in comprehension and grammatical encoding in production depend on different cognitive processes, structural repetition in discourse has been interpreted as evidence that syntactic priming may be due to priming of abstract grammatical representations that are common to the comprehension and production systems (Branigan et al., 1995, 2000).

In contrast to research in sentence production, the evidence for structural repetition effects in language comprehension is more limited. Noppeney and Price (2004) found that reading times for syntactically ambiguous sentences decrease if participants have recently read blocks of sentences requiring similar structural disambiguation. Readers are thus faster to process an "early closure" sentence like "While the woman was eating the creamy soup went cold." if they had recently read "While the lady was knitting the scarf fell off her lap." (Branigan, Pickering, & Stewart, 1995 as cited in Branigan et al., 1995). Kaschak and Glenberg (2004) found that participants showed faster reading times after repeated exposure to a novel construction (the *needs* construction, as in "The grass needs cut."). [Kaschak and Glenberg (2004) explain that the *needs* construction is acceptable to some individuals in some regions of the United States, but was novel to the participants in their studies.] Kaschak and Glenberg also observed that participants generalized their new understanding to a related *wants* construction (e.g., "The dog needs walked."), again finding faster reading times for restricted types of syntactic structures under repeated presentation. Although, the ability to generalize these findings to other types of syntactic structures is necessarily limited, these studies support the conclusion that parsing can be influenced by recent or frequent experience in processing specific types of ambiguous or novel syntactic structures (Mitchell, Cuetos, Corley, & Brysbaert, 1995). Evidence for comprehension priming of structurally diverse, non-ambiguous sentences is much sparser. Frazier, Taft, Roeper, Clifton, and Ehrlich (1984) found faster reading times for the second clause of a conjoined sentence when the first and second clauses had parallel structure, and argued for a type of perseveration in processing so that local syntactic consistency increased the ease of subsequent processing. The limited data concerning structural repetition effects in comprehension may be due to the restrictions imposed by the reading time measures that have been used to assess priming. Reading times are influenced by the number of words in a sentence, the frequency of usage of those words, and a variety of other factors (Rayner & Clifton, 2002). The necessity of matching these non-syntactic factors limits the range of sentence variability that can be feasibly examined. For instance, while the word sequence "the only mortal Gorgon" is structurally equivalent at one abstract syntactic level to the single lexeme "Medusa" (both are noun phrases), these phrasal variants cannot be substituted in a reading time paradigm. It has thus been difficult to strip away lexical, conceptual, and contextual factors from comprehension paradigms to examine structural repetition effects in isolation (Mitchell, 1994).

In contrast to studies using reading times, Mehler and Carey (1967) used accuracy of perception in noise as a dependent measure. They presented blocks of structurally homogeneous sentences auditorily in white noise and found that perception of an anomalous sentence structure presented at the end of each block was improved if the test structure shared similar phrase structure with the preceding block. Although Mehler and Carey (1967) relate their methodology to the click perception study of Fodor and Bever (1965), accuracy of detection for perceptually degraded stimuli is a standard measure of perceptual priming that has been related to implicit memory (Schacter, 1987). If their results were to be interpreted as conceptual facilitation for repeated phrase structure, their conclusions would be most relevant to the current investigation of structural facilitation in comprehension and its relation to implicit memory. However, there are a number of methodological concerns that make this conclusion unwarranted. First, each participant in Mehler and Carey's study ($N = 45$) responded to only two of four total test sentences, and no item analyses are provided. Second, the sentences were recorded in a monotone voice in order to reduce intonation cues, and there is some evidence that this non-natural articulation did not influence each of the sentence types equally. The four test sentences were not independently normed for equal comprehensibility in the

absence of white noise, and one of the four sentences had markedly different error rates in the control condition, indicating a disparity that is independent of the experimental manipulation. Third, their materials contrasted two sets of two sentence types ("They are forecasting cyclones." versus "They are conflicting desires." and "They are delightful to embrace." versus "They are hesitant to travel."). Participants were alerted to the fact that all of the sentences they would hear in the course of the experiment would begin with the phrase "They are…," making it difficult to extrapolate an effect to a wider range of sentence types. Furthermore, a paradigm using auditory sentence perception in noise introduces a number of additional factors that may be of concern, such as speed of presentation for each syllable, acoustic properties of word initial syllables, or cohort size. These concerns make an auditory perceptual or conceptual priming methodology even more challenging to pursue than methods using visual presentation of stimuli.

In the present experiments, we investigate a different, non-speeded measure of structural repetition in sentence comprehension, *structural facilitation*, defined as increased preference for the grammatical structure of a recently encountered sentence. We suggest that increased ratings of grammaticality acceptability after exposure to structurally similar sentences can serve as an indirect measure of facilitated processing. Analogous to the proposal of Bock and colleagues for production priming (Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000), we explore the possibility that structural facilitation in language comprehension is a form of implicit learning. Below, we suggest that exposure-induced changes in evaluative ratings—such as changes in judgments of grammatical acceptability—have something in common with the *mere exposure* effect (Zajonc, 1980), and with artificial grammar learning paradigms (Reber, 1967), and that these memory paradigms can be extended to examine sentence comprehension. We then present five experiments using the new method and interpret the results in the context of models of parsing that are sensitive to the frequency and recency of syntactic structures.

Evaluative responses index learning

Increased preference for repeated stimuli is an established measure of facilitated processing, but one that has been used only sparsely in psycholinguistics. Zajonc (1968) first defined the *mere exposure* effect as an increase in liking or preference for previously processed items, an effect he observed in response to a variety of stimuli, including words, nonsense words, abstract symbols, and faces. Increased liking of previously presented stimuli is evident even when stimuli are presented very rapidly or in perceptually degraded formats such that participants are unable to consciously recognize them (Kunst-Wilson & Zajonc, 1980). Zajonc (1980, 2001) argued that such judgments reflect affective processes that are at least partially independent of perceptual and cognitive systems, that precede cognitive judgments in time, and that are not dependent upon or derived from conscious recognition. These proposals by Zajonc are the topic of some debate, and the parameters that influence judgments of liking are still investigated, but the primary finding of increased preference due to prior exposure is well established (Forgas, 2000; Zajonc, 2001).

Preference judgments have also been used to assess learning in artificial grammar paradigms. These paradigms include a training phase during which subjects study meaningless letter sequences such as "XMVTRX." Participants are unaware that the letter sequences are constructed according to the rules of a finite-state (Markov) grammar. In the most typical testing procedure, they are later told that an underlying system of rules exists, and asked to classify letter strings as grammatical (created by the same finite state grammar) or ungrammatical (not conforming to the patterns produced by the underlying grammar). Participants may profess no conscious awareness of the grammar, but their ability to categorize letter strings is significantly better than chance, even when grammatical patterns are instantiated by entirely different letters (Reber, 1967, 1993; Stadler & Frensch, 1998). The relationship between the mere exposure effect and artificial grammar learning, a connection proposed in Zajonc (1980), was made explicit by Gordon and Holyoak (1983), who first demonstrated that increased preference did not require repetition of identical stimuli, but instead "generalizes to previously unseen stimuli that are similar along certain abstract dimensions to the exposed stimuli" (p. 492). After exposing participants to letter strings generated by a finite state grammar, Gordon and Holyoak asked them to provide both grammaticality judgments and judgments of liking for grammatical and ungrammatical strings. Grammatical strings were rated as more grammatical, but also received higher ratings of liking. Manza, Zizak, and Reber (1998) and Reber (1993) further argue that ratings of liking and ratings of grammatical acceptability are comparable in these paradigms, although ratings of liking have been found to be more sensitive in some circumstances (Manza & Bornstein, 1995) and less sensitive in others (Newell & Bright, 2001). These studies indicate that preference and grammaticality judgments both provide evidence that previously exposed structures have been learned (although not necessarily consciously).

Preference judgments as indirect measures of structural learning have also been productively employed in language acquisition research. With infants as participants, it is more difficult to explain preference for familiar stimuli as deriving from indirect task demands, supporting the view that preferences derive from fluent

processing of representations that overlap at some level. Experiments in this area use acoustic patterns of nonsense words (e.g., *PEL*, *JIK*). Infants first listen to patterns of word strings generated by a finite state grammar and then are tested using a head turn preference procedure (Fernald, 1985). Infants consistently prefer to listen to previously presented patterns as compared to novel patterns of word strings, even when the patterns are composed of different words (Gomez & Gerken, 1999; Marcus, Vijayan, Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996).

There may be multiple causes for increased preference and higher evaluative judgments of familiar stimuli. Zajonc (2001) has suggested that more positive affect toward recently experienced stimuli is independent of inferential processing, and may instead reflect a type of conditioning. This theory may be relevant to infant research on structural learning. In contrast, adult memory researchers have noted that a variety of laboratory judgments about stimuli can be influenced by prior exposure, and argued that such judgments reflect a process of attribution that is not always well-grounded. Numerous studies show that familiarity due to recent exposure can be misattributed to other sources (Bornstein, 1989). For example, Jacoby and colleagues have demonstrated that 24 h after exposure to a series of proper names in the laboratory, participants were likely to classify the familiar (but made-up) names as belonging to famous people (Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989). Other studies show that recently processed stimuli in a variety of modalities are erroneously reported by participants to be brighter, louder, or more likeable, whichever characteristic is queried in the context of the experiment (Jacoby & Dallas, 1981; Jacoby, Kelley, & Dywan, 1989; Kamas & Reder, 1995). In Jacoby's theory, there is no necessary opposition between conscious recognition of stimuli as repeated and such misattributions; instead both recognition judgments and a variety of other judgments can be based on the same fluent processing, if participants are given no motivation to analyze the source of an item's fluency (Jacoby et al., 1989; Johnson, Hashtroudi, & Lindsay, 1993).

The current experiments extend the mere exposure paradigm to grammatical structures in natural language. Participants were asked to evaluate sentences for grammatical acceptability after reading sentences that were similar in wording, similar in structure, or novel in both wording and structure. We argue that grammaticality judgments are a variety of evaluative judgment, much like judgments of likeability, brightness, loudness, etc. (see Luka, in press for extended discussion). If grammaticality judgments are influenced by exposure to related sentence types, not just sentence tokens, it will support the idea that the evaluation of syntactic structure is similar, in at least some respects, to the evaluative processes identified in the mere-exposure and artificial grammar learning literatures (Gordon & Holyoak, 1983). If we were to find consistent and predictable increases in grammaticality judgments for stimuli that are structurally related to previously presented exemplars, it would show that indirect tests of memory can index ongoing learning of a natural language, in much the same way that these tests index newly acquired representations of artificial grammars.

From the standpoint of experimental psycholinguistics, grammaticality judgments are instruments that are sensitive to a wide array of cognitive processes. They have frequently been used to evaluate theories of language comprehension or to examine the interdependence of language comprehension and memory (e.g., Blackwell & Bates, 1995; Ferreira & Henderson, 1991). This use of grammaticality judgments as a dependent measure is quite different from the way formal syntacticians use ratings of grammaticality, although our findings can be seen as relevant for linguistic theory as well (see General discussion). Luka (in press) examines the relationship between syntactic exposure and grammatical acceptability from a more linguistic perspective. For the sake of clarity and ease of exposition in the current context, we use the terms "grammatical" and "acceptable" in accord with standard use, in contrast to the technical sense that these terms evoke for syntacticians.

Overview of the experiments

Below we report five experiments in which participants first read a series of unrelated sentences presented individually (*reading task*). After a short break (5 min performing an unrelated analytic reasoning or arithmetic task) participants rated a set of *grammatical*, *moderately grammatical*, and *ungrammatical* sentences (*rating task*). These stimulus categories were normatively defined by grammaticality ratings previously collected from a different set of participants drawn from the same population. Critical grammatical and moderately grammatical stimulus sentences were counterbalanced across participants to be *familiar* or *novel* during the rating task, depending on which list had been exposed during the reading task. Familiar sentences were either identical or structurally related to those presented during the reading task. We predicted that previous exposure would increase grammatical acceptability. Ungrammatical sentences were excluded from the reading task to avoid calling attention to grammaticality and preserve a fairly natural reading mode. During the rating task, participants provided judgments for sentences that had received normative ratings across the full range of the response scale, from strongly preferred to frankly ungrammatical. Dispreferred but grammatical sentences (that is, sentences with intermediate, unimodal norma-

tive ratings of grammatical acceptability) were included in order to avoid ceiling effects in the rating task, and it was the ratings of these sentences that we expected to be most influenced by prior exposure.

Our first goal was to determine whether mere exposure effects are evident in ratings of grammaticality for identical sentence repetitions. Because Experiment 1 demonstrated facilitation for identical sentences, Experiment 2 examined whether facilitation would be observed for sentences that shared syntactic structure but not content, analogous to the results of artificial grammar learning paradigms. Experiment 2 also investigated the role of positional processing, excluding the possibility that facilitation could be attributed to lexical or semantic properties rather than syntactic structure. Experiment 3 contrasted identical versus structural repetition and also investigated the influence of multiple prior exposures, comparing one, three, and five repetitions of sentences that were either identically repeated or only structurally similar. Experiment 4 compared identical to structural repetition (single exposures) in a within-participant design, and reconfirmed the findings of Experiments 1 and 2. Experiment 5 investigated an alternative definition of multiple structural repetition and provided further evidence regarding the stability of structural facilitation effects in comprehension.

Experiment 1

Method

Stimulus norming

One hundred sixty-six grammatical and ungrammatical sentences were drawn from handbooks of English grammar, textbooks of theoretical syntax, and published linguistic articles, or were adapted from examples in these sources. The sentences exemplifying different types of errors were evenly distributed among two lists (83 sentences per list) with 2 orders of presentation for each list. Forty University of Chicago undergraduates received payment for providing grammaticality ratings of these sentences. The participants rated each sentence using an anchored 7-point rating scale, where the anchors were *very ungrammatical* (1) and *perfectly grammatical* (7). Thirty of the highest-ranked sentences were used as *grammatical* stimuli (mean = 6.28, $SD = 1.1$). Twenty *ungrammatical* sentences were selected from the lowest end of the distribution (mean = 2.23, $SD = 1.4$). Twenty sentences with ratings between 2.90 and 5.55 and with unimodal distributions of acceptability were designated the critical *moderately grammatical* stimuli (mean = 4.17; $SD = 1.4$). These included sentences such as "I miss having time to do anything.", "It's uncertain he'll arrive until after midnight.", and "Sam asked that we leave and he'd give us $10.". To the extent that these moderately grammatical sentences were largely drawn from articles and textbooks illustrating aspects of syntactic theory, it can be argued that the irregularities these sentence embody are structural in nature rather than exemplifying semantic or logical violations (see Appendix A for full set; examples of ungrammatical stimuli are included in Appendix E). Type of syntactic structure, mean rating, median rating, standard deviation, and the distribution of ratings were all considered when assigning sentences to stimulus versions used in the experiments. Normed sentences that had a bimodal response distribution were excluded from the materials.

Participants

The participants were 24 University of Chicago graduate and undergraduate students (13 women) who were native speakers of English. Participants provided informed consent and were paid for their participation.

Materials and procedure

Participants were tested individually and responses were collected using pencil and paper. The procedure had three parts: a reading task, a five minute distractor task (verbal analytic problem solving), and a rating task.

Reading task. Two stimulus lists were created from the critical moderately grammatical sentences, so that half the participants read 10 of these, whereas the other half of the participants read the other 10. Each participant read the same set of 20 highly grammatical sentences. No ungrammatical sentences were included to avoid drawing attention to sentence structure during the reading phase. Each sentence was typed on a 3 in. × 5 in. card and presented by the experimenter at the rate of approximately seven seconds per card. Participants were instructed to read each sentence silently and carefully and to imagine what the sentence was about. Parallel to instructions in artificial grammar learning paradigms, instructions emphasized attending carefully to the stimuli but did not call attention to grammaticality or to subsequent tasks that would require attending to the form of the sentences. There were two presentation orders for each stimulus list.

Rating task. Each participant rated the same set of 60 sentences, in one of two presentation orders, at his or her own pace. The rating set included all 20 moderately grammatical sentences, half repeated from the reading task and half new. These were combined with 40 fillers—20 ungrammatical and 20 highly grammatical—so that the rating list included a broad range of sentence acceptability. Although, the repetition manipulation was of interest only for the moderately grammatical sentences, we wanted to avoid the perception that the critical sentences were special in some way, which might occur if only those sentences re-occurred across the reading and

rating phases. Half of the highly grammatical fillers were thus also repeated from the reading phase. These ratings were expected to be at ceiling, so this design did not investigate novel versus familiar highly grammatical sentences. (The impact of both identical and structural repetition on highly grammatical sentences is addressed in Experiments 3, 4, and 5.)

Each sentence was presented separately on an 8.5 in. × 2.5 in. slip of paper to encourage participants to evaluate only one sentence at a time. The 7-point rating scale and the verbal labels anchoring the scale were provided below each sentence. Participants circled their ratings.

Results and discussion

The data were first analyzed for potential list effects using a repeated-measures ANOVA with grammaticality (grammatical, moderately grammatical, and ungrammatical) as a within-participants factor and reading list (2 levels) as a between-participants factor. In the item analysis, grammaticality was a between-items factor, and reading list was a within-items factor. There was a main effect of grammaticality, as would be expected. Stimuli classified as *grammatical*, *moderately grammatical*, and *ungrammatical* by the normative group received mean ratings of: 6.47 ($SE = 0.09$), 4.21 ($SE = 0.20$), and 2.28 ($SE = 0.19$), respectively; $F_1(2,44) = 334.83$, $p < .0001$; $F_2(2,57) = 512.4$, $p < .0001$. There was no main effect of reading list and no interactions involving this factor (mean rating and *SE* for reading list 1 and reading list 2, grammatical and moderately grammatical, respectively: 6.50(.13), 6.45(.14), 4.31(.23), and 4.10(.35); all Fs < 1.5, *ns*, by participant and by items).

The effect of identical sentence repetition at reading and rating was analyzed by one-way ANOVAs, and a main effect for repetition was found. Moderately grammatical sentences received a mean rating of 4.49 ($SE = 0.18$) when they occurred in the repeated context (when participants had previously encountered the sentence during the reading task), compared to a mean rating of 3.93 ($SE = 0.13$) for the same set of sentences when those sentences occurred in the unexposed (novel) context. This difference is significant by participants and by items, $F_1(1,23) = 15.31$, $p < .001$; $F_2(1,19) = 14.38$, $p = .001$. The comparative increase in ratings for identically repeated sentences is most clearly attributed to the predicted effect of previous exposure.

We interpret the finding of increased grammatical acceptability for familiar sentences as analogous to the finding of preference for familiar stimuli as established in *mere exposure* paradigms. The increased affective response found in mere exposure studies is attributed to evaluative mechanisms that are sensitive to non-reinforced, repeated exposure to the stimuli. In Experiment 1, the stimuli were identical for exposure and rating, and so the increased ratings observed may be attributed to perceptual facilitation for exact instances (Jacoby & Hayman, 1987; Kolers, 1975; Masson, 1984), or conceptual facilitation based on memory for gist (Bransford & Franks, 1971; Sachs, 1967). However, the facilitation cannot be task related because the prior exposure occurred in a task that did not require ratings of any sort. Whether the source of facilitated processing is perceptual or conceptual, fluent processing of repeated sentences influenced the subjective evaluation of the grammatical acceptability of the items.

Experiment 2

In Experiment 2, we examined whether mere exposure effects generalize to structurally related but not identical repetitions. In contrast to mere exposure effects for identical stimulus repetition, artificial grammar learning studies find increases in affective responses to structurally similar letter strings and nonsense words, even when the patterns presented during the training phase are instantiated by different symbols during the testing phase (Manza et al., 1998). In Experiment 2 we extended these research methods to grammatical structures in natural language, asking participants to rate the grammatical structure of sentences that are structurally related to sentences they read earlier.

Three structural variants of each given syntactic type were presented during the reading task, with a fourth variant of each cohort presented for rating. We included multiple structural variants during reading on the assumption that priming effects might accrue over repeated exposures (Hartsuiker et al., 1999; Monahan, Murphy, & Zajonc, 2000), an assumption explicitly tested in Experiment 3.

Experiment 2 used a conservative definition of structural similarity: related sentences varied in content words and morphology, whereas closed class words, including modal and auxiliary verbs, were the same (morphology of the verbs did change to preserve normal verb agreement). This method of stimulus construction preserved structural similarity as defined by superficial phrase structure configurations. For example, the following sentences were created as structural variants: "On the agenda were fourteen items"; "In the corner was a grizzly bear"; "In the closet were three midgets"; "On the field were twelve players." However, noun phrases were defined at an abstract level rather than word-for-word. That is, the sentences had the same structure above the level of the noun phrase but demonstrated a variety of configurations within the noun phrase itself (e.g., determiner-noun, determiner-adjective-noun, numeral-noun, and proper names were all accepted as variations of a noun phrase; postnominal modifiers were not investigated). Thus, the following sentences were also structural variants: "It is difficult for Kate to deci-

pher your handwriting."; "It is impossible for anyone to use the Pentagon's automated answering service."; "It was easy for the detectives to reconstruct the criminal's motives.". It is standard in the literature on syntactic priming in production to define structural similarity at this level of abstraction, but it is open to empirical inquiry whether this same assumption holds for structural facilitation in comprehension.

A second factor we examined in Experiment 2 is the role of positional configurations in structural facilitation. If structural facilitation is due to priming at an abstract level of syntactic representation rather than shared lexical or conceptual relations, then the linear order and hierarchical relationship between syntactic constituents should be critical. In a production task, Hartsuiker et al. (1999) showed priming for word order, finding, for example, that reading and repeating "Above the table hangs a lamp." increased the likelihood that participants would produce a sentence such as "On the shelf stands a book." in contrast to "A book stands on the shelf." in describing a relevant target picture. (These experiments were conducted in Dutch, where the primed sentence structure, a frontal locative, is more common than the comparable English translation.) These findings imply that the linear relationship of the constituents is relevant to priming, because classes of lexical items (for example, verbs having the same thematic roles or argument structures) do not prime the production of positional variants (see also, Pickering et al., 2002). However, comprehension priming of the type examined here may differ from production priming in this regard. For example, the repetition of identical or related lexical items, especially verbs, could activate shared participant role relations, facilitating comprehension in a way that is not related to grammatical encoding strategies. To examine the role of positional priming in comprehension, we must demonstrate that variations in the linear order of constituents do not result in facilitation in the absence of structural similarity.

We tested this proposal by including lexical control sentences in the reading condition of Experiment 2. For example, some participants read "On the agenda were fourteen items." while others read a *positional variant* "Fourteen items were on the agenda." Sentences presented at rating (i.e., "In the corner was a grizzly bear.") have a *structurally similar* relationship for participants who had previously read the first sentence, or share a *positional variant* relationship in the case of participants who read the second sentence. Thus, the prime sentences in the experimental and control conditions were matched for lexical and conceptual content, but are structurally distinct. We predicted increased grammaticality ratings (relative to novel sentences) for structurally similar sentences but not for positional variants.

This contrast primarily controls for priming that might occur due to indirect lexical associations, because structural and positional variants contain the same open class words. Therefore, if priming is not observed in the positional variant condition, it must be the case that lexical and conceptual properties in themselves are not sufficient to induce facilitation.

Examining the potential influence of lexical and conceptual commonality in contrast to structural configuration is an important first step toward investigating the influence of thematic roles in structural facilitation. Research in syntactic priming has established that priming effects in production cannot be attributed to thematic similarities (Bock, 1989, but cf. Chang, Bock, & Goldberg, 2003). In this experiment we do not directly manipulate thematic relationships (that is, the positional variants did not systematically alter the thematic role assignment of arguments), but we begin by isolating the potential effects of lexical and conceptual commonality. This is a relevant concern in light of recent findings of cross-linguistic priming in production (Loebell & Bock, 2003; Hartsuiker, Pickering, & Veltkamp, 2004). Due to the definition and experimental methods of syntactic priming in sentence production, cross-linguistic priming can only be observed for constructions that are similar between the two languages. While cross-linguistic priming cannot be mediated strictly via lexical and conceptual commonalities (the observed effects are structural), the possibility remains that structural priming is moderated either by shared information in the bilingual lexicon or shared conceptual scenarios (such as descriptions object-transfer specifying analogous relationships between discourse participants). Measures of structural facilitation may be more sensitive to lexical and conceptual similarity, such that the comprehension of positional variants will facilitate the evaluation of sentences that are lexically and conceptually related, even though the test sentence employs a different structural configuration. If priming were found for positional variants, it would indicate that structural facilitation is not, in fact, structural and that any facilitation effects in comprehension are most likely unrelated to the phenomenon of syntactic priming in sentence production.

Methods

Design and stimulus construction

Table 1 illustrates the experimental design for one stimulus unit: an experimental group read three sentences with a common structure (*structural similarity sentences*), and then rated sentences with both similar and dissimilar structures. A control group read three sentences expressing the same content in an alternate structure (*positional variants*), and then rated the same sentences as the experimental group. The positional variants comprised the same open-class words as their counterparts and retained the same gist.

Table 1
Example items from Experiment 2

Structurally Similar reading List A	*Positional Variant reading List A*
Amanda carried Fernando the package.	Amanda carried the package to Fernando.
Ramarez passed Santiago the ball.	Ramarez passed the ball to Santiago.
Eli schlepped Daniel some lox and bagels.	Eli schlepped some lox and bagels to Daniel.
Structurally Similar reading List B	*Positional Variant reading List B*
What Mark wanted is to look at your notes.	To look at your notes is what Mark wanted.
What Marilyn advised is to drive to his house.	To drive to his house is what Marilyn advised.
What Kathryn said is to button your shirt.	To button your shirt is what Kathryn said.
Rating, all lists	
Egor lugged Dr. Frankenstein the corpse.	
What the pharmacist recommended is to read the directions.	

For the control group, the commonality between sentences in the reading and rating phases was largely restricted to a global semantic level. For instance, sentences such as "To button your shirt is what Kathryn said." and "What the pharmacist recommended is to read the directions." both express that a specified agent indicated the performance of some action. In contrast, the commonality between read and rated sentences in the experimental group additionally included identical constituent structure.

The goal of an initial norming study was to create quartets of structurally related sentences such that three sentences of the quartet could be presented during the reading task, and the fourth presented for rating. Additionally, each quartet needed a control quartet of sentences matched for lexical and semantic content, but with a different syntactic structure. The other criteria for creating and selecting sentences were: (1) moderate grammaticality ratings in the normative study (4.0–5.5 on the seven-point scale); and (2) a wide range of syntactic structures so that no quartet of sentences would be structurally similar to some other quartet of sentences read by the same participant.

Sixty sets of paired quartets (480 sentences total) were initially created, and split into four lists such that no list contained structurally or semantically related items. Grammaticality ratings were collected from 80 University of Chicago undergraduates, who each rated 120 sentences. From these normative ratings, 40 sets of paired quartets were selected for the main experiment. From each of these units (of eight sentences) one sentence was selected for inclusion in the rating task (mean normed grammaticality of 4.49, $SD = 1.67$), and six were assigned to the reading phase.

From the selected quartets, four reading lists of 60 sentences were formed. Each list included 20 triplets of structurally related sentences. For the first structural similarity list (List A), these were similar in syntactic form to half of the grammatical sentences that would be rated later. Structural-similarity List B contained triplets that were similar in form to the other half of the grammatical sentences that would be rated later. One positional variant list contained sentences expressing the same content as structural-similarity List A; the positional variant List B contained sentences expressing the same content as the structural-similarity List B. All of the control sentences were, however, structurally dissimilar to the sentences that would be rated later. Appendix B includes additional examples of the stimuli.

Participants

Participation was contingent on responses to a screening questionnaire about linguistic experience. Individuals who had learned a second language before primary school or who had taken courses in syntax were excluded. Participants were 48 University of Chicago undergraduates who were native speakers of American English, and who had not participated in the norming study or Experiment 1. Data from one additional participant were excluded due to experimenter error during the reading task. Twelve participants were randomly assigned to each of the four reading lists.

Procedures

As in Experiment 1, participants first completed the reading task, then performed an unrelated task requiring verbal analytic problem solving for five minutes, and finally completed the rating task. One difference from Experiment 1 was that sentence stimuli in the reading and rating tasks were displayed on a Macintosh computer using SuperLab Pro 2.0 software (Cedrus Corporation). Sentences were presented individually, centered on the screen, in one or two lines of text. In contrast to Experiment 1, the reading task was self-paced, and participants were instructed simply to read each sentence carefully and to press the space bar when they were ready for the next sentence. During the rating task, the response scale was placed as a typewritten placard above the numbers 1–7 on the standard keyboard. The order of sentence presentation in both the reading phase and rating phase was individually randomized for each participant.

In both the structural-similarity and positional-variant groups, each participant read three variants of 20 syntactic types (60 sentences total). During the rating task, all participants received an identical set of 60 sentences: the 40 critical sentences and 20 ungrammatical fillers. For participants in the structural similarity group, half of the critical sentences were structurally similar to the reading phase items and half were novel (counterbalanced across participants). For participants in the positional variant group, all of the rated sentences were structurally novel.

Results

Initial ANOVAs investigated potential list effects, independent of the exposure manipulation, by comparing ratings of novel sentences across the experimental and control groups The ANOVAs included factors of Group (assignment to structurally similar versus positional variant condition during the reading task) and Grammaticality (novel moderately grammatical versus ungrammatical sentences). Both groups found the moderate sentences more acceptable, with an overall rating of 5.16 ($SE = 0.11$) as compared to 2.13 ($SE = 0.11$) for the ungrammatical sentences; $F_1(1,44) = 563.8$, $p < .0001$; $F_2(1,38) = 1450.5$, $p < .0001$. Neither the main effect for Group nor the interaction of Group and Grammaticality were significant for participants and items (means and *SE* for structural and positional variant reading conditions, respectively: 3.59(.24), 3.71(.25), significant for items $F_2(1,38) = 7.25$, $p < .02$, but not for participants $F_1(1,44) = 0.43$, $p = .52$. Means and *SE* for structural and positional conditions, respectively, by grammaticality: 5.06(.16), 5.26(.16), 2.11(.16), 2.15(.16), $F_1(1,44) = 0.37$, $p = .54$; $F_2(1,38) = 2.94$, $p < .09$).

For the critical moderately grammatical sentences overall, taking into account the exposure manipulation, participants in the structural similarity group assigned an average rating of 5.06 ($SE = 0.15$) to novel sentences, and a rating of 5.31 ($SE = 0.13$) to sentences that were similar to those they encountered in the reading phase. For these same sets of sentences, ratings from the control participants in the positional variant group differed very little, 5.26 ($SE = 0.14$) versus 5.22 ($SE = 0.15$).

ANOVAs were conducted with Group (assignment to reading condition: structural similarity versus positional variant) and Familiarity (familiar versus novel) as factors. For the structural group, the Familiarity factor indexes structural repetition. For the positional variant group, "familiarity" refers not to structural repetition, but instead contrasts the familiarity or novelty of the rating sentences according to lexical and thematic content of the positional variants that were presented in the reading task. A specific influence of structural similarity would thus appear as an interaction between Group and Familiarity, an interaction significant in both the subject analysis ($F_1(1,46) = 4.54$, $p < .05$) and the item analysis ($F_2(1,39) = 5.22$, $p < .05$). Ratings of the structurally similar sentences were significantly higher for previously encountered structurally similar sentences (familiar mean(SE) = 5.31(.13), novel mean(SE) = 5.06(.15)), but ratings did not differ significantly for participants who had previously read the positional variant sentences that controlled for lexical and thematic content but did not preserve structural similarity with the rating sentences (familiar mean(SE) = 5.22(.19), novel mean(SE) = 5.26(.14)).

Discussion

Experiment 2 tested two hypotheses. First, that participants' ratings of grammatical acceptability would be higher if participants had been exposed to syntactically related sentences during the reading task. Second, that structural facilitation in comprehension cannot be attributed to lexical or thematic influences, but rather is fundamentally dependent upon linear order at an abstract level of constituent structure.

The first hypothesis was supported: increased ratings of grammaticality did generalize to structurally related sentences, a finding analogous to results established in studies of stimulus generalization in mere exposure and artificial grammar learning. These results accord with the proposal that facilitated processing occurs for sentence structures that have been processed recently, and that this ease of processing is attributed to grammatical acceptability.

The second hypothesis was also supported, in that comparable effects were not found in the stimulus-paired condition controlling for lexical and conceptual content. If the increased ratings of grammaticality observed in the structural similarity condition were due to abstract semantic content or shared conceptual structure, the lexically matched sentences presented in the positional variant condition also would have facilitated comprehension of the rating sentences. These results imply that the observed effects of structural facilitation are truly structural in nature. We cannot yet infer that structural facilitation in comprehension depends upon the same mechanisms that elicit structural priming in sentence production paradigms, but the empirical results appear analogous thus far.

It must be noted that the positional variant condition did not provide a direct test of the role of closed-class words in structural facilitation. In some cases, positional variants could be created with almost no change in grammar or morphology (e.g., "In a cave in northern Missouri, under some mossy rocks, Clementine discovered the treasure." versus "Clementine discovered the treasure under some mossy rocks in a cave in northern Missouri."; "If the paper is here, bring it in." versus "Bring in the paper if it is here.") However, because the critical sentences instantiated a wide variety of phrase

structures, it was necessary for some positional variants to include morphological changes or the addition of closed class words (e.g., "HMOs **are** likely **to** advocate the new insurance regulations." versus "It **is** likely **that** HMOs advocate the new insurance regulations."). The role of closed class items in syntactic facilitation is open to further investigation.

These results may be considered analogous to the findings of Hartsuiker et al. (1999) for the priming of word order in sentence production. Hartsuiker et al. (1999) suggest a process of positional encoding (linearization) as a separate stage of grammatical encoding—a mechanism specific to language production. However, Pickering et al. (2002) explain the results of Hartsuiker et al. (1999) with a single-stage model of constituent structure in which structural representations are shared across systems of production and comprehension. This latter model is able to address the structural facilitation results of Experiment 2 while also explaining the basis of syntactic priming in language production, as elaborated in the General discussion.

The stimuli used in Experiment 2 provided an initial exploration of structural facilitation in comprehension using a method derived from indirect tests of memory. In this set of stimuli, a six-sentence cohort was defined by common phrase structure at an abstract rather than word-for-word level of representation (i.e., lexical configurations within a NP or PP did vary). Furthermore, structural configurations differed substantially across the 18 cohort sets, providing a test for facilitation across a wide variety of syntactic structures. In Experiments 3 and 4 we use different stimulus sets in order to extend the definition of structural similarity.

Experiment 3

Experiment 3 examined the impact of number of sentence exposures, for both identical items (*token repetition*) and structurally related items (*type repetition*). One group of participants rated both new sentences and sentences that had occurred once, or had been identically repeated three times or five times in the prior reading phase (token condition). Another group of participants rated the same sentences after reading zero, one, three, or five sentences that were structurally similar but not identical to the test sentences (type condition).

Repeated exposure to particular structures is clearly important in natural language learning, and statistical frequency or probability distributions of syntactic parameters have been assigned an important role in parsing strategies (e.g., Crocker & Brants, 2000; Johnson & Riezler, 2002; Jurafsky, 1996, 2003; Trueswell & Tanenhaus, 1994; Vosse & Kempen, 2000). Determining whether number of exposures yields a graded effect on grammaticality ratings is thus a first step toward forging a link between this laboratory paradigm and the real-world influence of long-term frequency. The impact of cumulative exposure to a given sentence structure has not received much attention in the production priming literature. Standard methods for investigating syntactic priming in production include tasks such as picture description, sentence completion, and cued recall (Bock, 1986; Pickering & Branigan, 1998; Potter & Lombardi, 1998). In these tasks, exposure to critical sentence types (such as double object versus prepositional object, passive versus active) alternate within an experiment. The impact of each prime sentence is measured on the production of the next utterance that could possibly utilize the same structure (the next opportunity to use the primed structure may follow some neutral intervening material, as in Bock & Griffin, 2000). This method does not naturally lend itself to examining the cumulative impact of exposures to similar primes, simply because over the course of the experiment sequences of one prime type (passive) are interleaved with exposures to the comparison structure (active), priming both of the alternative productions. There is also the question of whether alternate structures can be primed equally (whether priming effects are balanced), or if the baseline frequency of one alternative structure is influential (Bock, 1986; Pickering et al., 2002). One production priming study has hinted that a cumulative effect might occur nonetheless, because production of a low-frequency structure (frontal locative in Dutch) increased over the course of the experiment (Hartsuiker et al., 1999; cf. Scheepers, 2003, for relative clause attachment in German). Because the present method uses a variety of sentence structures that are not dichotomous alternatives to one another, it is much easier to manipulate the number of exposures to a given structure within a single block. By analogy with research on the mere exposure effect, it was predicted that acceptability ratings would increase in relation to the number of times a given sentence type is encountered (Monahan et al., 2000).

Methods

Participants

Forty-eight young adults from the University of Chicago (23 women) provided informed consent and were paid for their participation. All were native English speakers who had not taken any courses in syntax. None had participated in the previous experiments. Half of the participants were randomly assigned to the token-repetition group, and half to the type-repetition group. This factor was manipulated between-participants in order to avoid drawing attention to the repetition manipulation in the type-repetition group. (In Experiments 4 and 5, we evaluate whether both type and token repetition effects can be observed in the same session in a within-participant design.)

Materials

Sentences were constructed from those normed in Experiments 1 and 2. These included both those with high normative grammaticality scores (above 6.0 on the 7-point scale), and moderate normative ratings (3.5–5.0 range). Normative acceptability thus serves as a factor in the design, in order to determine whether highly grammatical sentences benefit from repetition in the same way as the moderately grammatical sentences tested in Experiments 1 and 2. We were concerned that ceiling effects might preclude observation of an exposure effect for these items, but Experiment 3 explicitly tests this possibility.

Structurally similar sentence variants were created in sets of six, by changing the content words of the normed item while maintaining surface phrase structure. Selection of verbs for a given cohort was based on similarity of meaning and of argument structure. For example, for the sentence "The president and the manager have outlined their plans," structurally similar alternates consisted of "*NP* and *NP* have presented/proposed/expressed/reviewed/discussed *NP*." A full set of six different verbs could not be established for all cohorts, and so some predicates were repeated. As established by Pickering and Branigan (1998), verb repetition is relevant for syntactic priming in production and this may also be the case for structural facilitation effects. (In Experiment 5 we investigate structural repetition effects when no verbs are repeated across reading and rating stimuli.) Closed-class words were left largely intact, although prepositions and pronouns varied across sentences in a structural cohort, as did the tense and number of auxiliary verbs. For example, variants in one cohort included: (a) "Lester is a better pianist than Janice and better singer than Edna.", (b) "Dante was a better story-teller than Virgil and better poet than Homer.", and (c) "Jackie Chan is a better stunt man than Bruce Willis and better actor than Tom Cruise." Appendix C provides examples of structurally related cohorts.

Procedure

Each participant read 54 sentences during the reading phase. In the Token condition, a reading list consisted of three highly grammatical sentences presented once, three highly grammatical sentences presented three times, three highly grammatical sentences presented five times, as well as moderately grammatical sentences that likewise occurred once, three times, or five times. The rating list included these 18 old items, and 18 novel sentences. Which sentences served as novel versus repeated was counterbalanced across participants. The subsets of sentences occurring once, three times, or five times in the reading list were also counterbalanced across participants, so that there were six Token reading lists altogether. Each reading list was constructed such that no sentence repeated consecutively, and the first and second halves of a reading list contained an equal number of repeated items. Finally, two different orders of each reading list were created, within the constraints above.

Six reading lists for the Type condition paralleled those of the Token condition, except that variants of a sentence type replaced the identical sentence repetitions. For both groups of participants, the reading lists additionally included three buffer sentences at the beginning and the end.

The rating list was the same for all participants, and consisted of 18 highly grammatical sentences (nine identical or similar to the reading phase, nine novel), 18 moderately grammatical (nine identical or similar to the reading phase, nine novel), and 20 ungrammatical sentences. There were four orders for the rating list, balanced across participants. Presentation methods and instructions for the reading and rating tasks were as in Experiment 2, as was the brief filler task between reading and rating.

Results

Mean ratings for all the conditions are shown in Table 2. An initial ANOVA examined the impact of repetition collapsed across number of repetitions, and included Grammaticality (highly versus moderately grammatical sentences) and Familiarity (repeated or similar versus novel) as within-participants factors, and Type/Token (structural or identical repetition) as a between-participants factor. In the item analysis, Grammaticality was a between-item factor while Familiarity and Type/Token repetition were within-item factors. As used here, "Familiarity" refers only to the objective relationship between items in the reading and rating phases, and not to the mental states of the participants. Highly grammatical sentences received higher ratings than moderately grammatical sentences, as expected, with overall means of 6.42 ($SE = 0.05$) and 3.92 ($SE = 0.08$), respectively; $F_1(1,46) = 556.4$, $p < .0001$; $F_2(1,34) = 295.4$, $p < .0001$. More relevant to the goal of the experiment, familiar sentences received generally higher ratings than novel ones, with overall means of 5.28 ($SE = 0.15$) and 5.05 ($SE = 0.17$), respectively; $F_1(1,46) = 12.59$, $p < .001$; $F_2(1,34) = 15.44$, $p < .001$.

The main effect of Familiarity was qualified by an interaction with Grammaticality; $F_1(1,46) = 8.69$, $p = .005$; $F_2(1,34) = 12.6$, $p < .001$. Table 2 shows that ratings of the moderately grammatical sentences were more strongly influenced by familiarity than the highly grammatical sentences. The interaction of Familiarity and Type/Token was also significant; $F_1(1,46) = 9.84$, $p < .005$; $F_2(1,34) = 9.19$, $p < .005$. Fig. 1 shows that structural (Type) repetition was much more effective than identical (Token) repetition.

Table 2
Grammaticality ratings in Experiment 3 (mean and *SE*)

Presentations in reading task	Grammatical	Moderately grammatical	All
Identical (token) repetition			
Zero	6.44 (.07)	3.95 (.12)	5.20 (.09)
One	6.35 (.11)	3.97 (.20)	5.16 (.15)
Three	6.32 (.12)	4.10 (.23)	5.20 (.16)
Five	6.51 (.10)	4.08 (.19)	5.30 (.07)
Mean repeated			5.22 (0.09)
Structural (type) repetition			
Zero	6.36 (.07)	3.46 (.11)	4.91 (.10)
One	6.31 (.13)	4.00 (.19)	5.15 (.15)
Three	6.54 (.12)	4.17 (.21)	5.34 (.16)
Five	6.53 (.10)	4.56 (.21)	5.50 (.15)
Mean repeated			5.33 (0.09)

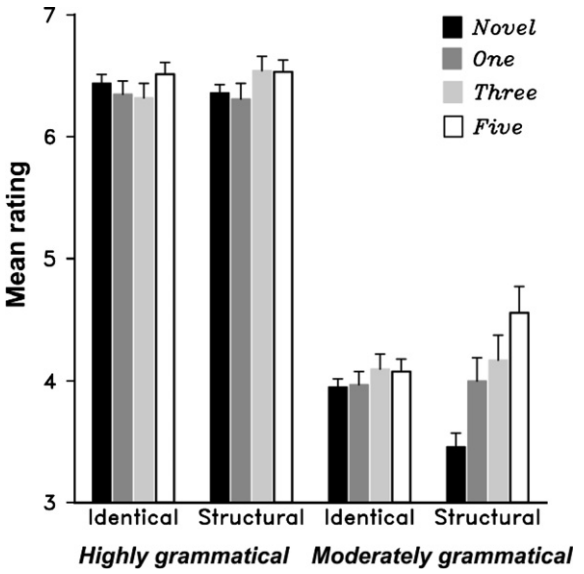

Fig. 1. Mean grammaticality ratings from Experiment 3 (error bars denote *SE*).

Followup analyses pursued these interactions by examining the familiarity effects separately for identical repetition of highly grammatical sentences, structural repetition of moderately grammatical sentences, etc. Replicating Experiment 2, the structural repetition effect was significant for moderately grammatical sentences; $F_1(1,23) = 30.3$, $p < .0001$; $F_2(1,17) = 18.7$, $p < .001$. The structural repetition effect was not evident for highly grammatical sentences. In contrast to Experiment 1, identical repetition was not effective for either level of grammaticality.

A second core analysis examined the impact of number of structural repetitions, for the moderately grammatical sentences. In this analysis, number of repetitions was defined as an ordered factor with levels of one, three, and five, so that a monotonic increase across number of repetitions could be examined. The relevant outcome is the linear trend in a polynomial contrast (F_{linear}). Exposure to multiple variants of a sentence structure boosted ratings of the moderately grammatical sentences in an orderly way, as seen in Table 2; $F_{1\text{-linear}}(1,23) = 6.46$, $p < .02$; $F_{2\text{-linear}}(1,17) = 3.09$, $p = .09$. A final analysis evaluated whether a single related sentence from the reading task was capable of altering subsequent grammaticality judgments. A comparison between moderately grammatical sentences that were new at test versus those preceded by one similar sentence yielded a significant difference; $F_1(1,23) = 8.01$, $p < .02$; $F_2(1,17) = 7.93$, $p < .01$.

Discussion

The design of Experiment 3 included replications of Experiments 1 (identical repetition) and 2 (structural repetition), together with a new manipulation of number of prior exposures. A structural facilitation effect—increased grammaticality ratings of sentences similar in form to those read—was again observed. This effect was restricted to sentences that received moderate grammaticality ratings in normative samples. Sentences that were found highly acceptable on initial exposure (averaging above 6 on a 7-point scale) were immune to further benefit, which may reflect a ceiling effect in the rating scale. For the moderately grammatical sentences, the impact of reading similar sentence structures proved to be cumulative across number of exposures. The ability to examine such cumulative effects is one advantage of the current methodology. As yet unknown are, first, the limit of this incremental benefit, and, second, the temporal window over which exposures to a given structure will continue to increase its perceived grammaticality; these factors are currently under investigation.

Experiment 3 also yielded a puzzling null result. Although, reading sentences that were structurally similar to the rated sentences increased grammaticality ratings, reading the exact same sentences was ineffective.

This null effect is also, of course, in conflict with the outcome of Experiment 1. Differences in presentation method in the reading task are unlikely to explain the discrepancies, however. While Experiment 1 was experimenter-paced using index cards, Experiment 2, like Experiment 3, was subject-paced with computer presentation. Minor differences in instructions are also unlikely to explain the different outcomes, because the instructions used in Experiment 3 were the same as those in Experiment 2.

However, there were a number of differences between the stimulus sets used in Experiments 1 and 2 compared to the Token repetition condition of Experiment 3 which may account for the disparate findings. Most obviously, Experiment 1 included no repetition during the reading task, while Experiment 3 included a set of only 18 sentences, six of which were repeated five times each, and six were repeated three times each. The degree of overt redundancy in the Token condition may have influenced the attention participants devoted to each identical repetition. It is possible that participants did not fully process the repeated items, but instead quickly advanced to the next item after recognizing a sentence. Indeed, reading times decreased from a mean of 1099 ms ($SE = 145$) for initial presentations to only 625 ms ($SE = 53$) for fifth presentations of identical sentences (a 43% decrease). In contrast, the decrease in reading time for structural repetitions was not as dramatic. Mean reading time for initial presentation was 1200 ms ($SE = 162$), while the fifth presentation of a structurally related sentence was 923 ms ($SE = 106$) (a 23% decrease). The greater decrease in reading time for identical repetition argues in favor of recognition rather than a more general speeding of responses. The benefit of identical repetition might require more extensive processing. We address this anomalous result in Experiments 4 and 5, which examine single and multiple exposures of identical and structural repetition in within-participants designs.

Experiment 4

Experiment 4 investigates token and type relationships in a within-participants design. One goal is to determine if sentences identical to those read a single time will receive more fluent processing and higher grammaticality ratings than new items (as in Experiment 1), given the failure to replicate this result in Experiment 3. We believe the odd results of the token repetition condition in Experiment 3 may have occurred because participants did not always re-read entire stimulus sentences once they had been recognized as repetitions during the reading phase. Therefore, in order to encourage attentive and careful reading of each sentence during the reading task, participants were asked to read each sentence aloud while being tape-recorded. A second goal of Experiment 4 was to verify that a structural facilitation effect can be observed after only a single exposure to a syntactically related sentence. Recall that in Experiment 2, participants read three related items before making grammaticality judgments. In Experiment 3, a significant effect was observed after only one exposure, but that observation was based on a small number of trials, and so we attempt to replicate that finding.

Methods

Materials

Construction and norming. Eighty distinct sentence frames were constructed to include a wide variety of construction types and constituent configurations. Each frame served as the structural basis for constructing five related sentences using different content words (only two from each cohort are used in the present experiment). Variants were defined by substituting a verb of the same class according to the taxonomies presented in Levin (1993). Structural identity was defined at the phrasal level, such that a Noun Phrase could be instantiated by a determiner-(adj)-noun combination or a proper noun (i.e., "The farmer peeled the toddler an apple" and "Isaac poured Spencer an arsenic-laced martini" were structural variants). Closed class words such as adverbs, comparatives, determiners, prepositions, pronouns, modal, and auxiliary verbs were repeated across sentences. Open-class words, including common nouns, proper names, numbers, and adjectives were not repeated across any of the sentences. The six sentence variants of each structural frame were distributed across six lists for norming. The same set of 40 ungrammatical sentences was added to each list to provide anchor points of gradient grammaticality across the lists. Each list was randomized and printed with two different presentation orders for each list. One hundred and thirty-two students at the University of Arizona received course credit or payment for providing acceptability judgments for these sentences. Data from an additional 22 of the students were excluded without analysis due to non-native speaker status or concurrent participation in remedial educational assistance programs, as established in an exit survey. Data from two participants were excluded due to non-compliance with instructions. The mean acceptability rating for each critical sentence was therefore established by the average of at least 22 participants' ratings; sentences presented for norming as ungrammatical anchor points were identical across lists and so were established by the average rating of 132 participants.

List construction. From the normed materials, 24 sentences were selected for the Token repetition condition. These included 12 with high grammaticality scores (mean of 5.75, $SE = .10$), and 12 with moderate

grammaticality scores (mean of 4.16, $SE = .15$). Twenty-four pairs of structurally related sentences were selected for the Type repetition condition, also divided into highly and moderately grammatical (mean normative scores of 5.75 and 4.13, $SE = .11$ and .15, respectively). Appendix D shows examples of the structurally related pairs.

All participants rated the same 48 critical sentences for grammaticality: 6 highly grammatical and 6 moderately grammatical sentences repeated identically from the reading phase, 6 highly grammatical and 6 moderately grammatical sentences that were structurally similar to those read, and 24 novel sentences (balanced for grammaticality with the complementary set). These were combined with 24 ungrammatical fillers to anchor the rating scale. Participants were assigned to one of two reading lists preceding the rating phase, so that each sentence in the grammaticality test was novel for half the participants, and related to the reading phase for the other half of the participants. In contrast to the design of the studies reported above, half of the sentences in these reading lists did not have counterparts in the rating list; this subset of the reading list stimuli served as materials for a different experiment. Reading lists included three buffer sentences at both the beginning and the end, which were not analyzed.

Participants

The participants were 26 University of Arizona graduate and undergraduate students (8 men) who were native speakers of English. Data from one additional participant were excluded without analysis because the individual reported concurrent enrollment in a remedial study assistance program for reading comprehension deficits. All participants provided informed consent and were paid for their participation.

Procedure

The experiment was run on a standard PC using DMDX software developed at the University of Arizona by J.C. Forster. Participants were first asked to read each sentence aloud and were tape recorded. Sentences were presented individually, centered on a single screen. Participants were asked to read at a normal rate, with normal intonation, articulating clearly but without exaggeration. They pressed the space bar when they were ready to read the next sentence. Next, participants did a speeded arithmetic response task, answering progressively more challenging addition problems on the computer for 5 min, after which they began the rating task. Sentences for rating were presented individually, centered on a single screen. The rating response scale (1–7) was displayed under each sentence, along with the verbal anchors "Very ungrammatical" (under the number 1) and "Perfectly grammatical" (under the number 7). After the grammaticality test, participants completed a questionnaire to assess their awareness of the relationship between the reading and rating phases.

Results and discussion

Mean grammaticality ratings are shown in Table 3 and Fig. 2. Repeated-measure ANOVAs were conducted with Grammaticality (high vs. moderate), Token/Type (identical vs. structural repetition), and Familiarity (repeated vs. novel) as within-participant factors. It was necessary to exclude one moderately grammatical item from analysis, after a typographical error was discovered in stimulus presentation. Sentences normed as highly grammatical received higher ratings than those normed as moderately grammatical, $F_1(1,25) = 139.7$, $p < .0001$; $F_2(1,43) = 512.4$, $p < .0001$. More importantly, both identical and structural repetition were effective, leading to a main effect of Familiarity unqualified by an interaction with Token/Type; $F_1(1,25) = 13.8$, $p < .001$; $F_2(1,43) = 8.41$, $p < .01$. No significant interactions involving Grammaticality were observed, indicating that repetition from the reading phase was effective for both the highly and moderately grammatical sentences. Fig. 3 shows, however, that the repetition effects were larger for moderately than highly

Table 3
Table of effects across experiments

	Grammatical		Moderately grammatical		All	
	Repeated	Novel	Repeated	Novel	Repeated	Novel
Identical repetition (mean and SE)						
E1	—	—	4.49 (*.18*)	3.39 (*.13*)	4.49 (*.18*)	3.39 (*.13*)
E3	6.39 (*.11*)	6.44 (*.07*)	4.05 (*.18*)	3.95 (*.12*)	5.22 (*.09*)	5.20 (*.09*)
E4	6.50 (*.09*)	6.22 (*.12*)	5.44 (*.13*)	5.05 (*.17*)	5.97 (*.11*)	5.63 (*.13*)
E5	6.45 (*.09*)	6.32 (*.13*)	5.75 (*.23*)	5.40 (*.18*)	6.10 (*.13*)	5.86 (*.13*)
Structural Repetition (mean and standard error)						
E2	—	—	5.31 (*.13*)	5.06 (*.15*)	5.31 (*.13*)	5.06 (*.15*)
E3	6.45 (*.11*)	6.36 (*.07*)	4.20 (*.18*)	3.46 (*.11*)	5.33 (*.09*)	4.91 (*.10*)
E4	6.27 (*.11*)	6.34 (*.10*)	5.04 (*.20*)	4.69 (*.21*)	5.65 (*.14*)	5.51 (*.16*)
E5	6.67 (*.09*)	6.50 (*.11*)	5.53 (*.22*)	5.13 (*.26*)	6.10 (*.15*)	5.81 (*.17*)

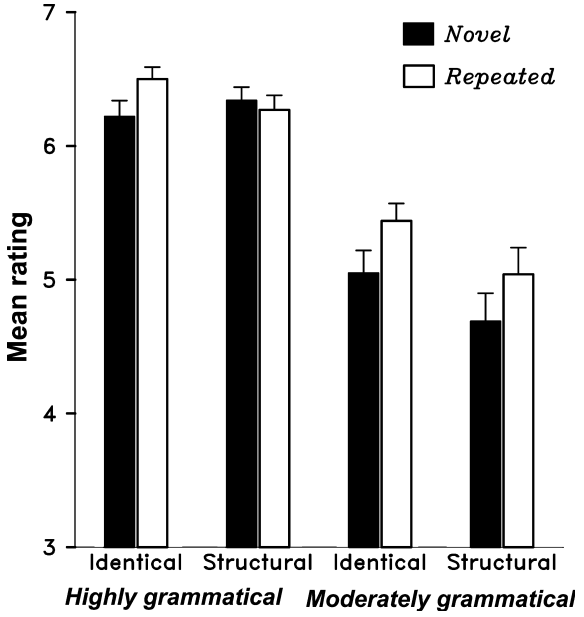

Fig. 2. Mean grammaticality ratings from Experiment 4 (error bars denote *SE*).

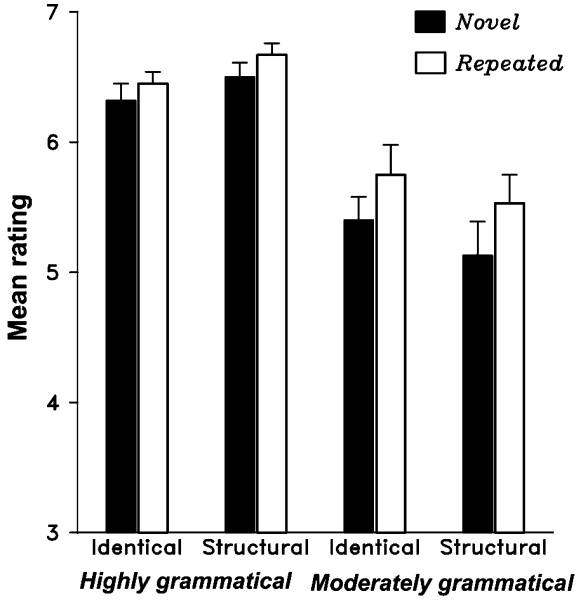

Fig. 3. Mean grammaticality ratings from Experiment 5 (error bars denote *SE*).

grammatical sentences, particularly for structural repetition. Although there were no interactions with Grammaticality, inspection of the means prompted follow-up tests. One way ANOVAs conducted on the means ratings of sentences familiar from the reading phase versus novel confirmed differences in perceived grammaticality, $F_1(1,25) = 10.34$, $p < .005$; $F_2(1,23) = 16.31$, $p < .001$. For structural repetitions, the follow-up contrast was not significant by participants and items, $F_1(1,25) = 0.98$, $p = .3$; $F_2(1,43) = 0.63$, $p = .4$. Thus, while structural facilitation was strong for moderately grammatical sentences, variability in the ratings for highly grammatical sentences reduced the overall effect of facilitation for structurally related sentences. As in Experiment 3, it seems that sentences with high initial ratings have less room for improvement due to prior exposure.

In the questionnaire following the main experiment, participants first answered the general question "Did you notice any relation between the sentences you read aloud during that tape recorded section and the sentences presented for rating?" All responded "yes." The questionnaire then included two questions about identical repetition: (a) "While you were doing the rating task, did you notice whether any of the sentences were repeated from the reading part of the task?" and (b) "You read about 50 sentences in the tape recorded part of the experiment. None, some, or all of these sentences may have appeared again for you to rate. How many of the original 50 sentences were identically repeated (word-for-word) in the rating part of the task?" Two additional questions asked about structural repetition: (c) "None, some, or all of the 50 sentences in the tape recorded task may have been *similar* to sentences in the rating task *without* being identical repetitions. When you were doing the rating task, were you aware of any type of similarity between the recorded sentences and the sentences you rated? Yes/No How would you describe the similarity?", and (d) "How many of the original 50 sentences were similar to sentences you rated (but not identically repeated)?"

The exit questionnaire indicated that nearly all of the participants were aware of the identical repetition manipulation; 22 of 24 responded "yes" when asked if they noticed sentences in the rating task that had occurred in the reading task. When asked how many of the reading-phase items re-occurred in the rating task, the mean estimate was 14.5 ($SE = 1.8$; the correct answer was 12). A smaller majority, 18 of 24 participants, indicated awareness of the structural repetition manipulation; the mean estimate of the number of structurally similar sentences was 11.5 ($SE = 2.0$; correct answer = 12). Not all of the participants chose to respond to the open-ended question asking them to describe any similarity they observed, but two of the responses included: "Grammatical errors were alike in both sections." "Wording of the sentences was in the same pattern." Other responses by participants were at least somewhat ambiguous: "Subject matter was similar.", "Words switched.", "Same type of words.", "Fairly close.", "Word type used in each.", and "General trend." It is unclear to what degree participants may have been guessing. The role of conscious awareness for structural similarity remains open for investigation.

Experiment 4 extends the results of the previous experiments by using a different set of stimuli, and confirms that a single exposure to a structurally related or identical sentence is sufficient to increase perceived grammaticality. The results support the conclusion that the failure to observe structural facilitation in the Token repetition condition of Experiment 3 was likely due to the highly redundant nature of the reading lists used in that experiment, combined with the self-paced silent reading task.

There were two procedural differences between Experiments 3 and 4: excluding repetitive material from the reading task, and asking participants to read aloud rather than silently. The rationale for excluding sentence repetitions from the reading phase was that such repetitions might highlight this manipulation in participants' perceptions of the experiment, and block a change in the perceived grammaticality of highly familiar items. In Experiment 4, however, the exit questionnaire suggested that the repetition manipulation was quite noticeable, indicating that the increase in grammaticality ratings due to prior exposure does not depend on recognition failure. The results cannot comment on the converse possibility, that identical and/or structural repetition effects require conscious recognition of similarities between the reading and rating phases. The reading-phase learning that leads to higher grammaticality ratings may be episodic, or it may be a form of non-episodic, procedural memory, as argued for syntactic priming in production (Bock & Griffin, 2000). Resolving this issue will require a good deal of additional research.

Experiment 5

Experiment 5 was designed to reexamine the role of multiple identical repetitions during the reading phase. If participants re-process identically repeated sentences at even the superficial level required by reading aloud, we predict that a facilitation effect will obtain. In Experiment 5, each sentence in the reading phase occurred four times, which should lead to strong memory traces for individual items. If identically repeated items still receive higher grammaticality ratings than novel sentences, it will more strongly suggest that the null result of Experiment 3 reflected only inadequate re-processing.

Experiment 5 also provides a more closely equated definition of structural and identical repetition, and provides a second perspective regarding the role of conscious recognition of such repetitions. In Experiment 3, multiple examples of structurally related sentences occurred in the reading phase, it may have been possible for participants to notice a structural dimension of variability in the stimuli. However, it is unlikely that participants would retain accurate episodic memory for phrase structure configurations between the reading and rating phase given the delay, the intermediate task, and studies such as Sachs (1967) that demonstrate poor episodic memory for sentence structure. The possibility for conscious awareness of repetition for structurally related sentences was further reduced by the fact that Experiments 2 and 3 used a between-participants design for identical versus structural repetitions. Thus, participants in the identical repetition were much more likely to be influenced by conscious awareness of the repetition than were participants in the structurally similar condition. While the exit interview of Experiment 4 indicated that awareness of similarity or identity between sentence presented at reading and rating does not reduce the effect of structural facilitation, Experiment 4 did not include an overt manipulation of repetition. The design of Experiment 5 was developed specifically to equate more closely the definitions of identical and structural repetition during the reading phase. Experiment 5 includes only Token repetition during the reading phase, with either identical or structurally related sentences presented at test. For example, participants read a single sentence like "Lydia pointed her finger at her naughty cousin." four times at various intervals within the reading phase and then rated either that same sentence or a structurally related sentence like "Portia fluttered her eyelashes at the surly waiter." Experiment 5 investigates this definition of repetition in a within-participants design so that during the rating phase all participants rate sentences that are either identical, structurally related, or novel depending upon the composition of the reading list they had seen earlier.

Participants

Twenty-four University of Arizona graduate and undergraduate students (11 men) who were native speakers of English participated in the experiment. All participants provided informed consent and were paid for their participation.

Materials and procedures

Materials and procedures were identical to Experiment 4 except that the stimulus lists in the reading condition were modified to incorporate four repetitions of each stimulus. Each reading list included 12 grammatical and 12 moderately grammatical sentences. Sentences were presented in a random order for each participant, cycling through the set of 24 stimulus sentences, sampling without replacement, until the entire block had been presented, for a total of four cycles, that is, 96 critical stimulus presentations during the reading phase for each participant. Intervals between blocks of repetition were not demarcated in any way. Two such reading lists were created with different sentences, so that each subsequently rated sentence would be novel for half the par-

ticipants, and similar to a reading-list sentence for the other half of the participants.

All participants rated the same 48 critical sentences: 6 highly grammatical items identically repeated from the reading phase, 6 highly grammatical items that were structurally similar to sentences from the reading phase, 6 moderately grammatical items identically repeated from the reading phase, 6 moderately grammatical sentences structurally similar to those read, 12 novel sentences with high normative grammaticality ratings, and 12 novel sentences with moderate normative grammaticality ratings. These were combined with 24 ungrammatical sentences.

Results and discussion

Mean grammaticality ratings are shown in Table 3 and Fig. 3, and were analyzed as in Experiment 4. It was necessary to exclude one moderately grammatical sentence from analysis after discovering a typographical error in the stimulus presentation list. As usual, sentences with high normative grammaticality ratings were also rated more highly by the experimental participants than those with moderate normative ratings; $F_1(1,23) = 51.9$, $p < .0001$; $F_2(1,43) = 66.1$, $p < .0001$. More importantly, the main effect of Familiarity between read and rated sentences was significant; $F_1(1,23) = 8.41$, $p < .01$; $F_2(1,43) = 8.39$, $p < .01$. Table 3 shows that both identical and structural repetitions were effective in increasing grammaticality ratings, and that highly and moderately grammatical sentences were similarly influenced, so that the main effect of Familiarity was unqualified by interactions with either Type/Token or Grammaticality. One unanticipated interaction was also significant, Type/Token by Grammaticality; $F_1(1,23) = 8.26$, $p < .01$; $F_2(1,43) = 3.2$, $p = .08$. This interaction results from highly grammatical sentences receiving somewhat higher ratings when assigned to the identical repetition condition as compared to when assigned to the structural repetition condition, and the reverse for moderately grammatical sentences. This interaction does not involve the primary manipulation of Familiarity between the reading phase and the rating phase, and its interpretation is unclear.

The increase in rated grammaticality for sentences that were identical or structurally similar to those encountered earlier confirms the results of prior experiments and helps to clarify the one null effect observed earlier—the absence of a token repetition effect in Experiment 3. Although the participants surely recognized the frequent repetition of items in the reading lists, this did not prevent an influence of identical repetition on grammaticality ratings. The results thus argue against the idea that failure to recognize stimuli in the rating phase is a necessary precondition for obtaining identical or structural repetition effects. Instead, the precondition for these phenomena appears to be attentive processing of the repeated or structurally related stimuli on initial exposure.

General discussion

The five experiments reported here investigated structural facilitation in sentence processing with a new methodology that is a hybrid of the mere-exposure paradigm and some studies of artificial grammar learning. Prior research with those methods has demonstrated that more positive evaluation of frequently or recently encountered stimuli is an indirect measure of facilitated processing due to learning. The current experiments extend these methods to natural language stimuli. Individuals who had read grammatical sentences found identical or syntactically similar sentences to be more grammatically acceptable in a subsequent rating task. Across the series of experiments, facilitation effects were almost as strong for structurally related sentences as for identical repetitions.

Four experiments examined grammaticality ratings of sentences identical to those presented in a prior reading task; higher ratings of repeated than novel sentences were observed in three of these (Experiments 1, 4, and 5). The single null effect occurred when the reading phase contained multiple repetitions of most of the items, and no incentive to carefully re-read the sentences (Experiment 3). When the requirement to read aloud was implemented, enhanced ratings of repeated sentences were again observed (Experiment 5).

Four experiments examined the more theoretically interesting case of exposure to sentences that shared phrase structure with the test sentences, but did not share content words. In all four (Experiments 2, 3, 4, and 5), sentences that were syntactically similar to recently read items received higher grammaticality ratings than sentences whose structures had not been recently encountered. The structural facilitation effect was uniformly observed for sentences that were deemed "moderately grammatical" by both normative and experimental groups of participants. These were items receiving mean ratings in the 3.5–5.5 range on a 7-point scale, and consisted of sentences that were grammatical, but would likely be revised by a good writer or editor. Less consistent evidence of exposure-induced facilitation was observed for sentence structures that received very high grammaticality ratings on initial exposure. Significantly, enhanced grammaticality ratings for sentences with high initial mean ratings were observed in Experiment 5 but not Experiment 3. Experiment 4 yielded an exposure-dependent change in rating for identical repetitions, but not structural repetitions. These equivocal results may well reflect ceiling effects in the rating scale. The highly grammatical sentences used in Experiments 4 and 5 had normative mean ratings of 5.75; although the stimuli in Experiment 3 were not independently renormed, they were developed from sentences used in Experiments 1 and 2 that had normative mean ratings of 6.47. An alternative explanation that requires further investi-

gation is that the moderately grammatical sentences used here are simply less frequent structural types and are therefore more sensitive to facilitation. This alternative proposal is analogous to arguments made with respect to syntactic priming in production tasks, namely, that priming effects are stronger for rarer or marked sentence types, such as for passives, double object datives, and locative fronting (Bock, 1986; Pickering et al., 2002).

Priming manipulations of one sort or another have been a very productive methodology across multiple research areas in cognitive psychology, and have been used to determine what attributes of stimuli are retained in memory versus discarded, and what attributes of stimuli are spontaneously processed versus typically ignored. Branigan et al. (1995) state this standard assumption in the context of their research on syntactic priming in production: "If two stimuli are related only along one particular dimension, and the processing of one stimulus affects the processing of the other for reasons attributable to that relationship (i.e., if priming occurs), then we can infer that the cognitive system is sensitive to that dimension, and that it treats the two stimuli as related within that dimension." (p. 491). In the present research, we found altered test performance for sentences similar to those previously encountered at only an abstract level of phrase structure. Experiments 2 through 5 used a wide variety of phrasal instantiations to define similar sentences, so that an NP could have been a pronoun, a proper name, or a determiner-adjective-noun sequence. Because the initial encounter occurred in the natural task of reading (silently or aloud), the shared stimulus attributes were accessed spontaneously rather than forced by an unusual laboratory task. The overlapping representations or processes shared by the reading and rating tasks were syntactic rather than lexical, given that open-class words were replaced in creating structurally similar sentences. Although closed-class words were largely similar between structural primes and matched test sentences, the same closed-class words also occurred in novel sentences (e.g., all sets of sentences would have included common determiners, prepositions, and pronouns). In contrast, the linear order of constituents was critical for the observation of a structural facilitation effect: Experiment 2 found no facilitation when the reading phase sentences were positional variants (same words, different phrase structure) of sentences that did yield facilitation due to shared structure with the rated sentences.

Our stimuli included a wide variety of phrase structure configurations. The benefit of this approach is that it minimizes the likelihood that participants will become sensitive to a single class of sentences or error types. Furthermore, the diversity of sentence structures that can be used as stimuli in an indirect memory paradigm of the current type permits investigation of many definitions of structural similarity, and allows targeted comparisons between minimally contrasting sentence pairs. The structural facilitation paradigm provides a means to test definitions of similarity at concrete and abstract levels of description. While other dependent measures such as reading time or accuracy have limited application for investigating abstract structural representations, structural facilitation provides a robust alternative.

The indirect methods used in the present experiments also reduce the possible interference of task demands. Although, participants may be aware of episodic or structural repetition of stimuli between reading and rating, they have no incentive to increase their ratings of grammatical acceptability for all and only those sentences for which they have declarative memory. Furthermore, recognition memory for the exact surface structure of a sentence is known to be comparatively poor (Sachs, 1967). It remains true, however, that the nature of the learning or memory system responsible for the structural facilitation effect observed here is unknown. Altered preference ratings in mere exposure paradigms are assumed to be independent of conscious memory because they occur after subliminal exposure (Buchner & Brandt, 2003; Zajonc, 2001). Artificial grammar learning paradigms frequently include dissociations between the ability to categorize stimuli as grammatical or ungrammatical and the ability to recognize previously presented strings, but the extent of this dissociation is still investigated and debated (Channon et al., 2002; Knowlton & Squire, 1996; Stadler & Frensch, 1998; Tunney, 2003; Tunney & Shanks, 2003). It will similarly be important to determine the relationship between structural facilitation and recognition memory.

The current results also encourage further investigation of the relationship between structural facilitation in sentence comprehension and syntactic priming in sentence production. While it is possible that that the two are unrelated, the hypotheses of the current research were guided by previous findings in production priming, and yielded compatible results. One concrete comparison that can be made between syntactic priming and structural facilitation is the relative duration of the respective production and comprehension effects. This characteristic is related to the contrast between *priming* as a short-term change in behavior due to recent experience and one criterion of *implicit learning* as a long-lasting change in behavior. Production priming effects are reported to span at least 10 intervening sentences (Bock & Griffin, 2000). For the present experiments including only a single prior exposure to a structurally similar sentence (Experiments 3 and 4), some 40–50 sentences intervened between reading and rating a given item (on average), in addition to the five-minute distractor task interposed between the two phases. The effects clearly reflect more than a transient activation (cf. Branigan, Pickering, & Cleland, 1999; Wheeldon & Smith, 2003), but additional studies are underway to determine their ultimate duration.

The current results are consonant with models of parsing during comprehension, to the extent that such models are sensitive to exposure based strategies (Mitchell, 1994). The results are also consistent with a model of Pickering and colleagues that accounts for priming in comprehension and production. Their model was developed to address comprehension-to-production priming, a finding that suggests a shared level of representation utilized by both comprehension and production (Branigan et al., 2000; Cleland & Pickering, 2003; Hartsuiker et al., 2004). Pickering and colleagues begin with a standard model of speech production in which information regarding the grammatical properties of lexical items, or lemma information, is shared between the comprehension and production systems (Levelt, Roelofs, & Meyer, 1999). To this model, they add a level of representation that specifies lexical category information (e.g., Verb) as well as information about how such categories combine, such as how phrasal categories such as noun phrases (NP) and prepositional phrases (PP) may occur in the context of specific verb such as "give" (e.g., NP,NP or NP,PP; Cleland & Pickering, 2003; Pickering & Branigan, 1998). This model represents combinatorial information of lexical categories that is shared across production and comprehension and thus would be sensitive to a wide variety of structural types, such as those used in the current experiments.

In the present research, grammaticality judgments were used as a tool to investigate structural facilitation. While this research concerns memory and language processing and not grammaticality judgments per se, our findings do have relevance for the use of grammaticality judgments as the basis of linguistic methodology (Levelt, 1974; McCawley, 1996; Schütze, 1996). We observed the strongest exposure effects for the class of sentences normed as in the moderately grammatical range (mean rating of approximately 3.5–5.5 on a 7-point scale of acceptability). Many of these stimulus sentences were taken directly, or derived from, examples in textbooks and articles on linguistic theory. Such sentences receive disproportionate attention in these venues exactly because debates about syntactic theory fundamentally depend upon whether a theorist considers a given sentence grammatical or ungrammatical, and "questionable" sentences are the most informative in distinguishing between theories. The results of these experiments thus have implications for linguistic intuition and its role in theoretical linguistics, which are discussed elsewhere (Luka, in press).

We argue that our research paradigm and our experimental results provide information about the interaction of memory and evaluative processing heuristics in language comprehension. Our method cannot be equated with rate-rerate study designs that have been used in the past to investigate the stability of repeated linguistic judgments over time (e.g., Greenbaum & Quirk, 1970; Nagata, 1988). In studies examining the stability of linguistic judgments, participants are asked to provide ratings of grammatical acceptability the first time they encounter a given stimulus sentence, and their initial ratings are compared to subsequent responses. In some cases, participants may retain explicit memory for their initial responses. Further research is necessary in order to establish differences in sentence comprehension and evaluation in structural facilitation contexts compared to rate-rerate paradigms.

Variability of grammaticality judgments under repeated exposure has also been investigated in the context of *syntactic satiation*, an effect in which "over time, certain types of sentences that were initially judged ungrammatical begin to sound increasingly acceptable." (Snyder, 2000, p. 575) Snyder defines syntactic satiation operationally as an increase in the number of "yes" (grammatical) judgments in the final two test blocks compared to the number of "yes" responses in the initial two test blocks (5 test blocks of 10 sentences each with order of presentation balanced across subjects). This definition is similar to the measure of test–retest reliability used in the studies of Greenbaum and Quirk (1970), except that the change in rating is evaluated over sentence types rather than repeated presentations of the same sentence. Snyder examined seven sentence types and found evidence of syntactic satiation in only two of the sentence types: wh-island (e.g., Who does John wonder whether Mary likes?) and complex-NP sentences (e.g., Who does Mary believe the claim that John likes?). Synder concludes that syntactic satiation is not a phenomenon which affects all sentence types equally. Snyder's results indicate that syntactic satiation is most likely unrelated to the effects we investigate here. Syntactic satiation applies only to sentences that are initially judged to be ungrammatical and is observed in response to only a small number of specific sentence types. In contrast, structural facilitation was observed across a variety of sentence types, including fully grammatical sentences.

While linguists encourage the careful and informed use of acceptability judgments in the development of syntactic theory (Cowart, 1997), it is unclear whether cautious use of acceptability judgments might overcome the tendency to find a sentence more and more acceptable with re-reading. For example, it remains to be tested whether conscious awareness of structural facilitation, such as an instructional manipulation calling participants' attention to repeated or related structures, will inhibit subsequent increases in ratings. A priori, such inhibition would seem to depend on participants accurately recalling all and only those structures which they had seen before.

It is tempting to extrapolate these findings to contexts of natural language use, such as grammatical changes in the dialect of an adult who moves to a new geographical area or who begins to move in new social circles. Such linguistic changes (long term accommodation in grammar) have been observed, and the source of such changes

is attributed to exposure to the new dialect or sociolect (Trudgill, 1986). Implicit learning may account for some aspects of language change (Bock & Griffin, 2000). In addition, for speakers functioning in multilingual social contexts, structural transfer arising from shared structural representations may account for flexibility in language use (codeswitching) or may form the basis for processes of diachronic language change (Loebell & Bock, 2003). However, a variety of sociological factors, as well as social psychological factors such as self concept and group identification, are known to influence whether speakers begin to produce types of grammatical structures to which they are exposed (Bybee, 2003; Thakerar, Giles, & Cheshire, 1982). Therefore, mere exposure to linguistic forms in comprehension need not necessarily result in linear, monotonic increases in rates of production for these structures. While the current research method cannot address sociological and psychological motivations for the production of specific grammatical structures, it may prove possible to examine long-term change in the perceived acceptability of a given structure, which is perhaps a precursor to production.

A brief summary of the current findings is that familiarization with a syntactic structure leads to greater acceptance of similar syntactic structures. This observation is elicited by experimental methods that have defined studies of memory, specifically research paradigms using indirect tests of memory to investigate the processing characteristics of non-declarative systems of memory. Our results emphasize that language comprehension is sensitive to recent processing experiences and the cognitive systems that evaluate grammatical acceptability are influenced by heuristics such as fluency of processing.

Acknowledgments

Financial support for this research was provided by NSF Grant SBR 9421326 to Lawrence W. Barsalou, by the University of Chicago Graduate Student Research Fund, and by a departmental award made available through the contribution of Leslie Shaw. We thank David McNeill, Howard Nusbaum, Jerry Sadock, and the late Jim McCawley for helpful comments on earlier versions of this manuscript, as well as Martin Pickering and two anonymous reviewers.

Appendix A

Moderately grammatical sentence stimuli used in Experiment 1.

We hate to bake pies anymore.
From the CIA, I assure you that I would never accept a penny.
What Christmas shopping means to me is that I walk my feet to pieces.
I miss having any time to do anything.
Sam recites poems as well as playing the piano.
Sam asked that we leave and he'd give us $10.
We want for most people not to catch on.
It's uncertain he'll arrive until after midnight.
What Mark wanted is to look at your notes.
A heart courageous never breathed scant.
My brother was kept tabs on by the FBI.
Which pope's reign was Copernicus born during?
Who did you hire because he said would work hard?
Go see whether the paper's here and bring it in.
You should hire a manager as efficient as the competition has.
You should absolutely come and visit me.
In China, wine is served in small cups, and in Turkey, coffee.
More physicists became farmers than went to law school, and more philosophers did too.
Lester is a better pianist than Janice and better singer than Edna.
Clinton and Dole both promised Dan Rather not to cause each other any trouble.

Appendix B

Example items used in Experiment 2. Sentences in plain font appeared in the reading task of the Structurally Similar condition, while the respective lexically matched sentence in italics were presented in the reading task for the Positional Variant condition. The sentences presented for rating (identical set across conditions) appear in bold.

What Mark wanted is to look at your notes.
To look at your notes is what Mark wanted.
What Marilyn advised is to drive to his house.
To drive to his house is what Marilyn advised.
What Kathryn said is to button your shirt.
To button your shirt is what Kathryn said.
What the pharmacist recommended is to read the directions.

Debbie ought to buy a car as reliable as that fireman had.
That fireman has a reliable car like Debbie ought to buy.
Rachel needs to get a tattoo as colorful as Bob has.
Bob has a colorful tattoo like Rachel needs to get.
I have to get a personal trainer as sexy as Oprah has.
Oprah has a sexy personal trainer like I have to get.
You should hire a manager as efficient as the competition has.

Your low grades were attributed by your mother to laziness and lack of motivation.
Your mother attributed your low grades to laziness and lack of motivation.
The annual deficit was credited by the government to inflation and bipartisan disagreement.
The government credited the annual deficit to inflation and bipartisan disagreement.
The discrepant results were ascribed by the biologists to unsterile lab materials and genetically damaged control groups.
The biologists ascribed their discrepant results to unsterile lab materials and genetically damaged control groups.

The data losses were imputed by Bill Gates to incompatible software and memory limitations.

Bring in the paper if it is here.
If the paper is here, bring it in.
Carry up the baggage if it arrived.
If the baggage arrived, carry it up.
Invite in the delegates if they came.
If the delegates came, invite them in.
Show out these gentlemen if they are finished.

Amanda carried Fernando the package.
Amanda carried the package to Fernando.
Ramarez passed Santiago the ball.
Ramarez passed the ball to Santiago.
Eli schlepped Daniel some lox and bagels.
Eli schlepped some lox and bagels to Daniel.
Egor lugged Dr. Frankenstein the corpse.

Appendix C

Example items for Experiment 3. Each set of sentences was used in the Type (structurally related repetition) reading condition, with either one, three, or five sentences from each set balanced across reading lists. The sentences in bold were presented in the reading lists in the Token (identical repetition) condition, repeated one, three, or five times across reading lists. The sentences in bold were also presented during the rating task for both the Type and Token conditions.

Lester is a better pianist than Janice and better singer than Edna.
Neil is a better cook than Jerry and better kisser than Alex.
Julie is a better programmer than Madeline and better consultant than Leslie.
Dante was a better story-teller than Virgil and better poet than Homer.
Jackie Chan is a better stunt-man than Bruce Willis and better actor than Tom Cruise.
Yitzak is a better composer than Andras and better violinist than Petrov.
Lester is a better pianist than Janice and better singer than Edna.

It is likely that Antoinette is in a convent.
It is probable that the embezzlers are in Uruguay.
It is likely that Kyle is at the movies.
It is plausible that the evidence is in the safety-deposit box.
It is undeniable that the accused was at the scene.
It is likely that Stan is at the bank.

Ruthless lawyers seem to favor the tobacco industry.
Very old doctors seem to support assisted suicide.
HMO's seem to support the new insurance laws.
Current inmates seem to disprefer electroshock therapy.
Unemployed steelworkers seem to distrust Union promises.
Smart businessmen seem to fear the IRS.

Which patient's surgery was Dr. Phillips sneezing during?
Which professor's class did you fall asleep during?
Which Chinese dynasty did Marco Polo live during?
Which Caesar's rule did the Massada incident occur during?
Which Star Trek episode did the VCR jam during?
Which pope's reign was Copernicus born during?

The red and blue species evolved independently.
The yellow and white sapphires are priced differently.
The snow and king crabs were harvested methodically.
The first and second graders were told separately.
These legal and illegal immigrants were pursued relentlessly.
The male and female employees are paid equally.

Appendix D

Example items for Experiments 4 and 5. In Experiment 4, one sentence of a pair was presented at reading and either the same sentence or its structurally similar counterpart was presented for rating (identical or structural repetition, respectively). In Experiment 5, all sentences presented at reading were identically repeated four times. The identical or structural counterpart was then presented at rating.

It was simple for the surgeon to hide the evidence.
It is difficult for Kate to decipher your handwriting.

Bob elbowed his way through the mob.
Sally sipped her way through a quart of vodka.

Swordfish steaks grill rapidly.
Moose briskets roast easily.

The guys saw the picture of themselves.
The Mafia don recognized the recording of himself.

Portia fluttered her eyelashes at the surly waiter.
Lydia pointed her finger at her naughty cousin.

The brewmaster who sprouted the barley selected the hops.
The veterinarian who vaccinated the horse harnessed the dappled mule.

The podiatrist who the boot salesman purposely kicked twisted his ankle.
The electrician who the plumber eventually resuscitated scorched his knee.

There dawned an unlucky day.
There erupted a horrible plague.

Antoinette being in a convent is outrageous.
The hijackers being in Yemen is possible.

The organizers praised the selflessness in the volunteers.
The beat poets mocked the self-righteousness in the authorities.

Austin nodded his thanks.
Angelica winked her approval.

Appendix E

Examples of ungrammatical sentences presented during the rating task.

Either Gail baked a flan or Scott a pie, and Nick some custard.
To treat them as inhuman are cruel.
This exam was a worse piece of cake than the last one.
Someone apparently vanished my wallet.
Franz spilled any of the containers.
Rover, which was growling, scare the mayor.
For that the world is flat to be widely figured is unthinkable.
He lives in a far from the city location.
Here's being a good place to hide the loot was not suspected.
Several has already been tested.
Suddenly the students began the faculty's shouting obscenities.
The ordeal was spared the inmate by the judge.
We distributed the gifts among they.
You should avoid of whatever guy you were drawing caricatures.
After the exams are the time to relax.
His attack on you viciously left me speechless.
Mr. Sahn and me are on the nominating committee.
How the walrus got into the pantry.
At no time did that we were short of cash worry us.
Pete will have at what Alex was staring.
Smith's maying reject our offer anguishes me.
Dayna quietly not heard away.

References

Blackwell, A., & Bates, E (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience, 7*(2), 1228–1257.

Bock, J. K (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31*(2), 1163–1186.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning. *Journal of Experimental Psychology: General, 129*(2), 177–192.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*(1), 1–39.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin, 106*(2), 265–289.

Branigan, H. P., Pickering, M. J., Liversedge, S. P., Stewart, A. J., & Urbach, T. P (1995). Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research, 24*(6), 489–506.

Branigan, H. P., Pickering, M. J., & Cleland, A. A (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review, 6*(4), 635–640.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), B13–B25.

Bransford, J. D., & Franks, J. J (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*(4), 1331–1350.

Buchner, A., & Brandt, M. (2003). Further evidence for systematic reliability differences between explicit and implicit memory tests. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56*(2), 193–209.

Bybee, J. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2, pp. 145–167). Mahwah, NJ: Erlbaum.

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research, 29*(2), 217–229.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition, 90*(1), 29–49.

Channon, S., Shanks, D., Johnstone, T., Vakili, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia, 40*(12), 2185–2197.

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*(2), 214–230.

Corley, M., & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an Internet-based study. *Psychonomic Bulletin & Review, 9*(1), 126–131.

Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage Publications.

Crocker, M. W., & Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research, 29*(6), 647–669.

Fernald, A (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.

Ferreira, F., & Henderson, J. M (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30*(6), 1725–1745.

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language, 48*(2), 379–398.

Fodor, J. A., & Bever, T. G (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 414–420.

Forgas, J. P. (2000). *Feeling and thinking: The role of affect in social cognition*. New York: Cambridge University Press.

Frazier, L., Taft, L., Roeper, T., Clifton, C., & Ehrlich, K (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition, 12*(5), 421–430.

Gomez, R. L., & Gerken, L (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 1109–1135.

Gordon, P. C., & Holyoak, K. J (1983). Implicit learning and generalization of the "mere exposure" effect. *Journal of Personality and Social Psychology, 45*(3), 492–500.

Greenbaum, S., & Quirk, R. (1970). *Elicitation experiments in English: Linguistic studies in use and attitude*. Coral Gables, FL: University of Miami Press.

Hartsuiker, R. J., & Kolk, H. H. J (1998). Syntactic persistence in Dutch. *Language and Speech, 41*(2), 143–184.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75*(2), B27–B39.

Hartsuiker, R. J., Kolk, H. H. J., & Huiskamp, P. (1999). Priming word order in sentence production. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 52*(1), 129–147.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish–English bilinguals. *Psychological Science, 15*(6), 409–414.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306–340.

Jacoby, L. L., & Hayman, C. A (1987). Specific visual transfer in word identification. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*(3), 456–463.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*(3), 326–338.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In E. Tulving, H. L. Roediger, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Lawrence Erlbaum.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*(2), 115–125.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28.

Johnson, M., & Riezler, S. (2002). Statistical models of syntax learning and use. *Cognitive Science, 26*(3), 239–253.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science, 20*(2), 137–194.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 39–95). Cambridge, MA: MIT Press.

Kamas, E. N., & Reder, L. M. (1995). The role of familiarity in cognitive processing. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 177–202). Hillsdale, NJ: Lawrence Erlbaum.

Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General, 133*(3), 450–467.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning Memory, and Cognition, 22*(1), 169–181.

Kolers, P. A (1975). Specificity of operations in sentence recognition. *Cognitive Psychology, 7*(3), 289–306.

Kunst-Wilson, W. R., & Zajonc, R. B (1980). Affective discrimination of stimuli that cannot be recognized. *Science, 207*, 557–558.

Levelt, W. J. (1974). *Formal grammars in linguistics and psycholinguistics: III. Psycholinguistic applications*. The Hague, Netherlands: Mouton.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences, 22*(1), 1–75.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*(5), 791–824.

Luka, B. J. (in press). A cognitively plausible model of linguistic intuition. In S. S. Mufwene, E. Francis, & R. Wheeler (Eds.), Polymorphous linguistics: Jim McCawley's legacy. Cambridge MA: MIT Press.

Manza, L., & Bornstein, R. F. (1995). Affective discrimination and the implicit learning process. *Consciousness and Cognition: An International Journal, 4*(4), 399–409.

Manza, L., Zizak, D., & Reber, A. S. (1998). Artificial grammar learning and the mere exposure effect. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 201–222). Thousand Oaks, CA: Sage Publications.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77–80.

Masson, M. E. (1984). Memory for the surface structure of sentences: Remembering with and without awareness. *Journal of Verbal Learning and Verbal Behavior, 23*(5), 1579–1592.

McCawley, J. D. (1996). Acceptability judgments in the teaching and doing of syntax. In L. McNair, K. Singer, L. M. Dobrin, & M. M. AuCoin (Eds.), *CLS 32: The parasession on theory and data in linguistics* (pp. 119–132). Chicago: Chicago Linguistic Society.

Mehler, J., & Carey, P. (1967). Role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior, 6*(3), 335–338.

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 375–409). New York: Academic Press.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24*(6), 469–488.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science, 11*(1), 462–466.

Nagata, H. (1988). The relativity of linguistic intuition: The effect of repetition on grammaticality judgments. *Journal of Psycholinguistic Research, 17*(1), 1–17.

Newell, B. R., & Bright, J. E. H. (2001). The relationship between the structural mere exposure effect and the implicit learning process. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A*(4), 1087–1104.

Noppeney, U., & Price, C. J. (2004). An fMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience, 16*(4), 702–713.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*(4), 633–651.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46*(3), 586–605.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38*, 265–282.

Rayner, K., & Clifton, C. J. (2002). Language processing. In H. Pashler & D. Medin (Eds.), Steven's handbook of experimental psychology (Vol. 2, 3rd ed., pp. 261–316) Memory and cognitive processes. New York: Wiley.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 1855–1863.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics, 2*(9), 1437–1442.

Saffran, J. R., Aslin, R. N., & Newport, E. L (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(3), 501–518.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition, 89*(3), 179–205.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production: An online study. *Cognition, 78*(2), 123–164.

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguistic Inquiry, 31*, 575–582.

Stadler, M. A., & Frensch, P. A. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage Publications.

Thakerar, J. N., Giles, H., & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. In C. Fraser & K. R. Scherer (Eds.), *Advances in the social psychology of language* (pp. 205–255). Cambridge: Cambridge University Press.

Trudgill, P. (1986). *Dialects in contact*. New York: Blackwell.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. J. Clifton & L. Frazier (Eds.), *Perspectives on sentence processing* (pp. 155–179). Mahwah, NJ: Lawrence Erlbaum.

Tunney, R. J. (2003). Implicit and explicit knowledge decay at different rates: A dissociation between priming and recognition in artificial grammar learning. *Experimental Psychology, 50*(2), 124–130.

Tunney, R. J., & Shanks, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory & Cognition, 31*(7), 1060–1071.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and lexicalist grammar. *Cognition, 75*(2), 105–143.

Wheeldon, L. R., & Smith, M. C. (2003). Phrase structure priming: A short-lived effect. *Language and Cognitive Processes, 18*(4), 431–442.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1–27.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*(6), 224–229.