Cognitive Psychology 58 (2009) 250–271

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Experience and sentence processing: Statistical learning and relative clause comprehension

Justine B. Wells [a], Morten H. Christiansen [b], David S. Race [c], Daniel J. Acheson [a], Maryellen C. MacDonald [a,*]

[a] *Department of Psychology, University of Wisconsin-Madison, Madison, WI 53706, USA*
[b] *Department of Psychology, Cornell University, Ithaca, NY 14853, USA*
[c] *Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627-0268, USA*

A R T I C L E I N F O

Article history:
Available online 14 October 2008

Keywords:
Sentence processing
Relative clauses
Individual differences
Working memory
Statistical learning

A B S T R A C T

Many explanations of the difficulties associated with interpreting object relative clauses appeal to the demands that object relatives make on working memory. MacDonald and Christiansen [MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54] pointed to variations in reading experience as a source of differences, arguing that the unique word order of object relatives makes their processing more difficult and more sensitive to the effects of previous experience than the processing of subject relatives. This hypothesis was tested in a large-scale study manipulating reading experiences of adults over several weeks. The group receiving relative clause experience increased reading speeds for object relatives more than for subject relatives, whereas a control experience group did not. The reading time data were compared to performance of a computational model given different amounts of experience. The results support claims for experience-based individual differences and an important role for statistical learning in sentence comprehension processes.

1. Introduction

George Miller's (1956) landmark description of the nature of short term memory was a characterization of both its limits (7 ± 2 units) and the modulation of these limits through learning, in that the

* Corresponding author. Fax: +1 608 262 4029.
E-mail address: mcmacdonald@wisc.edu (M.C. MacDonald).

doi:10.1016/j.cogpsych.2008.08.002

units were *chunks*, the size of which could change through a person's experience with the material being processed. In discussions of computational capacity since that time, different research paradigms have tended to vary in their attention to the claim of capacity limits vs. the claim that capacity changes through learning. For example, within adult sentence comprehension, many accounts have invoked capacity limits to explain people's difficulties in relative clause comprehension (e.g., Gibson, 1998; Just & Carpenter, 1992; Lewis, Vasishth, & VanDyke, 2006). All of these accounts have noted that experience could affect processing abilities, but the focus in these accounts has been on showing how a characterization of capacity limits explains certain aspects of sentence comprehension performance, rather than on investigating how capacity (and presumably performance) could change with learning over time.

In another sub-field of sentence processing, constraint-based accounts of ambiguity resolution have invoked comprehenders' detailed knowledge of the language (knowledge of verb biases, discourse plausibility, etc.) in explaining differences in processing difficulty across sentences (see MacDonald & Seidenberg, 2006, for review), but here too there has been little formal attempt to address the learning that must underlie the acquisition and use of the probabilistic constraints hypothesized to shape comprehension performance. This paper begins to address the lack of research on the role of learning in sentence processing. We take relative clauses as our domain of investigation, following prior work (MacDonald & Christiansen, 2002) in which we investigated experience-based changes in relative clause processing in a computational model.

Relative clauses are a good choice for investigating the role of learning in part because these structures have been central to sentence processing research since Miller and Chomsky (1963) observed that certain types are very difficult to comprehend. Examples such as (1) are termed *subject relatives* because the noun modified by the relative clause (the head noun), *reporter*, is the subject of the relative clause verb *attacked*. Reordering some of the words in (1) yields example (2); sentences of this type are termed *object relatives* because the head noun *reporter* is the object of the verb *attacked*. The contrast between the relatively easy subject relatives and the much more difficult object relatives has formed the basis for investigations in virtually every area of psycholinguistics, including studies of memory use in language comprehension (Gibson, 1998; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; Lewis et al., 2006), impairment after brain injury (Dickey & Thompson, 2004), comprehension changes in cognitive aging (Wingfield, Peelle, & Grossman, 2003), typical and atypical child language development (Booth, MacWhinney, & Harasaki, 2000; Friedmann & Novogrodsky, 2007; Kidd, Brandt, Lieven, & Tomasello, 2007), and individual differences in adults Just and Carpenter (1992) and King and Just (1991).

1. Subject relative clause: The reporter that attacked the senator admitted the error.
2. Object relative clause: The reporter that the senator attacked admitted the error.

Despite this extensive use of relative clause materials in language research, there is remarkably little agreement about what makes object relatives harder than subject relatives in English. Some researchers have suggested that object relatives are harder because the sentence's meaning is more complicated in object relatives than in subject relatives, in that the head noun (such as *reporter* in the object relative (2)) is simultaneously the object of the relative clause verb *attacked* and the subject of the main clause verb *admitted*, whereas the head noun is the subject of both verbs in subject relatives (Bever, 1970; MacWhinney & Pléh, 1988). Other researchers have pointed to processing difficulty as the sentence unfolds. Gibson and colleagues (Gibson, 1998; Grodner & Gibson, 2005; Warren & Gibson, 2002) have emphasized the *locality* of the thematic role assignments and the working memory burden of maintaining noun phrases in memory before they can be assigned thematic roles and integrated into the sentence. The word order of object relatives requires longer maintenance of unanalyzed noun phrases than in subject relatives. Others have suggested that these unintegrated noun phrases in the object relative clause interfere with each other in working memory (Gordon, Hendrick, & Johnson, 2004; Gordon et al., 2001; Lewis & Nakayama, 2002; Lewis et al., 2006).

A few accounts have emphasized the role of experience in object relative processing. Gennari and MacDonald (2008) argued that object relatives are much easier with inanimate head nouns as in (3) than with animate head nouns as in (2) because the animacy information is relevant to resolving ambiguities in the object relative sentences (see also Gordon et al., 2001; Mak, Vonk, & Schriefers, 2002; Traxler, Morris, & Seely, 2002; Warren & Gibson, 2002). Gennari and MacDonald linked the processes of object relative interpretation to constraint-based ambiguity resolution, in which comprehenders use information concerning the most likely interpretation to guide the interpretation of temporary ambiguities in object relative clause structures. On this view, comprehenders are exposed to the distributional patterns of noun animacy that tend to occur in relative clauses (which are themselves shaped by various constraints on the language production system, Gennari & MacDonald, submitted for publication), they encode these regularities via statistical learning, and they use this knowledge to guide their interpretation of new input. Similarly, Reali and Christiansen (2007a, 2007b) showed that comprehenders' knowledge of typical pronoun usage patterns in relative clauses predicted processing difficulty.

3. Object relative, inanimate head: The article that the senator attacked was retracted.

To date, the most explicit description of statistical learning and ambiguity resolution in relative clauses comes from the claims and computational modeling in MacDonald and Christiansen (2002), who examined effects of learning about relative clause structures themselves, independent of animacy or other lexical content. They argued that statistical learning would have different effects on comprehension of subject and object relative clauses in English and drew an analogy between sentence comprehension processes and the *Frequency* × *Regularity* interaction in word recognition (e.g., Seidenberg, 1985). This phenomenon refers to the ambiguity in English spelling patterns, such as the fact that the letter sequence int sometimes is pronounced /ɪnt/ (as in *mint*) and at other times is pronounced /aynt/ (as in *pint*). The effect of this ambiguity on reading speed and accuracy varies with both a word's frequency in the language and the number of other words that share the same spelling-sound correspondence (its regularity or consistency). Specifically, reading is faster and more accurate for regular words (ones with many "neighbors" with the same spelling-sound relationship, as in *hint, mint, lint, dint*, etc.) than for irregular ones like *pint*. Moreover, this effect of regularity is much larger for low frequency words than for high frequency words in the language. Regular words receive a benefit from the neighbors, which provide practice on similar spelling-sound relationships, so that computing the pronunciation of the rare regular word *dint*, for example, is affected not only by experience with *dint* itself, but also by experience with similar words *hint, mint, sprint*, etc. Irregular words such as *pint*, however, have few neighbors with similar spelling-sound computations, and thus ease of computing pronunciations for these words is strongly dependent on specific experience with the words themselves. Thus very common irregular words, such as *have*, are read quickly and on par with regular words of similar frequency, while rare irregular words are read more slowly than regularly spelled words in the same frequency range.

Several researchers have suggested that sentence ambiguity resolution also has aspects of the frequency × regularity interaction, that certain sentence types are more "regular" than others, meaning that there is a more consistent mapping between their surface form (the word order) and their meaning. On this view, sentence processing should also exhibit Frequency × Regularity interactions, such that interpretation of less regular sentence types (those with idiosyncratic syntax-meaning mappings) depend heavily on specific experience (frequency) of that exact structure (Juliano & Tanenhaus, 1993; Pearlmutter & MacDonald, 1995). MacDonald and Christiansen suggested that object relatives were an example of irregular sentences, while subject relatives are more regular, in that they adhere closely to the overwhelmingly frequent S–V–O (subject–verb–object) word order in English. Moreover, the pattern of thematic role assignments also follows the overwhelmingly common English pattern of the immediately preverbal noun being the agent of the action and the post-verbal noun being the patient. Readers' processing of subject relatives thus benefits from their many encounters with sentence "neighbors": simple transitive sentences that share the S–V–O word order and thematic role assignments of subject relatives. Object relatives, however, follow an irregular O–S–V word order and patient-agent order of thematic role assignment that are extremely rare in English. Object relatives thus have essentially no sentence "neighbors" in terms of word order and thematic role assignments. As a result, processing of object relatives benefits almost exclusively from direct experience with object relatives themselves, which are far lower in frequency than simple transitive sentences in English (Rohde, 2002; Roland, Dick, & Elman, 2007).

MacDonald and Christiansen used the Frequency × Regularity analogy to reinterpret previous research suggesting that working memory capacity limits comprehenders' interpretation of object relative clauses. King and Just (1991) and Just and Carpenter (1992) argued that individual differences in speed and accuracy of comprehending object relative clauses was predicted by performance on Daneman and Carpenter's Reading Span task, that those who scored high on this measure could comprehend object relatives faster or better than those with poorer reading span scores. Whereas King and Just (1991) and Just and Carpenter (1992) emphasized capacity limits in this case, arguing that the working memory capacity of lower span readers was not sufficient to process object relatives sufficiently, MacDonald and Christiansen emphasized modulation of performance through learning. Their argument was again inspired by results in the word recognition domain. Seidenberg (1985) had suggested that the exact nature of the frequency by regularity interaction might vary across individuals. He examined readers with different levels of reading skill and argued that highly skilled readers had a large band of irregular words that could be read as quickly as regular words. By contrast, the less skilled readers read irregular words more slowly than regular words for all but the highest frequency irregulars. Seidenberg attributed this effect to differential effects of reading experience on regular vs. irregular forms—effectively a Frequency × Regularity × Experience interaction. The highly skilled group, who presumably read more than the low skill group, had more experience with both regular and irregular words, but this extra experience was much more helpful for the irregular words than the regular ones.

MacDonald and Christiansen argued that King and Just (1991) relative clause results could also be an example of a Frequency × Regularity × Experience interaction. They suggested that people who scored well on the reading span task were those who read more and thus had more experience than those with lower reading span scores; this extra experience was hypothesized to be both the source of the high-span group's good reading span performance and their better comprehension of object relatives. More specifically, variations in reading experience were hypothesized to change the nature of people's Frequency × Regularity interaction for relative clauses. Amount of reading experience was predicted to have little influence on subject relative processing, because even inexperienced readers have encountered the regular S–V–O word order often enough in simple transitive sentences to be fairly adept at its processing. However, reduced experience should impair the processing of object relatives, because the latter's irregular O–S–V word order makes processing them highly dependent upon direct experience with object relatives themselves.

MacDonald and Christiansen tested these hypotheses through computational simulations in which connectionist networks were provided with differing amounts of experience on corpora generated by a probabilistic context-free grammar and a small vocabulary. The corpora included simple intransitive S–V sentences, simple S–V–O transitive sentences, and sentences with subject or object relatives, some with multiple embeddings. Importantly, subject and object relatives occurred with equal probability, each accounting for about 2.5% of the sentences in the experience corpora. MacDonald and Christiansen assessed network performance after different amounts of experience by measuring processing accuracy for novel test sentences. They predicted that subject relative processing would be facilitated by extensive exposure to simple transitive sentences in the corpora, which shared much of the S–V–O word order of subject relatives, and thus amount of experience would have little effect on subject relative processing. In contrast, the processing of irregular object relatives was predicted to be largely dependent on exposure to object relatives themselves, so that there should be a large effect of experience on object relative processing. These predictions were confirmed, lending support to the claim that the King and Just (1991) effects of reading span could have arisen from variation in reading experience.

These modeling results are suggestive, but they do not provide a direct test of the role of experience on relative clause comprehension. The current study therefore aims to investigate whether manipulations of people's experience with relative clause constructions will result in the same experience-based patterns of performance suggested by the model. We designed a study to manipulate readers' experience with relative clause constructions in four experimental sessions spaced over 3–4 weeks. Half of the participants were assigned to a Relative Clause experience condition and received reading experience with an equal number of subject and object relative clauses. The other half, termed the Control experience group, received experience with other complex sentences. All participants were administered the reading span task twice. The first administration was prior to the experience manipulation and was done in part to assure that the Relative Clause and Control experience groups were matched on reading span, a known correlate of verbal ability (Daneman & Carpenter, 1980). The second administration followed the experience manipulation and assessed the effect of experience on reading span scores. Reading performance on sentences involving subject and object relative clauses was also assessed before and after the experience manipulations, using a self-paced reading task similar to that used by King and Just. We predicted that the reading patterns of Relative Clause experience participants would resemble those of MacDonald and Christiansen's SRN models—namely, equivalent exposure to subject and object relatives would facilitate reading times on object far more than on subject relatives. In other words, we predicted that readers would initially show a strong effect of relative clause type in reading times, but that the size of this effect would diminish after exposure to an equal number of object and subject relative sentences—a Testing Session (pre- vs. post-test) × Relative Clause Type interaction in reading times in critical regions. No such result is expected for the control experience group, so that across both groups of subjects, we expect an Experience Group × Testing Session × Relative Clause Type interaction. These results would support claims for the importance of reading experience—permitting statistical learning about key properties of relative clauses—in adult relative clause processing.

2. Methods

2.1. Participants

Participants were 97 undergraduates from the University of Wisconsin-Madison. They received either monetary compensation or course credit for their participation in four testing sessions. All were native speakers of English.

2.2. Materials

2.2.1. Reading span task

Because each participant performed the reading span task twice (once in Session 1 and once in Session 4), two lists were constructed, each composed of 70 unrelated sentences ranging in length from 11 to 17 words. None of the 70 sentence-final words was repeated within a list. Seventy-two of the sentences were taken from Daneman and Carpenter (1980) reading span task. The remaining 68 sentences were constructed in a narrative style similar to that of the Daneman and Carpenter sentences. All sentences were ordered randomly and then assigned to two lists, each with 36 Daneman and Carpenter sentences and 34 new sentences. The assignment of lists to testing session was counterbalanced across participants. Each sentence was printed in a single line on an 8.5″ × 5.5″ index card. The cards were arranged in sets of increasing numbers of sentences. The first five sets contained two sentences each, which were followed by sets of three, four, and five distinct sentences. Blank cards were at the end of each set to signal the participant's recall of all sentence-final words within that set.

2.2.2. Pre-test and post-test for relative clause reading

For the self-paced reading pre- and post-tests, 40 pairs of subject and object relative sentences were constructed. All the words in a given subject relative/object relative pair remained the same; the only difference was the word order, as in "The clerk that trained the typist told the truth about the missing files." (subject relative version) and "The clerk that the typist trained told the truth about the missing files." (object relative version). The head noun phrase and object relative clause always comprised six words and had the form *The noun that the noun verbed*, while the subject relatives always had the form *The noun that verbed the noun*. The main verb (e.g., *told*) immediately followed the end of the relative clause, such that it was always the 7th word in the sentence. Material after the main verb continued plausibly and contained four or more additional words.

All relative clause sentences were constructed to have minimal pragmatic bias in noun–verb relationships, as the absence of plausibility information increases the difficulty of these structures (King & Just, 1991). For the clerk/typist example above, the assumption was that a clerk is no more likely to train a typist than a typist to train a clerk, and that a clerk is no more likely than a typist to tell the truth about missing files. Eight sentence pairs were adapted from materials in King and Just's no-bias condition, and the other 32 were developed anew. None of the nouns or verbs in the experimental materials was repeated in any other experimental or filler sentence.

Beyond the absence of plausibility information, the lexical properties of the experimental sentences were controlled in several respects. First, both the head noun and the noun in the relative clause were always animate, typically job descriptions such as *clerk* and *typist*. Head nouns are typically animate for subject relatives but inanimate for object relatives (Gennari & MacDonald, submitted for publication; Roland et al., 2007). We know of no animacy statistics for the other noun in the relative clause, but it is likely that the noun tends to be animate in object relatives (where it is the agent of the action) and inanimate in subject relatives (where it is the patient/theme). Second, no pronouns were used, though they are common in relative clauses, particularly in object relatives (Reali & Christiansen, 2007a). Third the relative pronoun was always *that*, though other choices (e.g., *who*, or the omission of the relative pronoun entirely) are also common in natural language (Jaeger, 2005; Race & MacDonald, 2003; Roland et al., 2007). These lexical choices, while certainly attested in natural language relative clauses, are not typically the most common pairings of word choice and structure. However, they represent the most common instantiations of both subject and object relative clause materials in many other studies (e.g., some or all conditions in Gennari & MacDonald, 2008, submitted for publication; Gordon et al., 2001; 2004; King & Just, 1991; Race & MacDonald, 2003; Traxler et al., 2002; Warren & Gibson, 2002), and thus these choices make the materials most comparable to existing research.

Eighty fillers without relative clauses were constructed to be generally similar in length and syntactic complexity to the experimental stimuli. Half of the fillers were assigned to the pre-test and half to the post-test. Each set of 40 fillers included 19 sentences with multiple prepositions (*The bush by the cemetery tower with steep stairs was pruned by the groundskeeper*) and 21 sentential complements (*The cooks gossiped that the manager flirted with everyone to amuse herself while working at the diner*).

Yes/no comprehension questions were constructed for each experimental and filler sentence. As in King and Just (1991), half of the comprehension questions for the experimental items interrogated the main clause, and the other half interrogated the embedded clause. An equal number of questions interrogating each clause type had "yes" answers.

Two pre-test scripts were constructed, each containing 10 subject relative sentences, 10 object relative sentences, and 40 filler sentences. Two post-test scripts were also constructed, with 10 subject relative sentences, 10 object relative sentences, and 40 filler sentences. A given participant was exposed to one pre-test script in Session 1 and one post-test script in Session 4. The assignment of filler items to pre- or post-test list was random; all participants saw the same 40 fillers in the pre-test and 40 different fillers in the post-test. Assignment of relative clause sentences to pre- and post-test lists was counterbalanced so that every participant read exactly one version of each of the 40 subject/object relative clause stimulus pairs.

2.2.3. Materials for the experience manipulation

Two sets of stimuli were developed to manipulate participants' reading experience, one set for the Relative Clause experience group and one for the Control experience group. The stimuli for the Relative Clause group consisted of 80 subject relative sentences (e.g., *The amateur golfer that had beaten many of the pros won the celebrated state championship.*), 80 object relative sentences (e.g., *The actor's daughter that the Italian ambassador met last year loved Sicilian food.*), and 80 complex filler sentences without relative clauses. All relative clause sentences were adapted from sentences in the *Wall Street Journal* and Brown corpora. They were modified where necessary to make them comprehensible when removed from their original contexts, to have animate common noun heads and relative clause nouns, to have the relative pronoun *that*, and to replace any nouns or verbs that were contained in the experimental sentences in the pre- and post-tests. Compared to the pre- and post-test stimuli, the subject and object relative experience sentences were longer and more variable, with many noun and verb phrase modifiers as well as other pragmatic cues. In no case did a relative clause sentence in the experience set have the structure of the tightly controlled sequences in subject and object relatives in the pre- and post-tests (*The noun that {verbed the noun/the noun verbed} verbed...*). Thus if participants in the Relative Clause experience group learn something about relative clauses from their exposure to additional sentences, this learning cannot be attributed to expectations about simple noun and verb adjacencies.

The stimuli for the control experience group consisted of 80 sentential complement sentences (e.g., *The organizers estimated that more than 100,000 people attended the peace rally last year*), 80 conjoined sentences (e.g., *The amateur golfer had beaten many of the pros and even won the celebrated championship*), and the same 80 fillers as seen by the Relative Clause group. The sentential complement sentences were taken from the *Wall Street Journal* and Brown corpora and were chosen to be roughly the same length as relative clause stimuli; they were modified as necessary in the same manner as for the relative clause sentences. The conjoined experience sentences were adapted from materials in the Relative Clause experience condition. Eighty of the relative clause experience items were restructured into conjoined sentences such that each of the 80 conjoined experience sentences was closely matched on topic to a relative clause experience sentence. For example, the conjoined sentence, *The amateur golfer had beaten many of the pros and even won the celebrated championship*, was created from a subject relative experience sentence, *The amateur golfer that had beaten many of the pros won the celebrated state championship*. Forty conjoined sentences were derived from subject relative experience sentences, and 40 were created from object relative experience sentences. Thus the Relative Clause and control experience materials differed greatly in sentence structure, but there was substantial overlap in the topics and content words across the two sets of materials.

Across both experience sets, the subject nouns of all sentences (including fillers) were animate, mostly referring to human occupations or roles. Though there was no overlap of subject nouns and verbs from experience to pre- and post-test materials, there was some repetition of the subject nouns and verbs within the experience stimuli. Where there were repetitions, each experience group saw each noun or verb approximately the same number of times.

To reduce potential effects of practice with the Yes/No question structure of the pre- and post-test stimuli, a different comprehension probe structure was used for the experience stimuli. After each experience sentence, two statements appeared, and participants selected which of two statements was true according to the information in the sentence. Answer options for relative clause sentences were paired such that one statement concerned the event associated with the main verb of the sentence, and the other concerned the event associated with the embedded verb. The correct statement addressed the embedded clause half the time and the main clause the other half.

To approximate the generally high difficulty for the relative clause comprehension probes, the probes for sentential complement and conjoined sentences were made as difficult as possible on a case-by-case basis, without requiring excessive inferencing about the situation described in the sentence. The correct answer option for these stimuli included the first critical verb roughly as often as it included the second one. The comprehension probes for the 80 filler sentences (seen by all participants) were also designed to be as difficult as possible on a case-by-case basis.

Five scripts were constructed for each experience group, one for each of five experience blocks to be presented over the first three experimental sessions. For the Relative Clause experience group, each script included 16 object relative sentences, 16 subject relative sentences, and 16 fillers. For the control group, each script contained 16 conjoined sentences, 16 sentential complements, and 16 fillers.

2.3. Procedure

2.3.1. Schedule of tasks and sessions

Each participant was tested in four 30- to 60-min sessions, which were spaced four to eight days apart, scheduled at the participant's convenience. The spacing of the sessions was designed to assess non-immediate effects of experience rather than immediate priming from materials in the experience manipulation to the items in the post-test. Most notably, the final post-test assessing effects of experience was conducted in Session 4, which was always scheduled at least four days after the last experience blocks in Session 3.

In Session 1, participants performed the reading span task, read sentences in the pre-test, and then were assigned to either the Relative Clause or control experience group. Assignment to experience groups was semi-random, with the only constraint being an attempt to balance the initial reading span scores in the two experience groups. Following the group assignment, participants completed one block of experience sentences in their assigned experience condition.

Participants completed two experience blocks each in Sessions 2 and 3, along with some paper-and-pencil measures from an unrelated experiment. In Session 4, participants completed the self-paced reading post-test, followed by a second administration of the reading span task. In sum, two experience groups completed the identical set of tasks and differed only in the set of stimuli to which they were exposed during the experience blocks in Sessions 1, 2, and 3. In total, the Relative Clause experience group read 160 relative clause sentences during the experience sessions (half each of subject and object relatives), while the control group read 160 sentences containing other complex structures.

2.3.2. Reading span

In the reading span task, conducted in both Session 1 and Session 4, participants were presented with individual sentences on large index cards to read aloud, followed by a cue to recall the sentence-final words of the sentences in the current set. Participants were instructed to begin to read a sentence aloud as soon as it was placed in front of them, and they were also told that they should not look to the end of the sentence to get a preview of the to-be-remembered word. Cards were turned over and set in front of the participant immediately after he or she had finished reading the previous sentence, allowing minimal rehearsal time between sentences. The blank card between sets served as a recall cue to recall the sentence-final words.

After practice with two sets of two sentences each, participants completed five sets of three sentences. They then were presented with four-sentence, five-sentence, and six-sentence sets until they failed all five sets at a given level. Participants were alerted whenever they were moving to sets with an increased number of sentences, and they were repeatedly encouraged to try hard, in an attempt to motivate all participants equally.

The highest level at which a participant correctly recalled all the sentence-final words of at least three out of five sets constituted a participant's reading span score. Half credit was added for any higher level at which the participant correctly recalled two out of five sets. For example, if a participant was correct on three out of five three-sentence sets and two out of five four-sentence sets, a reading span of 3.5 was assigned. Only one half-credit point could be earned.

2.3.3. Reading pre- and post-tests

Participants performed the self-paced reading pre-test in Session 1 and the post-test in Session 4. The procedure was identical in both sessions. Materials were presented on a computer screen using a word-by-word, self-paced moving window display (Just, Carpenter, & Woolley 1982). At the beginning of each trial, a series of dashes appeared on the computer screen, each dash representing a nonspace character in the sentence. When a participant pressed the space bar, the first group of dashes was replaced by the first word of the sentence. Each subsequent keypress caused the next word to appear and the previous word to return to dashes. When participants pressed the space bar following the sentence-final word, a yes/no question about the sentence appeared. Participants pressed keys labeled "Yes" or "No" to answer the question and received feedback on accuracy.

At the beginning of the task, participants received instructions that encouraged them to read quickly while maintaining good comprehension. After the ten practice trials, a script of experimental trials (10 object relative sentences, 10 subject relative sentences, and 40 fillers) was presented in a different random order for each participant. Assignment of experimental items to pre- or post-test was counterbalanced across participants.

2.3.4. Experience

Each participant completed one experience block in Session 1, two in Session 2, and two in Session 3. The procedure for each block was identical and was designed to expose participants to certain sentence types without replicating the reading and comprehension probe tasks used in the pre- and post-tests. Each experience trial began with a whole sentence appearing on the computer screen (in contrast to the single-word presentation in pre- and post-tests). After reading the sentence, participants pressed the space bar, and two statements appeared one above the other. Participants were instructed to select the statement that was true according to the information in the sentence. Participants pressed a key labeled "Top" to select the top statement and a key labeled "Bottom" to select the bottom statement. Participants received feedback about the accuracy of their answer.

Each experience block began with instructions followed by four practice sentences. A script of 48 trials was then run, with the sentences appearing in a different random order for each participant.

3. Results

For the following analyses, four participants (all in the Relative Clause experience group) were excluded due to experimenter or equipment error in some task, and ten (six Control experience and four Relative Clause experience) were excluded because they failed to return for all four experimental sessions. We also excluded those participants whose mean comprehension accuracy across experimental items and fillers was 75% or below on either the pre- or the post-test, removing data for 19 participants (9 in the control experience group and 10 in the Relative Clause experience group). These rates are somewhat higher than comprehension exclusion rates in other studies of complex sentences. This higher rate may be attributable in part to our efforts to avoid any plausibility information in the pre- and post-test relative clauses, which substantially increased comprehension difficulty (as in King & Just, 1991). A second source of participant loss reflects the difficulty of maintaining participants' interest and cooperation over four experiment sessions. While many participants completed all tasks as instructed, a subset of them expressed annoyance in Session 4 upon being presented with the longer single-word self-paced reading task in the post-test after several short sessions with whole-sentence reading tasks, and a few of these participants rushed through the post-test sentences without reading carefully and answered "yes" to every comprehension question without regard to actual sentence content. By retaining only those participants who had comprehension accuracy above 75% on both pre- and post-tests, we were able to focus on participants who appeared to have conscientiously performed all tasks in this lengthy study.

Following these exclusions, there were 32 control experience and 32 Relative Clause experience participants remaining. The two groups were well-matched on reading span and other measures, as shown in Table 1. Included in this table is information about some participants' ACT scores. The ACT is a standardized test taken by some college-bound high school seniors instead of or in addition to the SAT. Participants gave permission to access their student records, and ACT scores were available for 46 of the participants; these data also suggest that the groups were well matched.

3.1. Self-paced reading times

All analyses of reading times for the pre- and post-tests included only those trials on which the comprehension question was answered correctly. Before analysis, RTs that were greater than 2000 ms were removed (0.05% of the data) and the raw reading times were transformed into

Table 1
Mean scores and (standard deviations) of the two experience groups on various parameters recorded in session 1

Experience group	M:F ratio	ACT*	Pre-test reading span	Unadjusted word reading time on pre-test** (ms)	Pre-test comprehension accuracy** (%)
Relative clause	13:19	28.4 (2.8)	3.30 (0.83)	391 (103.2)	83 (0.13)
Control	11:21	28.2 (2.8)	3.26 (0.86)	390 (101.2)	82 (0.14)

* The ACT is a standardized test, comparable to the SAT, often taken by college-bound high school students. The ACT is scored on a 36-point scale, and the n for this measure was smaller ($n = 46$) because not all participants had ACT scores on file with UW-Madison.

** Includes experimental and filler stimuli.

length-adjusted reading times by calculating a regression equation for each participant to predict that particular participant's reading time for each word length (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). We calculated each participant's regression over all sentences except practice items across both pre-test and post-test, so that the length-adjusted times reflect changes in reading speed from pre-test to post-test. All length-adjusted reading times that were more than 2.5 *SD* from the mean reading time at each session, word position and sentence type were replaced with the cutoff value, affecting 2.6% of the data.

Participants' length-adjusted reading times on relative clause sentences at the pre- and post-tests are shown in Fig. 1, and unadjusted times are presented in Appendix A. The data were grouped into regions in the figures and in analyses using the same regions that King and Just (1991) and MacDonald and Christiansen (2002) had previously used. The first region contained four words: the head noun, the relative pronoun *that*, and the next two words of the relative clause. In the subject relative condition, this region therefore ended with a determiner, as in *senator that attacked the*, whereas in the object relative condition, the region ended with the embedded subject noun, as in *senator that the reporter*. Region 2 contained one word, the embedded object noun in the subject relative condition and the embedded verb in the object relative condition. The third and fourth regions were identical for both sentence types. Region 3 contained the main verb, and Region 4 contained the next two words of the sentence.

Examination of the figure shows large effects of Session, in that all participants read more quickly on the post-test (solid lines) than on the pre-test (dashed lines). There was also a main effect of Session for the filler sentences (not shown in the figure); length-adjusted reading times for fillers were a reliable 71 ms per word shorter on the post-test compared to the pre-test (*p*s < .001). This pattern is expected from participants' increased familiarity with the equipment and task.

The omnibus ANOVA for the Experience Group × Session × Relative Clause Type × Region interaction was reliable ($F_1(3,186) = 3.33$, $p < .05$, $F_2(3,117) = 4.73$, $p < .01$), indicating that the different types of experience did affect reading times in different ways. An examination of the effects at each region of the sentence revealed the predicted three-way interaction of Experience Group × Session × Relative Clause Type at Region 3, the main verb ($F_1(1,62) = 4.22$, $p < .05$; $F_2(1,39) = 3.93$, $p = .05$) but not at the other regions, *F*s < 1. The effect at the main verb was as follows. In the pre-test, both groups showed reliably longer reading times for Object Relatives than Subject Relatives at the main verb (*p*s < .001). The difference was numerically larger (68 ms) for the Relative Clause experience group than for the Control group (34 ms), but the interaction of Group and Sentence type was not reliable

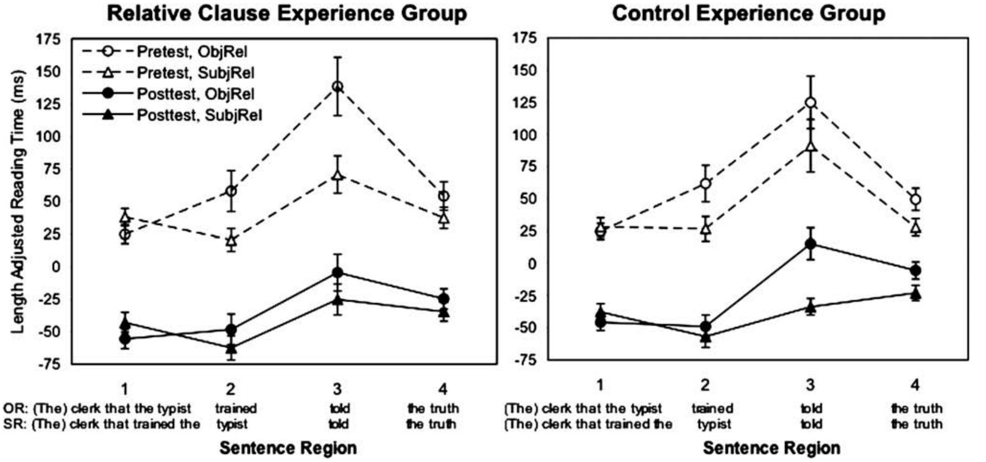

Fig. 1. Length-adjusted self-paced reading times for subject and object relative clauses in the Relative Clause Experience and Control Experience Groups.

($F_1(1,62) = 1.77$, $p > .15$; $F_2(1, 39) = 2.05$, $p > .15$). The pattern across pre- vs. post-test showed that the Relative Clause experience group benefited from their experience, while the control experience group did not: the control experience group had little change in the difference in reading time between object and subject relatives in the pre-test (34 ms) and the post-test (49 ms, a numerical increase), yielding no interaction between Session and Relative Clause Type, *F*s < 1. As with the pre-test, the effect of Relative Clause type was reliable for this group in the post-test, $F_1(1,31) = 24.62$, $p < .001$; $F_2(1,39) = 20.64$, $p < .001$. By contrast, the effect of relative clause type in the Relative Clause experience group was 68 ms on the pre-test but 21 ms in the post-test, yielding a reliable Session × Relative Clause Type interaction, $F_1(1,62) = 4.00$, $p = .05$; $F_2(1,39) = 4.92$, $p < .05$. The 21 ms effect for the Relative Clause experience group at the main verb was not a reliable difference ($F_1(1,31) = 2.56$, $p > .1$; $F_2(1,39) = 3.05$, $p > .08$), but we are not claiming that the Relative Clause experience completely removed the difference in difficulty between the relative clause types. Our prediction was that Relative Clause experience would reduce the difficulty of object relatives compared with subject relatives, and these analyses show that this prediction is supported at the main verb.

An important issue for interpreting effects of experience is how well-matched the two groups were at the pre-test. Examination of the dashed lines in Fig. 1 shows some variation in the patterns of pre-test reading times in the two groups. Across the sentence as a whole, the two groups were extremely well matched. There was a main effect of Relative Clause type in pre-test reading times (a 24 ms per word difference), $F_1(1,62) = 18.80$, $p < .001$; $F_2(1,39) = 41.00$, $p < .001$, and no interaction with Experience group, *F*s < 1. However, there was a small interaction between Relative Clause type, Experience Group, and Sentence Region, which was reliable only in the items analysis, $F_1(3,186) = 1.55$, $p > .20$; $F_2(3,117) = 2.87$, $p < .05$. This result suggests that the reading time contrast in the pre-test between subject and object relatives is distributed somewhat differently across the two experience groups. Because of this variation, because there is some concern about comparing word positions with different word types in subject and object relative clauses (e.g., Region 2 contains a noun in subject relatives and a verb in object relatives), and because different words in one region may create different spillover effects to a subsequent region such as Region 3 (the main verb in both conditions) (Mitchell, 1984; Vasishth & Lewis, 2006), we also examined the effect of experience averaged across Regions 2–4. These positions are natural ones to investigate because the main verb (Region 3) and the words just preceding and following it are the typical sites of reading time differences in subject and object relatives (e.g., in results reported by Gordon et al., 2001; King & Just, 1991; Traxler et al., 2002).

The three-way interaction of Experience Group × Relative Clause Type × Session was not reliable in the mean of Regions 2–4, $F_1(1,62) = 1.63$, n.s.; $F_2 < 1$. This result is not surprising, as this portion of the sentence contains two regions where the interaction was absent (Regions 2 and 4) one where it was present (Region 3). Given our hypotheses of a particular pattern of changes from pre-test to post-test, we pursued additional analyses as planned comparisons. First, in the pre-test, the effect of Relative Clause Type was robust (*p*s < .001) and similar in the two groups, i.e., there was no Relative Clause Type × Experience Group interaction, *F*s < 1. The Control Experience group had a 30-ms per word difference in the pre-test in regions 2–4 and a 25 ms per word difference in the post-test, yielding no interaction, $F_1 < 1$; $F_2(1,39) = 2.86$, $p > .10$. By contrast, the Relative Clause Experience group's reading times yielded a reliable interaction of Relative Clause type and Session; they had a 41-ms per word difference between subject and object relatives in the pre-test, reducing to a 15-ms per word difference in the post-test, $F_1(1, 31) = 4.11$, $p = .05$; $F_2(1,39) = 6.91$, $p < .05$. This remaining difference at post-test between subject and object relatives was reliable (marginal in the items analysis), $F_1(1,31) = 6.98$, $p = .05$; $F_2(1,39) = 3.33$, $p < .08$. Again, our claim is not that the relative clause experience in this study would remove all subject–object relative differences but rather that it would reduce those differences. These analyses show that this claim is supported when considering a larger region than the main verb.

3.2. Accuracy

Participants' accuracy rates on comprehension questions for the subject relatives, object relatives, and filler sentences in the pre- and post-test are shown in Table 2. For the experimental items, both groups had similar accuracy rates across the pre- and post-tests. The interaction of Experience

Table 2
Mean proportion correct and (standard deviations) on pre- and post-test comprehension questions

Experience group	Pre-test			Post-test		
	Subject relatives	Object relatives	Fillers	Subject relatives	Object relatives	Fillers
Relative clause	.84 (0.11)	.73 (0.16)	.89 (0.05)	.82 (0.12)	.76 (0.15)	.93 (0.05)
Control	.83 (0.14)	.73 (0.16)	.89 (0.05)	.79 (0.16)	.78 (0.16)	.92 (0.06)

Group × Session × Relative Clause Type was not significant (*F*s < 1). There was a main effect of Relative Clause Type ($F_1(1,62) = 19.91$, $p < .001$; $F_2(1,39) = 8.89$, $p < .01$), such that participants had better comprehension accuracy on subject than object relatives across Session and Experience group. Thus participants' overall shorter reading times on the post-test, and in particular the Relative Clause experience group's improved reading times on object relatives, did not result in lower accuracy on these items.

Accuracy rates did show an interaction of Session and Relative Clause Type, such that participants improved in their accuracy on object relative questions and got slightly worse on subject relative questions from pre-test to post-test, ($F_1(1,62) = 6.02$, $p < .05$; $F_2(1,39) = 7.72$, $p < .01$), This effect did not interact with Experience Group (*F*s < 1.2, *p*s > .20), and it is not clear why this result obtained.

The two groups improved by 3–4 percentage points in answering comprehension questions for the fillers from the pre-test to the post-test. This difference resulted in a main effect of Session for the filler items that was reliable in only the subjects analysis ($F_1(1,62) = 19.7$, $p < .001$; $F_2 < 1$). Note that filler items were randomly assigned to either the pre-test or post-test, and this assignment was not counterbalanced. Thus Session was a between-items factor in the filler analyses, and these slight differences in performance could result from overall improvements with practice and/or slight differences in the difficulty in the questions of the fillers assigned to the two tests. There was no Session × Experience Group interaction for the fillers (*F*s < 1), suggesting that the two types of experience did not differentially affect performance on the filler sentences.

3.3. *Reading span*

In Session 1, the mean reading span scores of the Relative Clause experience group (span of 3.30, *SD* = .83) and the Control group (3.26, *SD* = .86) did not differ. This outcome is expected because participants were assigned to Relative Clause and Control groups in a way that balanced the span scores in each group. Participants' Session 1 span scores correlated reliably with their scores on the second reading span test, administered in Session 4, $r = .46$, $F(1,62) = 16.64$, $p < .0001$. This fairly modest test–retest relationship is consistent with findings in previous studies; MacDonald, Almor, Henderson, Kempler, and Andersen (2001) reported a correlation of .54 between two Reading Span administrations about two weeks apart, and Waters and Caplan (1996) reported a correlation of .41 between two administrations separated by a somewhat longer interval. The general trend in the current study was for participants to improve their score from the first test (3.28, *SD* = .85) to the second test (3.58, *SD* = .96), $F(1,62) = 7.36$, $p < .01$. This effect was somewhat larger in the Relative Clause experience group (Session 4 span 3.79, *SD* = .98) than in the Control group (Session 4 span 3.38, *SD* = .91), resulting in a marginal interaction of Experience group and Session, $F(1,62) = 3.13$, $p < .09$. Given the relatively modest test–retest reliability of the task, it is unclear how to interpret these small differences in improvement in the two groups.

4. Discussion

The reading time results show a strong effect of experience in processing relative clauses. Importantly, the effect was asymmetrical, in that equal amounts of experience with subject and object relatives had a greater effect on the object relatives than on the subject relatives, reducing differences in reading times between the two conditions from pre-test to post-test. This pattern is consistent with the experience-based predictions in our account and is notable for its generality in several respects. First, the effects of experience were obtained after quite a minimal amount of exposure, reading only 160 relative clauses mixed with fillers of varying sentence constructions. Second, the effects obtained despite several days' interval between the last exposure to relative clauses or control sentences in the experience manipulation and the post-test. This result means that the effects cannot be attributed to short term priming of a syntactic structure or of any other linguistic element. Third, there was no overlap in either nouns or verbs from the experience sentences to the pre- and post-tests, so the effects cannot be attributed to lexical priming effects of content words. Fourth, the relative clause sentences in the experience manipulation were longer and more varied than the pre- and post-test items and never contained the exact word order of the relative clauses in the tests. Thus the effects cannot be attributed to learning how to handle the strict word sequence (*The noun that the noun verbed verbed*...) of the pre- and post-test object relatives. Fifth, the effects obtained despite the fact that in the experience manipulation, both the reading task (whole sentence reading) and the comprehension check (selecting a correct statement from two alternatives) differed from the reading and comprehension tasks in the pre- and post-tests. Thus the benefit for object relatives in the post-test cannot be attributed to learning strategies related to the self-paced reading task. Finally, these effects of experience obtained despite the fact that participants were never given an explicit task of learning anything; they were simply told to read sentences and respond to comprehension probes.

These results therefore appear to reflect powerful implicit learning of properties of the object relatives that are not driven by specific words or adjacent sequences of word types. To this point we have not discussed exactly what *is* learned as a function of reading these sentences, a topic to which we turn in the general discussion. First, we consider the implications for one of the issues that motivated the study, the Frequency x Regularity interaction in sentence processing and MacDonald and Christiansen (2002) computational simulations that yielded greater improvements on object relative clauses than on subject relatives in the face of equal training on both.

4.1. Relative clause experience and the Frequency x Regularity interaction

In their computational simulations, MacDonald and Christiansen (2002) investigated whether a Frequency × Regularity interaction in relative clause processing would emerge as a consequence of experience. They trained simple recurrent networks (SRNs; Elman, 1990) on corpora primarily consisting of simple sentences but which also included equal amounts of subject and object relative clauses. SRNs incorporate recurrent connections that allow past internal states to affect future processing, thus enabling these networks to process hierarchically organized sequential material such as natural language sentences. Prior work has indicated that such networks can learn to process subject and object relatives (Christiansen & Chater, 1999; Elman, 1991; Weckerly & Elman, 1992), and MacDonald and Christiansen predicted that differences in exposure would affect objective relatives more than subject relatives—even if exposure to the two types of relative clauses was equated.

The SRNs were trained on corpora generated by a probabilistic context-free grammar covering the minimal fragment of English grammar needed for exploring the Frequency × Regularity interaction in relative clause processing. This grammar fragment included subject noun–verb agreement, two verb tenses, variations in verb argument structure, and multiple relative clause embeddings with complex agreement structure. Sentences were generated using a very small 30-word vocabulary, including one determiner (e.g., *the*, though words were coded with single units and thus had no real semantics), one complementizer (*that*), singular and plural nouns (e.g., *lawyer, lawyers*), and present and past tense verbs that were obligatorily transitive (e.g., *praises, praise, praised*), obligatorily intransitive (e.g., *hesitates, hesitate, hesitated*) or used in both transitive and intransitive contexts (e.g., *phones, phone, phoned*). Sentences with present tense verbs employed number agreement between the subject noun and the verb (e.g., *The lawyers hesitate*. *The senator phones the reporter*.) Each training corpus consisted of 10,000 sentences of which 95% were simple intransitive or transitive sentences. The remaining 5% of the sentences were divided equally between the two relative clauses types—250 subject relatives and 250 object relatives—randomly interleaved with the simple sentences. The length of the sentences ranged from 3 to 27 words, with a mean length of 4.5 words reflecting the preponderance of simple intransitive (3 words) and transitive (5 words) sentences.

Ten SRNs, identical in architecture but each with different random starting weights, were trained on different randomly selected sets of 10,000 sentences generated by the probabilistic grammar. This roughly reflects the fact that language learners start out with different initial conditions and are exposed to different subsets of language. The networks were trained to predict upcoming input through the sentence and were provided with three epochs of training, corresponding to three increasing levels of experience with their respective training corpus. A separate test corpus was constructed involving 10 subject relatives (as in (1)) and 10 object relatives (as in (2)). These test sentences were all novel and thus had not been processed by any of the 10 networks prior to testing. Even though the grammar and the vocabulary used in these simulations were both very small, together they can nonetheless generate a large number of sentences with embedded relative clauses such as (1) and (2). The total number of possible sentences of the form used to test the networks—det$N$ comp [det$N V_{trans}$]/[V_{trans}-det$N$]V_{trans} det$N$—is 128,000, of which each network at most would have seen 0.4%. Thus, the SRNs were tested on completely novel sentences in which the words and their general distributional usage were familiar, similar in many ways to the situation facing the human participants in our training study.

The networks' prediction performance on the test sentences were assessed using the Grammatical Prediction Error (GPE) metric, as described in Appendix A of MacDonald and Christiansen (2002). GPE provides a conservative measure of how well the networks are able predict all and only the grammatical continuations at any given point in a sentence. The GPE for an individual word reflects the difficulty that the SRN experienced for that word given the previous sentential context and can be mapped qualitatively onto word reading times, with low GPE values reflecting a prediction for short reading times and high values indicating long predicted reading times.

In Fig. 2, we have replotted the GPE scores averaged across the 10 SRNs from MacDonald and Christiansen (2002) original simulations. The pattern of GPE scores across the three epochs indicate that the SRNs found object relatives harder to process than subject relatives, that increased experience facilitated processing, and that object relatives benefited more from increased experience than did subject relatives. MacDonald and Christiansen did not test the statistical significance of their SRN results, so we conducted a two-way ANOVA on the original main verb GPE scores. As suggested by Fig. 2, there was a main effect of Experience, $F(2,18) = 7.12$, $p < .005$, with decreasing GPE scores across epochs, and a main effect of Relative Clause Type, $F(1,9) = 8.25$, $p < .02$, with subject relatives eliciting lower GPE scores than object relatives. The Experience × Relative Clause Type interaction was only marginally significant, $F(2,18) = 3.10$, $p = .07$, but this is most likely due to a lack of power. Indeed, when we ran an additional 10 SRNs with new random starting weights and different 10,000-word cor-

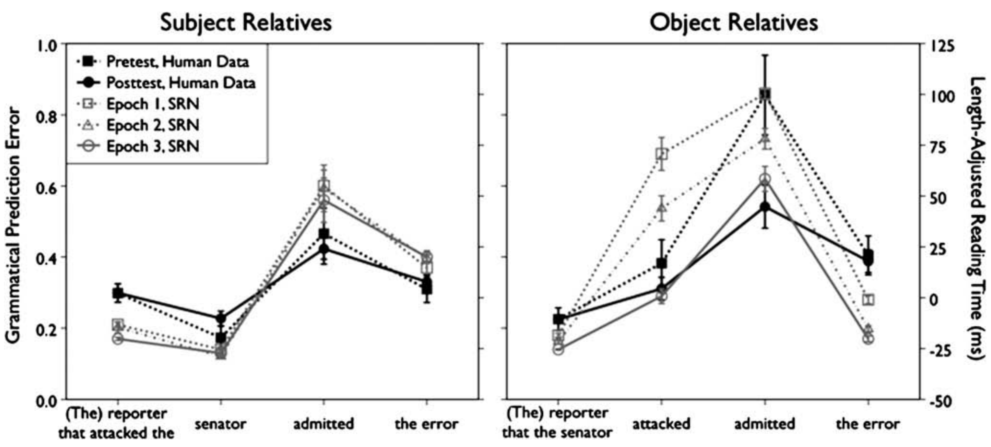

Fig. 2. Comparison of mean Grammatical Prediction Error (GPE) for the 10 models described in MacDonald and Christiansen (2002) and reading times for the Relative Clause Experience Group.

pora (with all other parameters being the same as in MacDonald & Christiansen, 2002), the results with $N = 20$ followed the same pattern and yielded a reliable interaction, $F(2,38) = 3.57$, $p < .04$.

The pattern of SRN results parallels the reliable Session × Relative Clause Type interaction found for our Relative Clause experience group. Fig. 2 further illustrates this parallel by including the pre- and post-test human data from the Relative Clause experience group (there is no equivalent to the Control experience group in the simulations). The Relative Clause experience data in Fig. 2 is normalized to remove the overall drop in reading times from pre- to post-test that was evident in Fig. 1. Much of this decrease in reading times across sessions is due to a growing familiarity with the computerized self-paced reading task; the control group also showed this overall decrease in times, and the effect also obtained for filler items. Because there is no mechanism whereby the SRN can capture reading changes owing to increased familiarity with the experimental apparatus, we recalculated each Relative Clause experience participant's length-adjusted reading times, performing the calculation for the pre-test and post-test data separately. The resultant data set reflects deviations from length-based predictions for each participant given their reading speed in each session. Fig. 2 thus shows these recalculated reading times for the Relative Clause experience group, indicating that there is a close parallel between the patterns of human and SRN performance as a function of experience and relative clause type. In other words, the humans' extra experience with relative clauses, reflected in reading times, closely resembles the SRNs' extra relative clause experience, reflected in GPE scores.

Of course, there are many important differences between the training experience provided to the SRNs in MacDonald and Christiansen (2002) and the kind of training experience that human participants received. For example, there were no semantic representations in the SRN simulations, whereas the human participants read sentences for meaning so as to be able to select the correct of two comprehension probes. Nonetheless, the original simulations do capture two important aspects of our experimental manipulations. First, there are the *within-sentence* differences; as is clear from Fig. 2, the GPE values from the SRNs provide a good fit to reading patterns through the sentence, and a good contrast between subject and object relatives. That is, the networks generally had higher error values (reflecting more uncertainty in the network) in those sentence regions where human readers had longer reading times. This reflects the SRN's ability to capture reading patterns throughout the sentence within subject and object relative clause sentences, and it is consistent with other SRN models of relative clause processing (e.g., Christiansen & Chater, 1999; Elman, 1991; Weckerly & Elman, 1992). Second, there are the *experience-based* differences; the SRNs show greater improvement on object relatives compared with subject relatives as training increases, consistent with the human data. In this context, it is important to note that the networks' training experiences indirectly capture some of the relevant lexical properties of the human Relative Clause experience materials. In particular, the human Relative Clause materials all involved animate common nouns both as head nouns and in the relative clauses. Because the networks would have come across every noun as the subject of a sentence, the nouns can therefore been seen as corresponding distributionally to animate nouns (though there were no explicit semantics associated with this; see Weckerly & Elman, 1992, for a similar distributional approach to animacy). This factor (along with the lack of distributionally inanimate nouns) is likely to have played a role in paralleling the human Relative Clause experience data. The SRNs' ability to show both within-sentence and experience-based effects that are similar to those of human readers is a clearly positive aspect of MacDonald and Christiansen's original modeling effort.

From a modeling perspective, the close match between the SRN predictions and the human reading time data from the Relative Clause experience condition is striking given the relative simplicity of the original simulations. MacDonald and Christiansen introduced the SRN simulations as a demonstration of the differential impact of experience on the processing of subject and object relative clauses. The models were trained on a small fragment of language and not meant to capture general aspects of sentence processing or even all aspects of relative clause processing. In particular, the absence of semantics in the model architecture means that semantic and discourse factors identified as important to relative clause processing, such as discourse load (Warren & Gibson, 2002), and agent-patient similarity (Gordon, Hendrick, & Johnson, 2004; Gordon et al., 2001), in principle cannot be captured by MacDonald and Christiansen's SRNs. Nonetheless, the networks were able to capture the patterns in reading times associated with both processing difficulty across relative clause type as well as the effect of experience. This result provides strong support for the importance of word order in the differences in difficulty in two relative clause types (a major feature of MacDonald and Christiansen's simulations), which has received less attention in recent years as researchers have begun more fine-grained analyses of subject- and object relative discourse and lexical properties. The results also support MacDonald and Christiansen's use of the SRN model to capture the role of experience on relative clause processing, and, more generally, the hypothesized role of Frequency $\times$ Regularity interaction in explaining the specific human reading time patterns at the sentence level.

5. General discussion

This work has investigated the role of experience, or learning, in interpreting complex relative clauses. More specifically, we have tested the claim that the effect of a given amount of sentence comprehension experience will vary with the nature of the sentences that are being comprehended and their relationship to previously experienced sentences. We hypothesized that subject relative structures, which share basic word order similarities with common simple transitive sentences, would receive less benefit from additional experience than would object relative sentences, which are quite idiosyncratic in their structure. This is the essential claim of a Frequency $\times$ Regularity interaction: that the effect of frequency of a linguistic element, such as a sentence type, varies as a function of the relationship of that element to other elements in the language. The results of the present study, together with the simulations, suggest that experience is a powerful factor in sentence comprehension expertise, even among young adult college students. These findings have implications for several different areas of learning, individual differences, and language comprehension.

5.1. Experience and sentence comprehension

The present empirical results and their close relationship to the SRNs argue for a powerful role for experience in language comprehension. This work fits within a long tradition of research emphasizing the importance of practice in gaining skill, including in memory tasks (e.g., Chase & Ericsson, 1982; Ericsson & Kintsch, 1995; Langley & Simon, 1981, chap. 12). Language comprehension is not generally described as a skill requiring practice, in contrast to more specialized abilities such as chess and playing a musical instrument. Indeed, extensive discussion in cognitive science of innate endowments for language in humans often seems to discourage the notion that practice is an important component, beyond the need to learn the unique aspects of the input language. However, whatever the innate components turn out to be, they are not mutually exclusive with an important role for practice, and abundant evidence attests to this view. For example, there is extreme variation in the amount of language input provided by caregivers to their infants, and there is a strong correlation between amount of input and speed of word learning in young children, even when controlling for other factors such as socio-economic status (Hart & Risley, 1995; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002).

Findings for a critical role of experience in sentence processing does not deny that innate individual differences may exist. Rather, it represents a shift in the weight that is given to alternative explanations for individual differences—a shift away from an explanation that invokes the size of a fixed capacity and a stronger emphasis on the amount that an individual has learned from past experiences. Gupta (2003) makes similar arguments for experience–learning–memory relationships at the lexical level, and it is likely that other domains are similarly ripe for an exploration of the role of prior learning in memory and other cognitive performance. We next pursue this role for prior learning with a discussion of statistical learning and its role in individual differences, relative clause interpretation, and sentence comprehension more generally.

5.2. Statistical learning

Over the past decade, statistical learning—the discovery of structure by way of statistical properties of the input—has emerged a general candidate mechanism by which a broad range of linguistic experience can be acquired (for reviews, see Gómez & Gerken, 2000; Saffran, 2003). This type of implicit learning has been demonstrated across a variety of natural and artificial language learning situations, including learning of information that is potentially highly relevant to sentence comprehension processes, such as discovering phonological and distributional cues to lexical categories (Monaghan, Chater, & Christiansen, 2005), acquiring gender-like morphological systems (Brooks, Braine, & Catalano, 1993; Frigo & McDonald, 1998), locating syntactic phrase boundaries (Saffran, 2001, 2003), using function words to delineate phrases (Green, 1979), integrating prosodic and morphological cues in the learning of phrase structure (Morgan, Meier, & Newport, 1987), and detecting long-distance relationships between words (Gómez, 2002; Onnis, Christiansen, Chater, & Gómez, 2003). Of particular relevance is an individual differences study by Misyak and Christiansen (2007) who found that statistical learning ability was a stronger predictor of relative clause comprehension than the reading span measure. Together, these studies suggest that statistical learning may play a strong role in the accumulation of linguistic experience relevant for sentence processing.

Moreover, within natural language comprehension and production studies, there is clear evidence that prior experience with sentences of a given syntactic structure affects both subsequent comprehension of similar structures and the probability that a speaker will utter a sentence with the same or similar structure, even when there is no meaning overlap between sentences (see Ferreira & Bock, 2006, for review). These effects, often termed *structural persistence* or *syntactic priming*, have been described as stemming from statistical learning at the syntactic level (Bock & Griffin, 2000; Chang, Dell, & Bock, 2006) or at the syntactic–semantic interface (Chang, Bock, & Goldberg, 2003; Hare & Goldberg, 1999). The experience effects from the present study can be viewed as examples of statistical learning of information relevant to sentence processing. The current results, in which effects of experience were observed over several days, extend previous findings of structural persistence effects, which have typically been assessed within a single testing session. To date, structural persistence effects in comprehension studies have typically been relatively weak, often requiring the repetition of a verb between prime and target sentences in order to observe a reliable effect on target sentence reading time (Arai, Van Gompel, & Scheepers, 2007). The present results suggest that given a multi-day presentation of "primes" (that is, the experience manipulation), robust effects can be observed on "target" items (sentences in the post-test) without any content word overlap between prime and target. Long-lasting, robust effects such as these are to be expected if statistical learning mechanisms are truly underlying important aspects of language acquisition and use.

5.3. Statistical learning and relative clause processing

Previous work has suggested that object relative clauses contain ambiguities in interpretation that are not present in subject relatives (Gennari & MacDonald, 2008), and that object relative comprehension difficulty correlates with the distributional patterns in corpora, such that higher frequency patterns are associated with more rapid processing (Gennari & MacDonald, submitted for publication; Reali & Christiansen, 2007a). Our work here directly manipulated relative clause frequencies and showed that these manipulations yielded more improvement with object than subject relatives (the Frequency × Regularity interaction). We suggest that additional experience with object relatives is particularly useful in helping comprehenders navigate the ambiguities inherent in this structure, which in turn raises the question of exactly what it is that the Relative Clause experience group learned about relative clauses from the experience manipulation.

As shown by corpus analyses, object relative clauses contain a complex web of lexical, structural and discourse-related distributional information (Gennari & MacDonald, submitted for publication; Jaeger & Wasow, 2005; Race & MacDonald, 2003; Reali & Christiansen, 2007a, 2007b; Roland et al., 2007). The present study was not designed to manipulate any particular subset of those constraints, and so definitive answers are not possible concerning exactly what the Relative Clause group learned. Several plausible hypotheses are available, however. One obvious possibility is that the Relative Clause readers, like the MacDonald and Christiansen (2002) model, learned about the unusual pattern of word order that is inherent in object relatives but is rare in the rest of the English language, in which a preverbal noun is the object of the embedded clause verb. Another source of learning that probably modulated object relative reading times concerns the pairing of noun types and syntactic structures. All of the relative clauses in the present study (and in most other psycholinguistic studies of relative clauses) had animate common nouns as the head of the relative clause and the noun within the relative clause, such as in *reporter* and *senator* in (1–2). In the language as a whole, however, object relatives more typically contain inanimate head nouns (Gennari & MacDonald, submitted for publication; Roland et al., 2007) and pronouns rather than common nouns as the embedded subject (Race & MacDonald, 2003; Reali & Christiansen, 2007a). Studies have shown that both of these noun properties affect relative clause processing (Gennari & MacDonald, 2008; Gordon et al., 2001, 2004; Mak et al., 2002; Race & MacDonald, 2003; Traxler et al., 2002; Warren & Gibson, 2002). The Relative Clause experience condition may have changed Relative Clause participants' knowledge about distributional patterns of noun type and relative clause type co-occurrences. Thus whereas the model's performance, following exposure to a relatively simple grammar and limited vocabulary, can be traced to learning about word order, the human case affords the opportunity to learn at other "grains" of statistics as well. We see this learning, both about word orders and about lexical–structural pairings, as changing what comprehenders implicitly know about the distribution of object relatives, consequently affecting constraint-based ambiguity resolution processes and reading patterns, as readers apply these updated probabilistic constraints to new text.

5.4. *Linking statistical learning and constraint-based comprehension*

Beyond implications for relative clause processing, the results of this experiment and simulation have implications for studies of constraint-based ambiguity resolution. In the constraint-based approach, language comprehension is accomplished through the rapid weighing of many probabilistic constraints concerning the likely interpretation of a sentence, given information about real-world plausibility, frequency, and other factors (see MacDonald & Seidenberg, 2006, for review). Constraint-based sentence processing research has primarily emphasized the time course of use of various kinds of constraints during on-line comprehension. Far less attention has been directed toward explaining how the comprehender becomes equipped with information about the constraints. Our results shed light on this process by showing that comprehenders rapidly learn from comprehension experience and apply it to subsequent linguistic input.

This approach to learning of probabilistic constraints has implications for studies of individual differences. For example, Pearlmutter and MacDonald (1995) investigated individual differences in use of probabilistic constraints in syntactic ambiguity resolution. They found that readers who scored high on the reading span task displayed reading patterns that reflected the application of complex probabilistic constraints relevant to interpretation of the ambiguous sentences, whereas low-span participants' reading times reflected the influence of only simple constraints. Pearlmutter and MacDonald speculated that the differences in performance stemmed from differences in experience, that the high-span comprehenders read more and therefore had sufficient exposure to complex combinatorial information that they could apply on-line during ambiguity resolution. Similarly, Farmer, Christiansen, and Kemtes (2005) found that when strong statistical biases existed in sentence materials, high-span individuals appeared to rely primarily on the statistical information, whereas their low-span counterparts ignored statistical information in favor of other cues, such as discourse context. An obvious extension of this work would be to examine the effects of experience in comprehenders with different abilities and/or experiences, such as readers with different amounts of reading experience or "print exposure" (Acheson, Wells, & MacDonald, 2008; Stanovich & West, 1989).

The importance of experience in relative clause comprehension, and the link between adult sentence processing and learning processes, are further underscored by work in child language acquisition, especially studies of relative clause acquisition. Roth (1984) provided children with extra experience and feedback on processing relative clauses, including subject and object relatives like those investigated here. A control group received experience processing other structures. The results showed that children with direct experience in comprehending relative clauses improved their comprehension rates over pre-test levels compared with the children in the control condition. Roth argued that the experience with relative clauses lead to improved processing efficiency and suggested that developmental changes in relative clause processing should be traced to linguistic experience, not growth of working memory. This view is consistent with experience-based continuity across the life span in which the use of constraints in adult language comprehension emerges as the product of a developmental process driven by the integration of multiple cues (Bates & MacWhinney, 1987; Farmer, Christiansen, & Monaghan, 2006; Gennari & MacDonald, 2006; Seidenberg, 1997; Seidenberg & MacDonald, 1999; Snedeker & Trueswell, 2004).

5.5. Nature and nurture in individual differences

Our approach in this research has been to emphasize the learning side, rather than the capacity limits, of Miller's (1956) observations concerning memory performance and its implications for language comprehension. We see this approach as a useful addition to the sub-field of sentence processing that has tended to emphasize capacity limits over the role of learning in relative clauses (e.g., Gibson, 1998; Just & Carpenter, 1992; King & Just, 1991; Lewis et al., 2006) and to the sub-field that has tended to emphasize rapid use of probabilistic constraints without exploring how the knowledge came to be learned (MacDonald & Seidenberg, 2006). This focus on learning does not mean that we do not see innate individual differences as a potential source of behavioral variation. Indeed, MacDonald and Christiansen (2002) argued against a conception of processing capacity as separate from learning about the language; they did not argue against innate variation in general. They discussed some potential sources of innate variation and how self-selection of experiences (such as how much people choose to read) may vary in complex ways with innate factors. We see computational modeling that explores parameter spaces of both innate constraints and amount of learning (e.g., Gupta & Tisdale, submitted for publication) as having an important role in future research on the complex joint effects of innate constraints and experience in understanding individual differences in linguistic performance.

Recent data in statistical learning suggest another possible source of innate variation with potential to have complex interactions with experience—individual differences in statistical learning itself. Whereas some researchers had previously assumed that there are not substantial individual differences in statistical learning (e.g., Reber, 1993), recent studies have shown individual differences in success in statistical learning tasks. For example, individuals with developmental or acquired language impairments appear less able to learn statistical regularities involving both linguistic and nonlinguistic stimuli (Christiansen, Kelly, Shillcock, & Greenfield, submitted for publication; Evans & Saffran, 2005; Hsu, Christiansen, Tomblin, Zhang, & Gómez, 2006; Plante, Gómez, & Gerken, 2002). It is currently not clear the extent to which these performance differences reflect individual differences in learning per se, or in perceptual acuity, other cognitive abilities, experiences, or some combinations of such factors. Nonetheless, these differences offer intriguing possibilities concerning the complex nature of interactions between innate abilities and linguistic experience in language comprehension processes. We suspect that the combination of computational modeling and experience manipulations that we have used will prove useful in this area as well.

Acknowledgments

This research was supported by NIH Grants P50-MH64445, R01-HD47425, and the Graduate School of the University of Wisconsin-Madison. We thank Thomas Farmer, Richard Lewis, Jennifer Misyak, Florencia Reali, Mark Seidenberg, and Tessa Warren for their comments and helpful discussion.

Appendix A

Reading Times (ms) not adjusted for length and (standard deviations) for Relative Clause and Control Experience Groups for Subject Relative (SR) and Object Relative (OR) sentences in Pre-test and Post-test sessions.

	Region 1	Region 2	Region 3	Region 4
Relative clause experience group (N = 32)				
SR, pre-test	383 (107)	381 (121)	425 (152)	376 (89)
OR, pre-test	371 (104)	415 (162)	491 (192)	393 (103)
SR, post-test	301 (69)	292 (72)	330 (103)	304 (60)
OR, post-test	285 (60)	310 (98)	345 (103)	313 (60)
Control experience group (N = 32)				
SR, pre-test	375 (95)	389 (125)	449 (174)	367 (94)
OR, pre-test	370 (92)	426 (144)	484 (182)	387 (113)
SR, post-test	302 (70)	298 (77)	319 (95)	312 (78)
OR, post-test	293 (66)	307 (91)	370 (124)	330 (81)

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*, 278–289.

Arai, M., Van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54*, 218–250.

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

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–332). New York: Wiley.

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

Booth, J., MacWhinney, B., & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development, 71*, 981–1003.

Brooks, P. J., Braine, M. D., & Catalano, L. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32*, 76–95.

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

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 234–272.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 16* (pp. 1–58). New York: Academic Press.

Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science, 23*, 157–205.

Christiansen, M. H., Kelly, L., Shillcock, R., & Greenfield, K. (submitted for publication). Impaired artificial grammar learning in agrammatism.

Daneman, M. E., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Dickey, M. W., & Thompson, C. (2004). The resolution and recovery of filler-gap dependencies in aphasia: Evidence from on-line anomaly detection. *Brain and Language, 88*, 108–127.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning, 7*, 195–225.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Evans, J., & Saffran, J. R. (2005). Statistical learning in children with specific language impairment. Paper presented at the Boston University Conference on Language Development, Boston, MA, November 2005.

Farmer, T. A., Christiansen, M. H., & Kemtes, K. A. (2005). Sentence processing in context: The impact of experience on individual differences. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society* (pp. 642–647). Mahwah, NJ: Lawrence Erlbaum.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences, 103*, 12203–12208.

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

Ferreira, F., & Clifton, J. C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348–368.

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language, 101*, 50–63.

Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language, 38*, 218–245.

Gennari, S. P., & MacDonald, M. C. (2006). Acquisition of negation and quantification: Insights from adult production and comprehension. *Language Acquisition, 13*, 125–168.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy and relative clause comprehension. *Journal of Memory and Language, 58*, 161–187.

Gennari, S. P., & MacDonald, M. C. (submitted for publication). Linking production and comprehension processes: The case of relative clauses.
Gibson, E. A. F. (1998). Linguistic complexity: Locality and syntactic dependencies. *Cognition, 68*, 1–76.
Gómez, R. (2002). Word frequency effects in priming performance in young and older adults. *Journals of Gerontology. Series B: Psychological Sciences and Social Sciences, 57B*(3), P233–P240.
Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*, 178–186.
Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Language, 27*, 1411–1423.
Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97–114.
Green, T. R. (1979). The necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior, 18*, 481–496.
Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29*, 261–291.
Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology (A), 56*, 1213–1236.
Gupta, P., & Tisdale, J. (submitted for publication). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate.
Hare, M. L., & Goldberg, A. E. (1999). Structural priming: Purely syntactic? In M. Hahn & S. C. Stones (Eds.), *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 208–211). Mahwah, NJ: Erlbaum.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
Hsu, H.-J., Christiansen, M. H., Tomblin, J. B., Zhang, X., & Gómez, R. L. (2006). Statistical learning of nonadjacent dependencies in adolescents with and without language impairment. Poster presented at the 2006 Symposium on Research in Child Language Disorders, Madison, WI.
Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337–374.
Jaeger, T. F. (2005). Optional that indicates production difficulty: Evidence from disfluencies. In *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop*, 10–12 September 2005. Aix-en-Provence, France, pp. 103–109.
Jaeger, T.F. & Wasow, T. (2005). Processing as a Source of Accessibility Effects on Variation. Proceedings of the 31st Berkeley Linguistics Society.
Juliano, C., & Tanenhaus, M. K. (1993). Contingent frequency effects in syntactic ambiguity resolution. In *Proceedings of the 15th Annual Meeting of the Cognitive Science Society*, pp. 593–598.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.
Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228–238.
Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes, 22*, 860–897.
King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.
Langley, P., & Simon, H. A. (1981). The central role of learning in cognition. In J. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.
Lewis, R. L., & Nakayama, M. (2002). Syntactic and positional similarity effects in the processing of Japanese embeddings. In M. Nakayama (Ed.), *Sentence Processing in East Asian Languages*. CSLI Publications.
Lewis, R. L., Vasishth, S., & VanDyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10*, 447–454.
MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain & Language, 78*, 17–42.
MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.
MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction approaches to lexical and sentence comprehension. In M. A. Gernsbacher & M. J. Traxler (Eds.), *Handbook of psycholinguistics*. Elsevier.
MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition, 29*, 95–141.
Mak, W. M., Vonk, M., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*, 50–68.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.
Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol II, pp. 419–491).
Misyak, J. B., & Christiansen, M. H. (2007). Extending statistical learning farther and further: Long-distance dependencies, and individual differences in statistical learning and language. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual cognitive science society conference* (pp. 1307–1312). Austin, TX: Cognitive Science Society.
Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 69–90). Hillsdale, NJ: Erlbaum.
Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition, 96*, 143–182.
Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498–550.

Onnis, L., Christiansen, M., Chater, N., & Gómez, R. (2003). *Reduction of uncertainty in human sequential learning: Evidence from artificial language learning. Proceedings of the 25th annual conference of the cognitive science society*. Mahwah, NJ: Lawrence Erlbaum. pp. 886–891.
Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language, 34*, 521–542.
Plante, E., Gómez, R., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders, 35*, 453–462.
Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In *Proceedings of the twenty-fifth annual meeting of the cognitive science society* (pp. 946–951). Mahwah, NJ: Lawrence Erlbaum.
Reali, F., & Christiansen, M. H. (2007a). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language, 57*, 1–23.
Reali, F., & Christiansen, M. H. (2007b). Word chunk frequencies affect the processing of pronominal object-relative clauses. *The Quarterly Journal of Experimental Psychology, 60*, 161–170.
Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford University Press, New York: New York.
Rohde, D. L. T. (2002). A connectionist model of sentence comprehension and production. Unpublished PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*, 348–379.
Roth, F. P. (1984). Accelerating language learning in young children. *Child Language, 11*, 89–107.
Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language, 44*(4), 493–515.
Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*(4), 110–114.
Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19*, 1–30.
Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275*, 1599–1603.
Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23*, 569–588.
Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238–299.
Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402–433.
Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*, 69–90.
Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33*, 285–318.
Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language, 82*, 767–794.
Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition, 85*, 79–112.
Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review, 103*, 761–772.
Weckerly, J., & Elman, J. (1992). A PDP approach to processing center-embedded sentences. In *Proceedings of the fourteenth annual meeting of the cognitive science society* (pp. 414–419). Hillsdale, NJ: Lawrence Erlbaum Associates.
Wingfield, A., Peelle, J., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging, Neuropsychology, and Cognition, 10*, 310–322.