The Processing Basis of Syntactic Persistence: We Repeat What We Learn

Victor S. Ferreira

University of California, San Diego

Please address correspondence to:

Victor S. Ferreira

Department of Psychology 0109

University of California, San Diego

La Jolla, CA 92093-0109

vferreira@ucsd.edu
Abstract

Speakers’ current descriptions tend to have the same syntactic structures as sentences they have recently processed. Learning-based approaches to language production attribute such syntactic persistence to the strengthening (learning) of message-to-syntax mappings – the knowledge that certain kinds of messages can be expressed with the primed syntactic structures. Learning-based approaches generally predict that less-preferred outcomes should show a stronger learning or persistence effect than more preferred outcomes. Three experiments (N = 192) using recall-based syntactic-persistence tasks showed that reduced embedded clauses (embedded clauses without complementizer that) show more syntactic persistence to the extent that they are less preferred in a given syntactic context, and that this pattern cannot be explained through similarity of syntactic context. Thus, syntactic persistence specifically and learning-based performance generally is inversely sensitive to degree of preference. This shows that persistence results from the strengthening or learning of message-to-syntax mappings, and reveals why such learning-based persistence does not lead to the unlearning of less preferred outcomes.
To use language, we must solve a formidable information-processing problem. At one end, language use involves meanings (sometimes called *messages*), the representations of which are rooted in, for example, propositions (e.g., Fodor, 1975), perceptually based representations (e.g., Barsalou, 1999; Johnson-Laird, 1983), or some other non-linguistic medium. At the other end, language use involves linguistic features, represented in terms of discourses, sentences, words, and phonemes (which eventually are realized in terms of articulatory gestures or acoustic features). These two domains are fundamentally unlike one another (a dog doesn't mean anything like the sounds /d/, /ʊ/, and /ɡ/), but yet, to be able to produce or understand linguistic expressions, we must be able to map elements of meaning onto features of language.

An especially complex aspect of this problem concerns how language users map the relational content of messages onto a particular language’s inventory of syntactic structures. For example, a *transitive* message structure involves a performer (often termed an *agent*) conducting an action upon a performed-upon entity (often termed a *patient*; e.g., a domestic canine actively pursuing a passenger vehicle). A transitive message structure can be expressed by, among other syntactic forms, an active syntactic structure, which maps the agent onto a grammatical subject, the action onto a verb, and the patient onto a direct object (e.g., “The dog chased the car”), or a passive syntactic structure, which maps the patient onto a grammatical subject, the action onto a verb with passive morphology, and the agent onto a by-object (e.g., “The car was chased by the dog”). Thus, producing or comprehending sentences requires knowledge that relates message-structure elements to constituents in particular kinds of syntactic structures. In this paper, such knowledge representations are called *message-to-syntax mappings*. 
A language-production phenomenon that has proven sensitive to syntactic knowledge is called *syntactic persistence* (or sometimes *syntactic priming*). This refers to the fact that speakers are especially likely to produce sentences that have overall forms that are identical or similar to sentences that they have recently produced (Bock, 1986) or comprehended (Bock, 2002). For example, a speaker who has recently said or heard a passive *prime* sentence like “The referee was punched by one of the fans” is more likely (relative to when the prime was not a passive) to subsequently produce a *target* sentence with another passive, like “The boy was awakened by the alarm clock” rather than an active like “The alarm clock awoke the boy” (Bock, 1986). Such persistence cannot be fully explained by the repetition of semantic, lexical, or prosodic properties (see, e.g., Bock, 1989; Bock & Loebell, 1990; Bock, Loebell, & Morey, 1992; Chang, Bock, & Goldberg, 2002; Ferreira, in press; Pickering & Branigan, 1998), suggesting that at least part of the effect is due to the repetition of features that are specifically syntactic in nature.

One explanation for syntactic persistence is that it reflects the learning or strengthening of message-to-syntax mappings (Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000). Specifically, the claim is that each time a speaker uses a particular message-to-syntax mapping, that mapping is strengthened by a small amount. The long term effect of strengthening successfully used message-to-syntax mappings is that with increasing experience, a language user will learn the possible mappings in the language they speak. A shorter term effect, however, is syntactic persistence: using a particular message-to-syntax mapping strengthens it, and thus when a speaker wishes to express that message structure again, the just-strengthened mapping is relatively more likely to be used.

Another relevant fact about sentence production is that a given message structure is
expressed by some syntactic structures much more commonly than others (e.g., actives express transitive ideas much more often than passives do). On the one hand, the claim that syntactic persistence reflects learning can partly explain such preferences, because preferred syntactic structures should be strengthened more and thus used more in the future (and that is one of the things that may lead them to continue to be preferred). But what prevents this trend from causing the learning of more preferred options to overwhelm the knowledge of less preferred options, thereby extinguishing those less preferred options completely? (For a review of the relationship between the frequency of a structure’s usage, language change, and grammar change, see Kroch, 1989, 1994)

In this paper, I evaluate this issue by describing and testing the predictions of such learning-based approaches to syntactic persistence. The essential claim is that sentence-production mechanisms confront this challenge with a kind of syntactic affirmative action: Less preferred options exert a stronger impact on the sentence production system than more preferred options, thereby allowing them a basic level of viability with a minimum degree of use. It should be noted at the outset that there are explanations of syntactic persistence that do not claim that it reflects longer-term learning in this manner (see especially Pickering & Branigan, 1998). While the focus here will be on learning-based accounts, I will discuss implications of the data described below for other accounts in the General Discussion.

Syntactic Persistence as the Learning of Message-to-Syntax Mappings

Learning-based approaches explain syntactic persistence as a manifestation of the mechanisms responsible for acquiring message-to-syntax mappings. An important assumption
of such accounts (see especially Chang et al., 2000) is that the amount that a particular mapping is strengthened is inversely proportional to how well learned that mapping was before it was strengthened. This follows from the basic learning strategy that some learning-based models of cognitive performance generally use, which is to modify existing knowledge by an amount proportional to how much the model’s performance differs from an ideal or target level of performance (this is sometimes called error-based learning). Thus, for a given input, if a learning-based model computes an output that differs greatly from the appropriate target output (i.e., the input-output mapping is less well-learned), its knowledge is adjusted to a greater extent, but if the model computes an output that differs less from the appropriate target output (i.e., the input-output mapping is better learned), its knowledge is adjusted to a lesser extent. This implies that mappings that are less well-learned – that are less entrenched – should be subject to greater strengthening, whereas mappings that are better learned – that are more entrenched – should be subject to less strengthening. This entrenchment-based learning strategy acts as a kind of representational affirmative action, whereby less common mappings exert a greater impact on a learning-based system than more common mappings do.

If syntactic persistence is due to the learning of message-to-syntax mappings within an entrenchment-based learning framework, then a specific prediction follows: less preferred message-to-syntax mappings – less well-learned mappings – should show greater syntactic persistence and more preferred message-to-syntax mappings – better learned mappings – should show less syntactic persistence. Very little current evidence bears directly on this point. In one investigation, Hartsuiker and Kolk (1998) showed with a correlational analysis that the degree of persistence of a particular syntactic form was not strongly related to a measure of the baseline preference for that form in the language. They did, however, find a correlation in the predicted
direction (–.21), even if that correlation was not significant.

One previous result that is consistent with the entrenchment-based learning prediction comes from Ferreira (in press). That report described experiments exploring the syntactic persistence of sentence-complement structures like (1) and (2):

(1) The jury believed [that the witness told the truth].

(2) The jury believed [the witness told the truth].

Sentence-complement structures consist of a main verb (“believe”) followed by an embedded clause, between which a speaker can mention an optional complementizer that, thereby producing a sentence-complement structure with a full embedded clause like (1), or with a reduced embedded clause like (2). Such full and reduced embedded clauses are syntactic constituents that express embedded propositions – predicate-argument statements that are subordinated within other message elements. For example, in a sentence-complement structure, a full or reduced embedded clause expresses a proposition that is subordinated within an action.

Important for present purposes is that overall, reduced embedded clauses appear to be substantially less preferred than full embedded clauses. In a number of reports (Ferreira, in press; Ferreira & Dell, 2000; Ferreira & Firato, in press; Garnsey, Lotocky, & McConkie, 1992), when speakers have the option of producing either a full or reduced embedded clause, they produced full embedded clauses more often than reduced embedded clauses with almost a two-to-one ratio. Furthermore, if we look at English sentence types other than sentence-complement structures that include embedded clauses, most require full embedded clauses and forbid reduced embedded clauses (e.g., the noun-complement structures described below), whereas very few require reduced embedded clauses and forbid full embedded clauses (with one notable exception to be exploited in Experiments 2 and 3). Across the language, these grammatical patterns will
lead to a bias that favors the production of full embedded clauses whenever an embedded clause is to be produced.

If full embedded clauses are indeed preferred more than reduced embedded clauses, then according to an entrenchment-based learning account, full embedded clauses should cause less syntactic persistence and reduced embedded clauses should cause more syntactic persistence. Results from Ferreira (in press) are consistent with this prediction. In those experiments (as in the experiments below), speakers produced target sentences that were sentence-complement structures, in which they could include a full embedded clause (like (1)) or a reduced embedded clause (like (2)). In one experiment, the production of such sentence-complement targets was measured in two different priming conditions (see Figure 1). In one priming condition, speakers produced prime sentences that were also sentence-complement structures with full or with reduced embedded clauses. Results showed that such primes caused syntactic persistence: Speakers produced target sentence-complement structures with full embedded clauses about 24% more often after prime sentence-complement structures with full embedded clauses, compared to after prime sentence-complement structures with reduced embedded clauses. Note that this effect could be due to the persistence of full or of reduced embedded clauses (or both), since this condition compared the persistence of full embedded clauses to the persistence of reduced embedded clauses. In the other priming condition in the same experiment, speakers produced noun-complement structures or control sentences like (3) and (4):

(3) The theory [that penguins built the igloos] was completely false.

(4) The theory [of the melted igloos] was completely false.

Note that noun-complement structures like (3) include a noun (e.g., “theory”) that takes a full embedded clause argument (“that penguins built the igloos”), but that the prepositional phrase
controls like (4) do not involve producing a reduced embedded clause (or any embedded clause at all). If full embedded clauses cause less syntactic persistence, then we should find that compared to the priming effect from prepositional phrase controls like (4), noun-complement structures like (3) cause little syntactic persistence. Indeed, as shown in Figure 1, such noun-complement structures were ineffective primes: speakers were not more likely to produce a sentence-complement target with a full embedded clause after they produced a noun-complement structure like (3), which also included a full embedded clause, compared to after a matched control prime like (4), which included a prepositional phrase modifier.

--- Figure 1 about here ---

Further results from Ferreira (in press) are consistent with this interpretation. Note that since the prepositional phrase controls like (4) do not include any kind of embedded clause, they can be taken to roughly indicate a level of neutral or baseline priming. Figure 1 shows that compared to this baseline, sentence-complement primes with reduced embedded clauses decreased full-embedded clause production more than sentence-complement primes with full embedded clauses increased it. Consistent with this, an independent experiment in the same report compared the priming effect from sentence-complement structures with full and reduced embedded clauses to the priming effect from neutral primes — sentences that had similar lexical content but could not take an embedded clause argument at all (“The jury nodded together when the young woman told the truth”). Results showed that after the neutral primes, speakers produced full embedded clauses in sentence-complement targets 73% of the time. Compared to this, sentence-complement primes with reduced embedded clauses strongly decreased full-embedded clause production in sentence-complement targets (59%), but sentence-complement primes with full embedded clauses increased full-embedded-clause production in targets only
slightly (77%). These results suggest that reduced embedded clauses might indeed exert a stronger persistence effect than full embedded clauses, supporting the predictions of an entrenchedment-based learning account generally.

However, the results in Figure 1 can also be explained without appealing to entrenchedment, and thus they do not exclusively support an entrenchedment-based learning account of syntactic persistence. An obvious alternative explanation is that the results are due to the overall similarity of the syntactic structures in the prime and target sentences. Note that in the sentence-complement targets, the critical embedded clause constituent is an argument of a verb, whereas in the noun-complement primes, the critical embedded clause constituent is an argument of a noun. It is possible that for an embedded clause in a prime to influence an embedded clause in a target, those embedded clauses might need to be produced in similar syntactic contexts. Note that if this alternative interpretation is correct, it would make an important claim about the representation of message-to-syntax mappings, as it would suggest that such knowledge is especially influenced by the syntactic details of the structures being mapped to.

Some evidence supports the idea that similarity in overall syntactic structure may determine whether syntactic persistence is or is not observed. For example, Bock and Loebell (1990) showed that a *prepositional locative* like (5) can prime a *prepositional dative* like (6):

(5) The wealthy widow drove an old Mercedes to the church.

(6) The boy is giving an apple to the teacher.

This priming occurred despite the fact that prepositional locatives and prepositional datives differ in a number of respects (e.g., thematic structure). Note however that the postverbal material (the material that exhibits persistence) in prepositional locatives and prepositional datives is syntactically similar (consisting of a noun phrase followed by a prepositional phrase), and so the
priming effect can be explained if we assume that syntactic persistence is sensitive to such syntactic similarity. Complementarily, Bock and Loebell also found that prepositional dative targets like (6) were unaffected by an infinitive form like (7):

(7) Susan brought a book to study.

Here, the lack of effect can be attributed to syntactic differences between the prime and target sentences, as the critical portion of the infinitives consisted of a noun phrase followed by a verb phrase (“a book to study”), whereas the corresponding portion of the prepositional dative structures consisted of a noun phrase followed by a prepositional phrase (“an apple to the teacher”).

It is important to note, however, that within-constituent syntactic differences are unlikely to influence syntactic persistence. Bock (1986) found syntactic persistence despite the fact that prime and target sentences were not controlled for the presence of modifiers on the constituents that vary between the syntactic alternatives (e.g., note that in the prime (5) above, the direct object is modified by the adjective “old,” but in the target (6), it is not). Fox Tree and Meijer (1999) confirmed this systematically by measuring whether target datives with simple postverbal arguments were differentially influenced by primes with postverbal arguments that were simple or complex (modified by a relative clause). No difference was found, suggesting that syntactic persistence was insensitive to such variations in local syntactic similarity. These results thus show that whether primes and targets are syntactically similar within the varying syntactic constituents is unlikely to be relevant to the syntactic persistence effect (e.g., whether the direct object is modified by a prepositional phrase, a relative clause or whatever), though it still leaves open the possibility that the syntactic similarity of those constituents with respect to neighboring constituents (e.g., whether the direct and indirect objects are configured as noun-phrase-plus-
prepositional-phrase, or as a noun-phrase-plus-noun-phrase) determines when persistence is and is not observed.

In sum, learning-based accounts of syntactic persistence predict that degree of syntactic persistence should be sensitive to degree of entrenchment, so that less preferred syntactic structures (e.g., passives, reduced embedded clauses) should exert a stronger persistence effect than more preferred structures (e.g., actives, full embedded clauses). Some results are consistent with this possibility (Ferreira, in press), though those results may be explained in terms of syntactic similarity rather than in terms of entrenchment. Other results fail to support this prediction (Hartsuiker & Kolk, 1998). Experiment 1 here tested whether representations of message-to-syntax knowledge are especially sensitive to syntactic similarity, and then Experiments 2 and 3 determined specifically whether degree of syntactic persistence is indeed sensitive to degree of entrenchment.

Experiment 1

Speakers produced sentence-complement targets like (8), which could be produced with either a full or reduced embedded clause:

(8) The director announced (that) Hollywood’s hottest actor would play the part.

These were produced in two priming conditions. In one, targets were preceded by the production of full or reduced sentence-complement primes with full or reduced embedded clauses like (9) and (10):

(9) The teacher noticed that the failing student skipped class.

(10) The teacher noticed the failing student skipped class.
In the other priming condition, sentence-complement targets were preceded by the production of full or reduced object-relative-clause primes like (11) and (12):

(11) The teacher noticed the student *that* the school nurse said was sick.

(12) The teacher noticed the student the school nurse said was sick.

The sentence-complement primes, like the sentence-complement targets, included a full or reduced embedded-clause that is an argument of the main verb of the sentence, whereas the object-relative primes, unlike the sentence-complement targets, included a full or reduced embedded relative clause that modifies the object noun of the sentence. To the extent that syntactic persistence is sensitive to the similarity of the syntactic relationship of the embedded clause to its neighboring constituents, the production of sentence-complement targets should be influenced more by the production of the syntactically similar sentence-complement primes than by the production of the syntactically dissimilar object-relative primes. In contrast, both the sentence-complement primes and the object-relative primes tested priming sentences that included less preferred reduced embedded clauses. To the extent that syntactic persistence is greater when speakers produce such less preferred reduced embedded clauses, both the syntactically similar sentence-complement priming condition and the syntactically dissimilar object-relative priming condition should yield significant priming effects.

All experiments used a recall-based methodology to elicit production. Speakers produced prime and target sentences shortly after encoding those whole sentences into memory. The validity of this task for assessing syntactic production depends on the well-known observation that people have weak memory for the surface forms of sentences they encode, at the same time as they have good memory for the meaning or gist of those sentences (e.g., Sachs, 1967). This implies that when speakers recall sentences from memory, they regenerate the syntactic
structures of those sentences based on the relatively well-encoded gist representation, rather than simply recall the sentence verbatim based on memory of the syntactic structure per se. Speakers do have some very short-term memory for the perceptual characteristics of sentences, but since the target sentences produced in these experiments were never produced immediately after they were encoded, any such perceptual memory is unlikely to influence target-sentence production. Hence, a recall-based sentence-production task controls for the meaning or gist representation that drives syntactic production, so that any variation that is observed as a function of the experimental manipulations is likely to derive from the sensitivity of the processes that generate syntactic structures, rather than the sensitivity of the processes that generate, say, the meaning or gist representation that drives sentence production in the first place. In support of this, investigations of syntactic persistence specifically (Chang et al., 2002; Ferreira, in press; Fox Tree & Meijer, 1999) and language production more generally (e.g., Bock, 1977; Bock & Irwin, 1980; Bock & Warren, 1985; Ferreira & Dell, 2000; Lombardi & Potter, 1992; Potter & Lombardi, 1990, 1998) have shown that syntactic persistence as well as other language-production phenomena are commonly observed when speakers produce sentences from memory.

**Method**

*Speakers.* Sixty-four members of the UCSD community participated either for class credit or cash payment ($6 or $7). All speakers reported learning English as their first language.

*Apparatus.* The experiment was administered with PsyScope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993), run on Macintosh 6500/250 computers with 17-inch Applevision 750AV color monitors set to a resolution of 832 x 624. Voice responses were collected with Shure SM10A unidirectional headworn microphones inputting to Marantz PMD201 cassette recorders
and PsyScope response boxes. The voice key was calibrated separately for each speaker.

**Procedure.** Speakers were tested individually. On each critical trial, speakers encoded first the target sentence into memory (like (8)) followed by the prime (like (9–12)). Speakers then produced the prime sentence when prompted by its main subject and verb (e.g., “teacher noticed”), and finally they produced the target when prompted by its main subject and verb (e.g., “director announced”). Note that the production of the prime immediately followed its encoding, increasing the likelihood that it would be produced as presented; in contrast, the production of the target was separated from the encoding of the target by both the encoding and the production of the prime, increasing the likelihood that target mention will vary more freely and so be influenced by the production of the prime. All stimuli were presented in bold Courier 14-point font, centered vertically and horizontally on the screen. An equal number of filler trials (which represented a variety of non-embedded-clause-taking structures) were also used with the identical trial structure as the criticals, except that the sentences were prompted in the same order as originally presented, so that across the entire experiment, speakers could not anticipate which sentence would be prompted first in a trial.

The main experiment was preceded by an instruction session that included a sample sentence pair and a practice block of five sentence pairs. Sample and practice sentences were similar to the fillers. An experimental session took about 40 minutes.

**Materials and design.** Materials were adapted from Ferreira (in press). Forty-eight sentence-complement-taking verbs were taken from the set normed by Garnsey, Lotocky, and McConkie (1992). Each verb was combined with a unique main subject and then either a full or reduced embedded clause to create a sensible, meaningful sentence. To create object-relative priming sentences, each main-subject-plus-verb sequence was combined with a direct object
(that often was the same as the embedded subject of the corresponding sentence-complement) which itself was modified by an object-relative clause that was sensible and meaningful in both the full and reduced forms. Thus, each verb was used to create a set of four sentences: a sentence-complement structure and an object-relative structure each with a full and reduced embedded clause constituent. The 48 sets of four sentences were paired into 24 meaning-unrelated sentence-set pairs for prime-target presentation. All critical materials for all experiments are reported in the Appendix. A set of 48 filler sentences and 10 practice sentences that represented a variety of declarative sentence structures were also created and randomly paired into sentence pairs.

Two factors of interest each with two levels were manipulated in counterbalanced fashion within speakers and items: prime type (sentence-complement or object-relative) and prime reduction (full or reduced). Also, target reduction (whether the target sentence complement was originally presented as a full or reduced structure) was counterbalanced across speakers, items, and the factors of interest.

Sixteen presentation lists were created. The form of a given critical sentence on each list was determined by rotating it (in counterbalanced fashion across sentence-set pairs) through (a) whether it was a prime or target sentence (with its pairmate serving the other role), (b) whether the target in its pair was presented as a full or reduced structure, (c) whether the prime in its pair was a sentence complement or an object relative, and (d) whether the prime in its pair was presented as a full or reduced structure. Note that the unit of analysis for the items analyses was the sentence-pair. Thus, for the 24 critical item pairs on each list, speakers saw six pairs in each priming condition (three with targets with full embedded clauses and three with targets with reduced); hence, the two prime-condition factors and the target reduction factor were
manipulated within list, counterbalanced across item-pair. Whether a particular sentence was prime or target was manipulated between lists. The filler sentence pairs were added to each list, resulting in 48 prime-target pairs per list. Four speakers were tested on each list, so that across the experiment, each sentence pair appeared 16 times in each priming condition. The same fixed, randomly generated order of presentation was used for all lists, constrained so that no more than four critical or filler trials and no more than two critical trials from the same condition were presented consecutively.

Scoring and analysis. All prime and target utterances were transcribed. Prime transcriptions were coded either as sentence-complement structures (where the main verb took a clausal complement that could be either full or reduced), object-relative structures (where the direct object was modified by an object-relative clause that could be either full or reduced), other structures (any other kind of structure), or forgotten (where speakers either said “I forgot,” produced nothing, or produced only the main-subject and verb prompt). Target transcriptions were coded either as sentence-complement structures, other structures, or as forgotten. All sentence-complement and object-relative productions were coded for whether they included a full or reduced embedded clause. The proportion of prime and target productions coded into each of these categories is reported in Table 1. Any trial in which the produced prime was coded into a category different from originally presented was excluded from the main analysis, as was any trial where the target was not produced as a sentence-complement structure.

The proportions of sentence complement targets produced with a full embedded clause in each condition for each speaker (across items) and for each item (across speakers) were calculated and entered into analyses of variance (ANOVAs) using speakers \( F_1 \) and items \( F_2 \) as random factors. The production of full versus reduced sentence-complement targets was
assessed with repeated-measures 2 x 2 ANOVAs with prime type and prime reduction as factors. The effects of prime reduction were evaluated separately for sentence complement and object-relative primes with planned comparisons. Variability is reported with 95% confidence-interval halfwidths based on single degree-of-freedom comparisons. All significant effects achieved the .05 level or better. Reported means were calculated across speaker means. Because of excluded trials, 7 speakers did not have at least one observation in one cell of their designs, and so were excluded from the analysis by speakers.

In addition, to assess whether reading a full versus reduced embedded clause in the originally presented target affected embedded clause production, separate three-way repeated-measures ANOVAs were conducted with the additional factor of target reduction, using list and item as random factors (speaker cannot be used as a random factor because of excessive numbers of missing values; also, two items were excluded from the items analysis in Experiment 1 due to missing values). Only significant effects of and interactions with target reduction are reported from these analyses.

Results

Overall performance. Prime and target production performance is shown in Table 1. Speakers mostly produced primes with the same structure as they were originally presented with. The notable exception is that speakers produced full embedded clauses when originally presented with reduced ones, both in sentence-complement primes (151 times or 39.3%) and in relative-clause primes (104 times or 27.1%). The opposite is observed less (37 times or 9.6% with sentence complements, and 41 times or 10.7% with relative clauses). Note that this reflects a bias to produce full embedded clauses generally, supporting the claim that full embedded clauses
are preferred to reduced ones both with sentence-complement and object-relative structures. Targets were nearly always produced as full or reduced sentence-complement structures (though again, speakers were biased to produce full embedded clauses).

--- Table 1 about here ---

*Priming effects.* In all, 885 trials (57.6%) were included in the main analysis. This level of data loss is in line with other investigations of syntactic persistence (e.g., Bock, 1986; Ferreira, in press), and is due to “exuberant responding” (Bock, 1996) – to allow variability in speakers’ responses (i.e., so that speakers can produce full or reduced embedded clauses), the task must permit enough freedom so that sometimes, fully distinct structure types are produced, as are prime structures different from what was originally presented. However, with respect to the current methodology, it has been shown (Ferreira, in press) that the degree of persistence is unrelated to degree of data loss, and so the conclusions below are unlikely to be compromised by this aspect of performance.

The mean percentage of full embedded clauses produced in sentence-complement targets as a function of the different primes with full or reduced embedded clauses is illustrated in Figure 2. The experiment revealed very similar priming effects in both priming conditions: Speakers produced about 16% more full embedded clauses in sentence-complement targets after they produced sentence-complement primes that also included full embedded clauses, compared to after they produced sentence-complement primes with reduced embedded clauses. Similarly, speakers produced about 18% more full embedded clauses in sentence-complement targets after they produced object-relative primes with full embedded clauses, compared to after they produced object-relative primes with reduced embedded clauses. This pattern led to a significant main effect of prime reduction ($F1(1,56) = 22.3, CI = \pm 7.1\%$, $F2(1,23) = 45.2, CI = \pm 7.5\%)$. 
The main effect of prime type was not significant ($F_1(1,56) < 1, CI = ±6.4\%$; $F_2(1,23) < 1, CI = ±7.0\%$), nor was the interaction between these two factors ($F_1(1,56) < 1, CI = ±8.3\%; F_2(1,23) < 1, CI = ±6.6\%$). The effect of prime reduction was significant both within the sentence-complement priming condition ($F_1(1,56) = 14.3; F_2(1,23) = 66.8$) and within the object-relative priming condition ($F_1(1,56) = 18.2; F_2(1,23) = 51.1$).

Supplementary analyses revealed that whether the target was originally presented as full or reduced caused speakers to produce full sentence-complement targets more often ($F_1(1,15) = 46.8, CI = ±6.9\%; F_2(1,21) = 48.1, CI = ±6.4\%$). No interaction with this factor approached significance (all $p$s > 0.10). Also, sentence-complement-target production after primes that were inaccurately produced with full embedded clauses was similar to that after primes that were accurately produced with full embedded clauses. Of the 151 sentence-complement primes produced with full embedded clauses when originally presented with reduced ones, 130 were followed by the production of sentence-complement targets. Of these, 106 (82\%) were produced as full sentence-complements. Similarly, of the 104 object-relatives produced with full embedded clauses that were originally presented with reduced ones, 85 were followed by the production of a sentence-complement target, and of these, 67 (79\%) were full sentence-complements.

Discussion

Experiment 1 revealed that the mention of full versus reduced embedded clauses in sentence-complement targets was influenced by full versus reduced embedded clauses in object-relative primes just as much as it was influenced by full versus reduced embedded clauses in sentence-complement primes. This suggests that embedded clauses show persistence.
independently of whether they are produced in a syntactically similar or dissimilar context, since object-relatives differ syntactically from sentence-complements in (at least) two ways. First, the embedded clause in a sentence-complement structure is an argument of the main verb, whereas the embedded clause in an object-relative structure is a (non-argument) modifier of a noun. Second, a (declarative) sentence-complement structure does not involve any movement or trace government, whereas an object-relative clause structure does (at the object position of the relative clause verb). It is especially revealing to contrast the effectiveness of the object-relative primes observed here with the ineffectiveness of the noun-complement primes (“The theory that penguins built the igloos was completely false”) from Ferreira (in press) described above. If anything, noun-complement structures are more syntactically similar to sentence-complement structures than object-relative structures are: Noun-complements, like sentence-complements (but unlike object-relatives), do not involve any movement or trace government, and the that in noun-complements, like the that in sentence-complements (but again unlike in object-relatives), cannot be analyzed as a relative pronoun. Nevertheless, the object-relatives primed sentence-complement production as much as sentence-complements did, but the noun-complements did not prime sentence-complement production at all.

The fact that embedded clauses can show syntactic persistence despite differences in overall syntactic similarity suggests that the ineffectiveness of noun-complement primes shown by Ferreira (in press) cannot be explained by appealing to overall syntactic similarity. Instead, the explanation that remains for the pattern shown in Figure 1 follows from an entrenchment-based learning account: Noun-complement primes were ineffective because they include full embedded clauses and not reduced embedded clauses, and full embedded clauses cause less syntactic persistence than reduced embedded clauses do.
If indeed the persistence of an embedded clause is inversely proportion to its preferred status per se, then we should observe that if speakers produce generally less preferred reduced embedded clauses in specific syntactic contexts in which they are more preferred, their priming effectiveness should diminish accordingly. Experiment 2 was designed to explore this prediction by investigating the production of sentences that have reduced embedded clauses that are as preferred as possible. In English, there is one structural subtype where the reduced embedded clause is (at least nearly) categorically preferred: In an interrogative sentence-complement structure in which the subject of the embedded-clause is questioned (e.g., “Who did the teacher notice was visiting the school nurse?” hereafter a subject-extracted sentence-complement interrogative), not only is a reduced embedded clause complement preferred, it is pretty much obligatory (as most consider “Who did the teacher notice that was visiting the school nurse?” to be ungrammatical; this is sometimes called the that-trace effect, referring to the ungrammaticality of a that followed by a trace or moved-from location; see Perlmutter, 1971). If it is true that reduced embedded clauses prime effectively because less preferred structures exert a greater persistence effect, then when speakers produce subject-extracted sentence-complement interrogative primes, in which a reduced embedded clause is more preferred, priming effects should diminish or disappear.

Experiment 2

Speakers again produced (declarative) full or reduced sentence-complement targets, like (8) above, in two different priming conditions: In one priming condition, the immediately preceding prime was an object-extracted sentence complement interrogative (sentence-
complement structures that questioned the object of the embedded-clause verb) with a full or reduced embedded clause, like (13) and (14):

(13) Who did the teacher notice that the student was visiting?

(14) Who did the teacher notice the student was visiting?

Note that because such object-extracted sentence-complement interrogatives can be produced with either full or reduced embedded clauses, there is no reason to believe that reduced embedded clauses are any differently preferred in such structures as they are in the declarative sentence-complements and object-relative structures already tested.

In the other priming condition, primes were subject-extracted sentence-complement interrogatives with reduced embedded clauses like (15):

(15) Who did the teacher notice was visiting the school nurse?

Subject-extracted sentence-complement interrogatives are ungrammatical with full embedded clauses (“*Who did the teacher notice that was visiting the school?”), and thus speakers should (perhaps) categorically prefer to produce reduced rather than full embedded clauses in such structures (note that recall performance of these structures as prime sentences can provide an indication of the degree of preference for production of a reduced embedded clause constituent).

Note that to evaluate priming from the subject-extracted sentence-complement interrogatives, primes with reduced embedded clauses cannot be compared against corresponding structures with full embedded clauses (“*Who did the teacher notice that was visiting the school nurse?”); this would require that speakers produce ungrammatical that-trace structures, and thereby defeat the very linguistic effect that the experiment exploits to make its prediction. Instead, in Experiment 2, subject-extracted sentence-complements with reduced embedded clauses were compared to wh-in-situ structures with full embedded clauses, like (16):
The teacher noticed that who was visiting the school nurse?

The wh-in-situ structures were chosen because they carry a similar overall meaning as the subject-extracted sentence-complement primes and do not include a reduced embedded clause. Nevertheless, the wh-in-situ structures are not ideal controls for the subject-extracted sentence-complement primes, both because of the surface differences between the two sentence types (the subject-extracted sentence-complement primes begin with a wh-word, the wh-in-situ primes do not), and because it is unclear how preferred a full embedded clause is in such structures (at least, both a full or reduced embedded clause complement is possible). Because of these concerns, Experiment 3 evaluated the priming effect of the subject-extracted sentence-complement interrogatives against a different control.

The results of Experiment 1 suggested that independent of syntactic similarity, reduced embedded clauses cause robust syntactic persistence. If this robust persistence is due to the fact that reduced embedded clauses are less preferred syntactic structures, then subject-extracted sentence-complement primes, which prefer reduced embedded clauses more, should cause a smaller priming effect. Indeed, because reduced embedded clauses are effectively required in the subject-extracted sentence-complement primes (i.e., reduced embedded clauses may be completely preferred), it may be that after such primes, no priming effect will manifest at all.

**Method**

*Speakers.* Sixty-four members of the same population as Experiment 1 were tested.

*Apparatus and procedure.* Apparatus and procedure were as in Experiment 1.

*Materials and design.* The sentence-complement structures from Experiment 1 were changed so that their embedded clauses were transitive. Each of these was then used to create a
set of six sentences: declarative sentence-complement structures with full and reduced embedded clauses like (8) above, and four interrogative sentence-complement structures like (13–16). Sets were paired into meaning-unrelated sentence-set pairs, as in Experiment 1. Filler and practice sentences were modified to also include interrogative structures (so that interrogative sentences generally were not restricted to the prime position of the critical trials).

The design was the same as Experiment 1’s, except that the sentence-complement and object-relative conditions were replaced with object-extracted and subject-extracted sentence-complement interrogatives. Sixteen presentations lists were created in the same way as in Experiment 1. The same fixed, randomly generated order of presentation was used as in Experiment 1.

Scoring and analysis. Primes and targets were transcribed and coded. Declarative sentence-complement targets were coded as in Experiment 1. For the interrogative primes, an utterance was coded as an object- or subject-extracted sentence-complement if it was a sentence-complement structure (i.e., it contained a main verb with a clausal argument), with a wh-element in the sentence-initial position and a gap in the object or subject position for object- and subject-extracted forms respectively. An utterance was coded as a subject-extracted full embedded clause (wh-in-situ) structure if it was a sentence-complement structure with a wh-element after the main verb. Other and forgot codings were determined as in Experiment 1. All declarative and interrogative sentence-complements were coded for whether they included a full or reduced embedded clause. The coding of prime and target utterances is shown in Tables 2 and 3. Exclusions were determined with the same criteria as Experiment 1.

The dependent variable was calculated and statistical analyses were conducted as in Experiment 1. Because of excluded trials, 8 subjects were left with a missing value in their
designs, and so were excluded from the analysis by speakers. Also, due to a stimulus design error, 1 to 3 sentences were presented to 30 of the speakers with an incorrect main verb (sense), affecting 64 of 1536 trials across the experiment. Productions resulting from these presentations were treated as other productions (the results are no different if they are excluded).

Results

Overall performance. Tables 2 and 3 show production performance for all prime and target trials. For primes, speakers nearly always produced sentences with the same structure as originally presented. The only exception is that in keeping with the expectation that reduced structures are generally less preferred, speakers produced object-extracted sentence-complement primes with full embedded subjects when originally presented with reduced ones on 100 trials (26.0%). Notably, speakers produced subject-extracted sentence-complements presented with reduced embedded clauses (i.e., the that-trace structures) quite accurately, as in such structures, speakers produced full embedded clauses only 16 times (4.2%). Targets were nearly always produced as full or reduced sentence-complement structures.

--- Tables 2 and 3 about here ---

Priming effects. Overall, 1072 trials (69.8%) were analyzable. For these trials, the percentages of full sentence-complement targets that speakers produced as a function of prime type and prime reduction are shown in Figure 3. While the object-extracted sentence-complement primes yielded an effect that was as big as the effects observed in Experiment 1 (20%), the subject-extracted sentence-complement primes yielded an effect that was only half as big (11%). These differences led to a significant effect of prime reduction ($F_1(1,55) = 19.2, CI = \pm7.0\%$; $F_2(1,23) = 28.8, CI = \pm6.5\%$) and an effect of prime type that was marginally significant.
only by speakers \((F1(1,55) = 3.18, p < .09, CI = ±5.2\%; F2(1,23) < 1, CI = ±7.0\%)\). The interaction between prime type and prime reduction – the difference between the two priming effects – was marginally significant \((F1(1,55) = 3.53, p < .07, CI = ±7.1\%; F2(1,23) = 3.82, p < .07, CI = ±7.6\%)\). The effect of prime reduction was significant both within the object-extracted sentence-complement priming condition \((F1(1,55) = 31.9; F2(1,23) = 35.6)\) and within the subject-extracted sentence-complement priming condition \((F1(1,55) = 8.95; F2(1,23) = 10.2)\).

--- Figure 3 about here ---

The supplemental analyses revealed that speakers were more likely to produce full sentence-complement targets when originally presented with full sentence-complement targets \((F1(1,15) = 51.0, CI = ±8.7\%; F2(1,23) = 87.7, CI = ±6.1\%)\), and that this factor interacted with prime reduction \((F1(1,15) = 14.6, CI = ±8.1\%; F2(1,23) = 15.1, CI = ±8.7\%)\). The latter interaction reflects a ceiling effect: When the presented target was a full sentence-complement structure, speakers’ overall level of full sentence-complement production was high enough that prime reduction exerts a smaller effect (88% vs. 83%) than it does when the presented target was a reduced sentence-complement structure (69% versus 44%). Importantly, however, target reduction did not interact with prime type \((F1(1,15) = 2.18, CI = ±8.6\%; F2(1,23) < 1, CI = ±7.1\%)\), nor did the three-way interaction between prime type, prime reduction, and target reduction approach significance \((F1(1,15) < 1, CI = ±9.5\%; F2(1,23) = 1.3, CI = ±10.7\%)\).

Thus, target reduction does not differentially affect the two priming conditions. Finally, as in Experiment 1, targets were produced as full sentence-complements about as often when the full embedded clause in the prime was inaccurately recalled as when it was accurately recalled: Of the 100 object-extracted sentence-complement primes that were produced with full embedded clauses when originally presented with reduced ones, 81 were followed by the production of a
sentence-complement target, and of these, 64 (79%) were produced with full embedded clauses.

Discussion

Experiment 2 sought to determine whether the priming effectiveness of a reduced embedded clause is sensitive to its degree of preference in its specific production context. It is first important to note that the production performance of the two types of interrogative primes themselves suggests that they indeed differ in their degree of preference for full versus reduced embedded clause complements. When producing object-extracted sentence-complement interrogative primes, speakers were quite likely to erroneously recall a reduced embedded clause as a full embedded clause (26%); this reflects the already noted preference for producing full rather than reduced embedded clauses. However, when producing subject-extracted sentence-complement interrogative primes, speakers were quite unlikely to produce full instead of reduced embedded clauses (4%). This suggests that speakers are very weakly biased to produce full embedded clauses in subject-extracted sentence-complement interrogatives, confirming that reduced embedded clauses are relatively more preferred in such structures.

As predicted, the priming effectiveness of the two kinds of primes followed this degree of preference. The object-extracted sentence-complement primes, in which reduced embedded clauses are less preferred, yielded a robust priming effect (20%), whereas the subject-extracted sentence-complement primes, in which reduced embedded clauses are more preferred, yielded a numerically smaller priming effect (11%). This pattern confirms the expectation of learning-based approaches to syntactic persistence that the degree of priming that is observed from a structure should be inversely proportional to its degree of preference, since the same syntactic constituent (i.e., a reduced embedded clause) causes a smaller priming effect when it was more
preferred than when it was less preferred. It should be noted, however, that the interaction effect that supports this interpretation just failed to reach conventional levels of significance by speakers and by items. Experiment 3, as a replication of Experiment 2, provides another opportunity to evaluate the reliability of this interaction, and moreover, the results of Experiments 2 and 3 can be combined to gain an especially powerful test.

It is also notable that despite the difference in the amount of priming in the two conditions, the priming effect from the subject-extracted sentence-complement primes was still significant (as assessed by the planned comparison). This latter effect is especially interesting in light of the results from Ferreira (in press) described above, whereby noun-complement primes ("The theory that penguins built the igloos was false"), in which the production of a full embedded clause is obligatory, yielded no priming effect whatsoever (a 3.3% difference in the opposite direction). Together with Experiment 2 here, this suggests that when speakers produce a prime in which a full embedded clause is obligatory, priming is eliminated, but when speakers produce a prime in which a reduced embedded clause is obligatory, priming may be reduced but it is not eliminated. If this difference is real, it should appear again in Experiment 3.

Experiment 3

Experiment 3 replicated Experiment 2 with one change: Instead of \textit{wh}-in-situ controls, reduced subject-extracted sentence-complement primes (e.g., \textit{that}-trace structures like (15) above, repeated as (17)) were compared to simplex interrogatives with the same \textit{wh}-word, main subject, and main verb, like (18):

(17) Who did the teacher notice visited the school nurse?
(18) Who did the teacher notice visiting the school nurse?

Note that simplex interrogatives like (18), again while not ideal controls for the *that*-trace structures, carry pros and cons complementary to those of the *wh*-in-situ structures tested in Experiment 2. Whereas the *wh*-in-situ structures began differently from the *that*-trace structures, the simplex interrogatives began identically, up to and including the first verb. Also, the degree of preference for a full or reduced embedded clause in simplex structures is known, since neither a full nor a reduced embedded clause can be produced at all. Furthermore, there may be a general effect on full versus reduced sentence-complement targets when a prime includes a sentence-complement-taking verb without a following *that* (e.g., whenever *notice* is produced without a following *that* – in a sentence-complement structure or otherwise – the degree of *that*-mention in a subsequent sentence-complement structures may be depressed independent of any effect of producing the reduced embedded clauses per se). If so, then the simplex interrogative structures should control for such effects. In contrast, unlike the *wh*-in-situ controls, the simplex interrogative structures are less similar in overall meaning and structure to the *that*-trace structures that they were compared to. Overall, the intention is that the different controls from Experiments 2 and 3, when combined and contrasted, should provide a complete picture of the priming effectiveness of the *that*-trace structures.

If the preference to produce reduced embedded clauses in *that*-trace structures indeed mitigates their priming effectiveness, then Experiment 3, like Experiment 2, should reveal that subject-extracted sentence-complements influence the production of sentence-complement targets less than the object-extracted sentence-complements do.

*Method*
Speakers. Sixty-four members of the same population as Experiment 1 were tested.

Apparatus and procedure. Apparatus and procedure were as in Experiment 1.

Materials and design. The materials were the same as in Experiment 2, with the exception of the full subject-extracted sentence-complement primes. These were replaced with simplex interrogatives that began with the same wh-word, main subject, and main verb as their matched reduced subject-extracted sentence-complement primes, but were then followed by non-clausal material (like (18) above). As much of the embedded clause lexical content from the reduced subject-extracted sentence-complement interrogative was used also in the simplex interrogative structure, so that the overall meaning and superficial forms of the two sentences were as similar as possible.

The design corresponded to that in Experiment 2. Because the subject-extracted sentence-complement priming condition in Experiment 3 does not include a full embedded clause structure, the prime reduction factor from Experiment 2 is described in Experiment 3 in terms of whether or not the prime includes a reduced embedded clause. Lists and presentation scheme were as in Experiment 2.

Scoring and analysis. Scoring and analyses were as in Experiment 2. A production was coded as a simplex interrogative if it included the correct wh-word, main subject and main verb, and then continued without an embedded clause complement. Prime and target codings are shown in Tables 4 and 5 below. Exclusions were determined as in Experiments 1 and 2.

Dependent variables and statistical analyses were measured and conducted as in Experiments 1 and 2. Because of excluded trials, two speakers were left with missing values in at least one cell of their designs; these speakers were excluded from the analyses by speakers.
Results

Overall performance. Production performance on prime and target sentences is shown in Tables 4 and 5. Overall performance was similar to that in Experiment 2. For primes, overall recall accuracy was good, though speakers were less accurate after object-extracted sentence-complement primes than after subject-extracted ones (139 sentences were shifted between the full- to reduced embedded-clause categories for the former, but only 20 for the latter; note that the simplex interrogatives are extremely unlikely to be erroneously recalled with a reduced embedded clause). For target production, speakers nearly always produced full or reduced sentence-complement structures.

--- Tables 4 and 5 about here ---

Priming effects. Overall, 1111 trials (72.3%) were analyzable. For these, the percentages of full sentence-complement targets produced as a function of prime-type and whether the prime included a reduced embedded clause is shown in Figure 4. The pattern is the same as in Experiment 2: Whereas a sizable priming effect is observed after object-extracted sentence-complement primes (17%), a smaller effect is observed after subject-extracted sentence-complement primes (6%). ANOVAs revealed a significant main effect (marginal by items) of whether the prime was in the object- or the subject-extracted sentence-complement condition \((F1(1,61) = 6.12, CI = \pm 4.8\%; F2(1,23) = 3.22, p < .09, CI = \pm 6.3\%)\) and a significant main effect of whether the prime included a reduced embedded clause \((F1(1,61) = 16.1, CI = \pm 5.6\%; F2(1,23) = 34.6, CI = \pm 4.5\%)\). The interaction between priming condition and whether the prime included a reduced embedded clause was marginally significant by speakers and significant by items \((F1(1,61) = 3.62, p < .07, CI = \pm 7.7\%; F2(1,23) = 5.15, CI = \pm 7.9\%)\). The effect of whether the prime included a reduced embedded clause was significant within the object-
extracted sentence-complement priming condition ($F_1(1,61) = 18.1; F_2(1,23) = 24.8$), but not within the subject-extracted sentence-complement priming condition (just marginally significant by items; $F_1(1,61) = 2.46; F_2(1,23) = 3.15, p < .09$).

--- Figure 4 about here ---

Supplemental analyses revealed that more full sentence-complement targets were produced when the originally presented target was a full sentence-complement ($F_1(1,15) = 76.6$, CI = ±8.4%; $F_2(1,23) = 107$, CI = ±6.9%), and that this interacted with whether the prime included a reduced embedded clause ($F_1(1,15) = 7.89$, CI = ±9.2%; $F_2(1,23) = 14.1$, CI = ±6.8%). As in Experiment 2, this reflects a ceiling effect, whereby full sentence-complement targets were produced so often when full sentence-complements were presented that the effect of whether the prime included a reduced embedded clause was small (87% vs. 83%), relative to when reduced sentence-complement targets were presented (61% vs. 40%). Again, however, target reduction did not interact with the prime sentence-type ($F_1(1,15) < 1$, CI = ±5.9%; $F_2(1,23) < 1$, CI = ±9.4%), and the three-way interaction of target reduction, whether the prime included a reduced embedded subject, and prime sentence-type was marginally significant only by lists ($F_1(1,15) = 3.65, p < .08$, CI = ±9.0%; $F_2(1,23) < 1$, CI = ±10.2%). Thus, the effect of target reduction does not appear to differentially affect the object- versus the subject-extracted sentence-complement primes. Finally, the production of full versus reduced sentence-complement targets was similar after inaccurately recalled primes as after accurately recalled primes. Of the 83 object-extracted sentence-complement primes produced with full embedded clauses when originally presented with reduced ones, 66 were followed by the production of sentence complement targets, of which 53 (80%) were produced as full sentence-complement targets. Similarly, of the 56 object-extracted sentence-complement primes produced with
reduced embedded clauses when originally presented with full ones, 42 were followed by the
production of a sentence-complement targets, of which 23 (55%) were produced as full sentence-
complement targets.

_Omnibus analyses of Experiments 2 and 3._ In Experiment 3, the interaction between
prime sentence-type and whether a prime included a reduced embedded clause was significant by
items, but it just failed to reach significance by speakers. Furthermore, the priming effect within
the subject-extracted sentence-complement priming condition was not significant in Experiment
3, unlike in Experiment 2. However, because Experiment 3 was a replication of Experiment 2,
the data from each can be combined into a single set of analyses. Collapsed across experiment,
the mean percentages of full sentence-complement targets produced as a function of prime
sentence-type and whether the prime included an embedded clause are shown in Figure 5. The
pattern of results looks similar to that in Experiments 2 and 3: The difference in the object-
extracted sentence-complement priming condition (18%) is about twice as big as that in the
subject-extracted sentence-complement priming condition (8%). Speaker and item ANOVAs
were conducted with the additional between-speaker factor _experiment_ (note that the use of a
between-speaker factor makes the calculation of a within-speaker confidence interval impossible,
and so variability is indicated with mean squared-error values instead). These revealed that the
main effect of whether a prime included a reduced embedded clause was significant (_F_1(1,116) =
35.8, MSE = 0.05832; _F_2(1,46) = 60.3, MSE = 0.01755), whereas the main effect of prime
sentence-type was not (_F_1(1,116) < 1, MSE = 0.03694; _F_2(1,46) < 1, MSE = 0.02495). Most
important is that the interaction between prime sentence-type and whether the prime included a
reduced embedded clause was significant (_F_1(1,116) = 7.05, MSE = 0.04109; _F_2(1,46) = 8.94,
MSE = 0.01682). The effect of whether the prime included a reduced embedded clause was
significant both within the object-extracted sentence-complement priming condition \((F_1(1,116) = 47.8; F_2(1,46) = 59.7)\) and within the subject-extracted sentence-complement priming condition \((F_1(1,116) = 9.99; F_2(1,46) = 12.2)\). Finally, the only effect of the experiment factor that was significant (marginally by items) was the interaction between experiment and prime sentence-type \((F_1(1,116) = 8.96, \text{MSe} = 0.03694; F_2(1,46) = 3.40, p < .08, \text{MSe} = 0.02495)\). This reflects the fact that in Experiment 2, speakers produced sentence-complement targets with more full embedded clauses in the subject- than in the object-extracted sentence-complement priming condition (72% vs. 69%), whereas the reverse was true in Experiment 3 (66% vs. 72%). This happened because speakers produced more full sentence-complement targets after the \(wh\)-in-situ structures that were used in the subject-extracted sentence-complement priming condition of Experiment 2 (which included full embedded clause constituents) than after the simplex interrogative structures used in the subject-extracted condition in Experiment 3 (which did not include full embedded clause constituents). Note that this implies that full embedded clauses in interrogative sentence complement structures cause some syntactic persistence (which, as will be shown, is consistent with the entrenchment-based-learning account; it is embedded clauses in noun-complement structures that should show negligible syntactic persistence).

--- Figure 5 about here ---

Discussion

Experiment 3 revealed the same pattern of results as Experiment 2: When reduced embedded clauses were produced in primes in which they were more preferred (the \(that\)-trace structures), the priming effect was smaller than when they were produced in primes in which they were less preferred (the object-extracted sentence-complement interrogatives). The
interaction supporting this claim was marginally significant by speakers and significant by items in Experiment 3, and was significant by speakers and items in the omnibus analysis of Experiments 2 and 3. This result thus supports the claim that the priming effectiveness of the reduced embedded clauses derives from their less preferred status. Furthermore, the second result of interest from Experiment 2, namely, the smaller but still significant priming effect that was observed from the embedded-clause preferring subject-extracted sentence-complement interrogatives, was corroborated by the results of the combined analysis of Experiments 2 and 3. This unexpected result is discussed further in the General Discussion.

General Discussion

These experiments yielded three primary observations. First, Experiment 1 showed that when speakers produced full versus reduced embedded clauses, object-relative primes influenced the production of sentence-complement targets just as much as sentence-complement primes did. This implies that syntactic persistence is not especially sensitive to the similarity of the syntactic context of the primed constituents between the prime and target sentences, since the embedded clause in an object-relative structure occurs in a different syntactic context from the embedded clause in a sentence-complement structure. Second, Experiments 2 and 3 showed that reduced embedded clauses were less effective primes when produced as parts of structures in which they were more preferred (subject-extracted sentence-complement interrogatives – that-trace structures) than when they were produced as parts of structures in which they were less preferred (comparable object-extracted sentence-complement interrogatives). This is predicted by entrenchment-based learning accounts of syntactic persistence, since such accounts predict that
the degree to which a syntactic constituent shows persistence is an inverse function of its production preference (i.e., the message-to-syntax mapping’s degree of entrenchment). Finally, even though reduced embedded clauses were categorically preferred in the that-trace structures tested in Experiments 2 and 3, a significant (but reduced) persistence effect was yet observed. This contrasts with prior results (Ferreira, in press), in which it was observed that when a full embedded clause was categorically preferred, no syntactic persistence effect was observed at all. I discuss below why this unexpected result was observed.

The observation that syntactic persistence can occur despite syntactic differences between prime and target sentences (the results of Experiment 1) carries important implications for any account of syntactic persistence. At minimum, it implies that the processing mechanisms that drive the persistence effect do not inherently reflect the relational structure that is a defining feature of syntactic representations. For example, Pickering and Branigan (1998) attribute syntactic persistence to residual activation of what they term combinatorial nodes. Within this account, the present data imply that such combinatorial nodes cannot be defined in terms of relational structure either with respect to the primed constituents’ syntactic context (since here, it was observed that full and reduced embedded clauses showed persistence independent of whether they were arguments of a verbs or modifiers of nouns) or with respect to the internal structure of the primed constituents themselves (as suggested by the results of Bock, 1986 and Fox Tree & Meijer, 1999). Of course, this does not imply that syntactic structures generally are non-relational in nature (see, e.g., Franck, Vigliocco, & Nicol, 2002 for a recent relevant observation); it only implies that syntactic persistence is not especially sensitive to any relational structure that syntactic structures possess.

Furthermore, any syntactic persistence account must also explain the sensitivity of the
persistence effect to the degree of preference for the primed constituent in the prime sentence. Outside of an entrenchment-based learning account, this sensitivity could be explained with some additional assumptions. Specifically, if it is assumed that less preferred structures (e.g., reduced embedded clauses) are more difficult to access for production than more preferred structures (e.g., because of asymmetric competition between the structures; see Ferreira, 1996 vs. Stallings, MacDonald, & O'Seaghdha, 1998 for different views on this point), then it may be useful to maintain the accessibility of less preferred structures more than the accessibility of more preferred structures. Thus, the present results support only the class of syntactic persistence models that make these kinds of assumptions so that they can account for the observed sensitivity.

Turning to learning-based accounts, an approach like that proposed by Chang et al. (2000) can naturally explain the observations described above. Most important is the observation that less preferred syntactic options cause greater syntactic persistence than more preferred syntactic options, as reflected by the smaller effect in Experiments 2 and 3 of reduced embedded clauses when they were produced in primes in which they were more preferred. This can be straightforwardly understood as a consequence of entrenchment-based learning, because a message structure expressed by a less preferred syntactic option should be less entrenched and thus subject to greater syntactic persistence than a message structure expressed by a more preferred syntactic option. Thus, these results support the claim that message-to-syntax knowledge generally derives from an entrenchment-based mechanism.

Furthermore, this entrenchment-based effect can also explain why uncommon syntactic structures remain viable members of speakers’ syntactic vocabularies. That is, these observations suggest that less common syntactic options remain part of speakers’ grammars
because those less common options, when used, exert a more marked learning or strengthening effect on the language-processing system. It worth noting that the kind of error-based learning that leads to entrenchment effects was originally motivated on acquisition grounds – that learning-based models learn better when the more accurate outputs they generate cause smaller knowledge changes and the less accurate outputs cause greater knowledge changes. To this, the present results add a functional justification relevant to adult (i.e., steady-state) performance: Less preferred structures affect the representation of syntactic knowledge more than more preferred structures do, allowing those less preferred structures to remain viable members of the relevant mechanism’s output vocabulary. Note that it is important for a number of reasons that uncommon syntactic structures be available for production. The existence of uncommon structures allows for greater syntactic diversity, which in turn affords a more nuanced communication process (since different structures can be used to communicate different meanings; e.g., Goldberg, 1995). Furthermore, Ferreira (1996) showed that syntactic diversity leads to faster and more error free production. The promotion of uncommon structures by an entrenchment-based learning mechanism allows for these language-processing benefits.

The observation from Experiment 1 that syntactic persistence can be observed in spite of syntactic differences between the prime and target sentences follows from the fact that learning-based models will generalize across syntactic details to map similarly structured messages onto similar classes of syntactic structures. In concrete terms, the idea is that full or reduced embedded clauses express embedded propositions in a similar manner across a variety of syntactic contexts. This implies that producing a reduced embedded clause to express an embedded proposition as a modifier of a noun (e.g., in an object-relative structure) is similar in critical respects to producing a reduced embedded clause to express an embedded proposition as
an argument of a verb (e.g., in a sentence-complement structure). As a result, the mapping itself between the embedded proposition message-structure and the reduced embedded clause syntactic structure is defined at least partially independently of the syntactic difference between being a modifier of a noun and being an argument of a verb. In short, to the extent that message-to-syntax mappings are similar across multiple syntactic contexts, the more those mappings will be defined independently of the specific syntactic details that vary among those contexts, and thus the more that syntactic persistence will also be independent of those same syntactic details. Of course, syntactic persistence should be sensitive to some syntactic features; after all, one end of the message-to-syntax mapping is syntactically defined. The point here is that syntactic persistence will not necessarily be sensitive to all syntactic features of a particular structure, but rather, it should be sensitive only to those syntactic features that are systematically related to the message-structure distinctions that trigger the use of particular families of syntactic structures.

The other observation noted above was that categorically preferred reduced embedded clauses (e.g., reduced embedded clauses in that-trace structures) yielded still significant (though reduced) syntactic persistence effects. This is in contrast to what Ferreira (in press) observed with categorically preferred full embedded clauses (in noun-complement structures), which yielded no syntactic persistence effect at all. One way to explain this observation is to claim that the overall preference for a syntactic option is jointly determined by its preference in its particular syntactic context as well as its preference across the entire language. That is, in noun-complement structures, the specific (categorical) preference to produce full embedded clauses over reduced ones combines with the general preference across the entire language to also produce full embedded clauses instead of reduced ones. The net effect is a double pressure to produce full embedded clauses, implying that the mapping will be very strongly entrenched and
thereby lead to little or no syntactic persistence. In contrast, in *that*-trace structures, the specific (categorical) preference to produce reduced embedded clauses will compete with the preference across the language to produce full embedded clauses. The net effect of these two opposing forces is a diminished pressure to produce reduced embedded clauses (compared to what happens in the other direction with noun-complement structures), and thus relatively less entrenchment, leading to an intermediate but still observable syntactic persistence effect.

In fact, a learning-based model of syntactic persistence like that proposed here and in Chang et al. (2000) will naturally exhibit this pattern, for the same reason that it shows persistence independent of overall syntactic similarity. That is, given that a mapping to a given kind of embedded clause is similar in many respects in all of its syntactic contexts, the properties of the mapping to an embedded clause in any specific syntactic context will reflect a combination of the properties of the mappings to that constituent in all its syntactic contexts. Thus, the properties of a mapping to a reduced embedded clause in the context of a *that*-trace structure will reflect the properties of mappings to reduced embedded clauses in all kinds of syntactic structures, and so the categorical bias for reduced embedded clauses that comes from the *that*-trace context will be diminished by the bias against reduced embedded clauses that exists across the language as a whole. This will lead to only a moderately biased mapping, and thus an intermediate syntactic persistence effect. Similarly, the properties of a mapping to a full embedded clause in the context of a noun-complement structure will reflect properties of mappings to full embedded clauses in all syntactic structures, and so the categorical bias favoring a full embedded clause in a noun-complement structure will be buttressed by the bias favoring full embedded clauses across the language as a whole, leading to an especially biased mapping, and thus especially weak syntactic persistence.
Another possibility, however, is that the *that*-trace structures in Experiments 2 and 3 showed a small persistence effect because of some unique aspect of the internal syntactic structure of the embedded clause in such constructions. While this explanation is in principle possible, it does run against the observation noted above from Fox Tree and Meijer (1999) that the internal structure of a constituent had no measurable influence on syntactic persistence. Furthermore, across Experiments 1–3, the internal structure of the embedded clauses varied widely, yet no systematic effect of that internal structure appeared relevant to the persistence of full or reduced embedded clauses. For example, embedded clause structures that were part of object-extracted interrogative structures or object-relative structures showed as much persistence as declarative sentence-complement structures, despite the fact that the former kinds of structures included traces whereas the latter did not. These observations imply that there is little reason to expect that *that*-trace structures showed a small persistence effect because of something about the internal constitution of the embedded clause constituent in such structures. In contrast, the learning-based explanation outlined above follows naturally from the assumptions that were made to account for other aspects of the persistence effect (i.e., that the persistence effect as a whole is not inherently sensitive to syntactic similarity). Together, this suggests that based on the current evidence at least, the learning-based explanation is a more satisfactory one.

It is worth noting that an interesting consequence of entrenchment is that it gives the notion of linguistic alternation (e.g., active vs. passive, prepositional dative vs. double-object dative; full vs. reduced embedded clause) a processing basis and an observable consequence. That is, linguistic alternations arise when a common underlying message structure (e.g., what might be considered a single d-structure in transformational theories of syntax) can give rise to more than one syntactic form (e.g., more than one s-structure). Note that in a learning-based
model, where alternations exist there will necessarily be less entrenchment than where
alternations do not exist, since an alternation requires that a given message structure be
sometimes mapped to one structure and sometimes mapped to another. As a result, where there
is an alternation and thus less entrenchment, syntactic persistence is more likely to be observed,
whereas where there is not an alternation and thus more entrenchment, syntactic persistence is
less likely to be observed.

Of course, this notion of alternation begs an important question of whether distinct
syntactic alternatives (e.g., actives vs. passives) differ also in meaning. If so, then the message-
structure inputs that trigger the use of alternative forms should differ, and thus no alternation
actually existed in the first place. Two points are relevant to this observation. First, whereas
actives and passives may differ in meaning, it is far less clear whether structures with full versus
reduced embedded clauses do (see Ferreira & Firato, 2002). Second, even if alternative
structures differ in meaning, it is unclear whether (a) that is true only for comprehenders and not
for speakers (comprehenders may infer meaning differences that speaker do not necessarily
represent), and (b) whether such meaning-level distinctions are always sufficient to determine
which structure a speaker uses. Concerning the latter point, if instead the meaning-level
differences between syntactic alternatives exert only a probabilistic influence over which
structure a speaker produces, then a kind of quasi-alternation (at least) follows, which is enough
to underlie the relationship between alternation and syntactic persistence noted above.

A Learning-Based Model of Embedded Clause Production

To bear out the theoretical claims made in this paper, a simple learning-based model was
implemented to simulate the results of the present experiments. The model, illustrated in Figure 6, was based directly on the model in Chang et al. (2000), simplified to demonstrate just the critical points described above. The goal was to show that when the assumptions regarding entrenchment-based syntactic persistence are instantiated in a computational mechanism, the observed pattern of results follows.

--- Figure 6 about here ---

The model was designed to map inputs, in this case message structures, onto outputs, in this case syntactic representations. The set of message-to-syntax mappings the model was trained on was chosen to represent the kinds of sentence types that speakers produced in the experiments above and in Ferreira (in press), namely, declarative sentence-complement structures, noun-complement structures, object-relative structures, object-extracted sentence-complement interrogatives, subject-extracted sentence-complement interrogatives, and simple transitives. The model was given a message structure that corresponded to each of these kinds of sentence types (i.e., one message structure corresponded to declarative sentence-complements, another corresponded to noun-complement structures, and so forth), and trained to activate the appropriate syntactic output (a full or reduced embedded clause, or a noun-phrase argument).

The patterns that the model was trained on are illustrated in Figure 7. The input to the model, shown at the left of the figure, is a distributed representation of the message structure of the to-be-produced constituent. For example, the first message-structure input shown in Figure 7 (the first two rows) represents a declarative sentence-complement message-structure, which should cause the model to produce a (full or reduced) embedded clause; the last message-structure input shown in Figure 7 (the last row) represents a transitive message-structure, which should cause the model to produce a noun phrase. It is important to note that a given message
structure varies in a one-to-many manner with possible syntactic outputs. This represents the above-noted notion of syntactic alternation – the idea that a given message structure can be expressed with more than one syntactic structure.

--- Figure 7 about here ---

These distributed representations of message structures were defined in the following way. First, each kind of message-structure was characterized in terms of a set of message-structure features, shown in Table 6. Specifically, each message structure took particular values of mood, element type, subordinating element, and bound element message-structure features. The mood feature represented whether the message-structure represented a declarative meaning (e.g., declarative sentence-complement structures), an interrogative meaning (e.g., interrogative sentence-complement structures), or was a modifier (e.g., relative-clause structures). The element-type feature represented whether the message structure represented a proposition (as in any structure that takes an embedded clause) or an argument (as in a simple transitive structure). The subordinating-element feature represented whether the message structure represented an element of an action (as when an embedded clause or noun phrase was an argument of a verb) or of an entity (as when an embedded clause was an argument of or a modifier of a noun). Finally, the bound-element feature represented whether some element represented within the message-structure was identical to another (here unspecified) element of the message, and thus should be expressed with a gap (i.e., it should not be phonologically realized). The bound-element feature represented either that the agent or patient of the represented proposition should was a bound element (as in a subject-extracted and object-extracted sentence-complement interrogative respectively), or that neither the agent nor the patient was (and thus that both should be realized, as in any declarative sentence-complement structure). Generally, the idea was to instantiate the
notion common to theories of sentence production (e.g., Bock, 1982; Levelt, 1989) that the input to the sentence-production mechanism includes a specification of the relational structure of a speaker’s intended message, so that an appropriate syntactic structure (e.g., an active, or a sentence-complement structure, or whatever) can be selected.

--- Table 6 about here ---

To generate a fully distributed message-level representation from the feature-based representation shown in Table 6, a set of 50 input units was defined, such that each input unit was sensitive to specific message structure features. In particular, each input unit was given a kind of receptive field, such that that input unit would respond (take a value of 1) only for message structures that included particular message-structure features. Which features the unit responded to was determined randomly, by selecting at most one feature from each of the message-structure feature-types described above. For example, the first input unit (the leftmost unit shown in Figure 7) was selectively sensitive to any message input that included a declarative mood feature, or that represented a message-structure that was an element of an action. As a result, this message unit was only inactive for object-relative message structures, since object relatives have the mood feature ‘modifier’ and the subordinating-element feature ‘entity’ (every other message type either is a declarative, represents an element of an action, or both). The last input unit (the rightmost input unit shown in Figure 7) was selectively sensitive to any message input that took a mood value of interrogative, that expressed an argument, that was an element of an action, or that included an agent that was a bound element. All message inputs except noun-complement inputs and object-relative inputs have at least one of these four message-structure features, and so this input unit was active for all message-structure inputs except noun-complement inputs and object-relative inputs. Generally, this input scheme results in a
distributed representation of message structure: Different kinds of message structures were represented with overlapping sets of active input units, but representations of relatively similar message structures (e.g., subject- versus object-extracted sentence-complement interrogatives) overlapped more, whereas relatively dissimilar message structures (e.g., declarative sentence-complements versus object relatives) overlapped less.

The 50 input units connected through 25 hidden units onto three output units. Each output unit represented a syntactic schema, so that one output unit represented full embedded clauses, another represented reduced embedded clauses, and the third represented noun phrases (for transitives). The model was trained by presenting it with a particular message structure, so that the particular input units that (in distributed fashion) represented a given message structure were set to a value of 1. The activations of these input units were spread to the hidden units via a set of connection weights, and the resulting activations of the hidden units were spread to the output units via another set of connection weights. The activations of the output units were then used to train the model by computing the difference between the actual syntactic output (as computed by the model) and the desired syntactic output (the structure the model was to be trained on, as shown in Figure 7), and then using that difference to modify the connection weights via the back-propagation learning algorithm (Rumelhart & McClelland, 1986). This training method implements an entrenchment-based-learning strategy, since message-to-syntax mappings that are highly learned will generate syntactic outputs that are close to their target outputs and thus will lead to small weight changes, whereas message-to-syntax mappings that are less learned will generate syntactic outputs that differ more from their target outputs, and thus will lead to larger weight changes. (Also, the model was trained with a learning rate of 0.1, and with no momentum and no bias units. Connection weights were initialized to small random
values. The simulation was implemented in the *tlearn* package; see Plunkett & Elman, 1997).

The model was trained by presenting it with a training vocabulary in random order for 2000 passes (epochs). The training vocabulary consisted of 56 input-output mappings, chosen to represent each relevant sentence type (shown in Figure 7) with a frequency that corresponded to a rough estimate of its frequency in the language. The vocabulary included 12 sentence-complement structures (9 full, 3 reduced), four noun complements (all full), four object relatives (3 full, 1 reduced), four object-extracted sentence-complement interrogatives (3 full, 1 reduced), four subject-extracted sentence-complement interrogatives (3 reduced, 1 full; the latter corresponds to a passive form like “Who did the teacher noticed that the school nurse was visited by?”), which would follow from a message-structure that defines the bound-element entity in terms of agent versus patient features), and 28 transitives (all noun-phrases).

Overall model performance on the patterns shown in Figure 7 is illustrated in Figure 8, which shows the model’s error in activating the target output syntactic schemas for each kind of message-structure input. The performance of the model reflects the learning-based claims above: Common message-to-syntax mappings (e.g., sentence-complement structures expressed with full embedded clauses) are activated relatively accurately so that they lead to relatively little error, whereas less common message-to-syntax mappings (e.g., sentence-complement structures expressed with reduced embedded clauses) are activated less accurately so that they lead to relatively more error. Note that the basic accuracy profile illustrated in Figure 8 cannot improve substantially (2000 epochs of training represented asymptotic accuracy, and adding more hidden units does not improve this performance), since the model is trained to map a given input message-structure onto more than one syntactic output (because learning one mapping necessarily causes the model to unlearn the other mapping, implying that the model can never
learn all of its mappings perfectly).

--- Figure 8 about here ---

To illustrate the primary effects reported here and in Ferreira (in press), syntactic persistence was simulated as in Chang et al. (2000). The priming event was simulated by training the model on a single input-output pattern, and then target production was simulated by presenting the model with another input pattern, and examining how much the model activated each syntactic schema node. For example, to simulate the priming effect from a noun-complement structure, which is expressed with a full embedded clause, the model was trained once to activate the full embedded clause schema from a message-structure input that corresponded to the noun-complement structure. This prime event causes the strengthening of the weights from the primed message structure to the trained syntactic schema. Then, the model was presented with the input pattern that corresponded to a declarative sentence-complement structure, and the degree to which the model activated the full embedded clause syntactic schema was measured (by dividing the activation of the full embedded clause syntactic schema by the sum of the activations of all three syntactic schemas). This target event thus measures the model’s production preferences after the prime sentence has exerted its strengthening effect. Priming is then evident to the extent that the model is differentially likely on the target trial to activate the full embedded clause schema, compared to after a different priming (training) event (e.g., compared to after the model was trained to produce a transitive structure).

Because the performance of the model is relatively variable (because of the one-to-many nature of the input-output mappings described above), the model was trained on 5 separate runs of 2000 epochs, where each run differed in terms of the initial random values assigned to the connection weights. Syntactic persistence was measured after each run. Averaged across these
five runs, the average activation ratio of the full embedded-clause schema on target trials (when
the input was a declarative sentence complement) after each kind of prime is illustrated in Figure
9. The model captures the basic pattern of results: After the model was primed by sentence-
complement structures with full versus reduced embedded clause constituents, it was relatively
more versus less likely to produce a full embedded clause constituent respectively. This is
evident both when the priming effect from both kinds of sentence-complement structures are
compared against each other, and also when the priming effect from each kind of sentence-
complement structure is compared to baseline, as measured by priming from the simple transitive
structure (indicated in the figure with a dashed line). Furthermore, note that the model captures
the asymmetry in priming, in that reduced sentence-complement primes caused target sentence
complements to be produced with full embedded clauses less often than baseline, compared to
how much full sentence-complement primes caused targets to be produced with full embedded
clauses more often than baseline. This same basic pattern is also observed after the other
structures that allow both full and reduced embedded clause constituents, namely, object-relative
structures and object-extracted sentence-complement interrogatives. On the other hand, noun-
complement structures, which allow only full embedded clause constituents, yielded almost no
syntactic persistence effect, as targets were produced as full sentence complements about as
often after noun-complement primes as after baseline. Finally, subject-extracted sentence-
complement interrogatives – *that*-trace structures – revealed an intermediate effect. After the
model was primed with a *that*-trace structure, it was somewhat less likely to produce a full
sentence-complement target, but full embedded clause production was not reduced as much as
when the model was primed with a declarative sentence-complement structure with a reduced
embedded clause.
The performance of the model follows from the factors described above. Structures that can be produced with both full and reduced embedded clauses represent a syntactic alternation; thus their mappings in the model are relatively less entrenched, and so syntactic persistence is observed. The structure that can only be produced with full embedded clauses (the noun-complement structure) is strongly entrenched, because the pattern-specific preference to produce the structure with a full embedded clause conspires with the language-general preference to produce full embedded clauses, and so little or no syntactic persistence is observed. Finally, the structure that can be produced with a reduced embedded clause only (the that-trace structure) is only somewhat entrenched, because even though its pattern-specific preference is to produce reduced embedded clauses (which should lead to strong entrenchment), this must work against the language-general preference to produce full embedded clauses, weakening the degree of entrenchment, and thus leading to an intermediate syntactic persistence effect. In sum, when the model is designed in a way that implements the learning-based assumptions described above, the appropriate pattern of results is observed.

Two aspects of the model performance are important for challenges that they pose for learning-based accounts of syntactic production generally. First, the scale of the ordinate in Figure 9 shows that though the model characterizes the relative differences in priming reasonably well, the overall magnitude of priming is notably small. For example, full versus reduced sentence complement primes led to differences in Ferreira (in press) and in the experiments above of about 20%. However, the relative-activation difference in the model is 3.8% (0.786 – 0.748). This discrepancy reflects a fundamental challenge that is faced by learning-based approaches like those characterized here. Specifically, learning-based approaches
must assume that the same learning that leads to life-long knowledge causes the relatively shorter term syntactic persistence effect. A challenge associated with that assumption is that production mechanisms must learn gradually enough so that long-term knowledge is not overly disrupted by recent processing, but rapidly enough so that recent processing can cause what are sometimes large persistence effects. This challenge might be addressed by assuming that learning is subject to a form of consolidation (so that older knowledge is less subject to modification), or relatedly, that weight changes gradually decay (McClelland & Rumelhart, 1985) or that learning occurs through both fast-decaying short-term and permanent long-term weight changes (e.g., Cleeremans & McClelland, 1991; Hinton & Plaut, 1987).

The second aspect of the model’s performance concerns the overall accuracies shown in Figure 8. In particular, consider that model performance is relatively inaccurate for subject-extracted sentence-complement interrogatives (i.e., that-trace) structures. This reflects the fact that when the model is presented with a message structure for a subject-extracted sentence-complement interrogative, the activation of the full embedded clause schema is relatively high (reflecting the bias across the language to activate that full embedded-clause schema). This implies that the model is somewhat likely to select that full embedded clause schema, which would result in an ungrammatical structure (i.e., a that-trace violation). Of course, speakers rarely produce that-trace violations (as shown by prime-structure performance in Experiments 2 and 3). It may be relevant that the syntactic vocabulary of the model in Figure 6 is restricted, so that the model’s rate of selecting full embedded clauses might be inflated (i.e., speakers may paraphrase when producing that-trace structures, avoiding embedded clauses entirely; the model, in contrast, cannot paraphrase). Another way to accommodate this issue is to assume that the outcomes of selection may need to be monitored and, when leading to ungrammatical
productions, reprocessed. This latter idea portrays a sentence-production apparatus that retrieves syntactic representations much like memory processes generally retrieve knowledge – specifically, with some trial and error.

Conclusions

The experiments presented here show that when speakers produce sentences, their tendency to repeat syntactic structures comes from an impetus to describe a given message feature with a particular syntactic structure from one sentence to the next in the same way. This impetus reflects a kind of syntactic affirmative action, in that it is greater with less preferred message-to-syntax mappings than with more preferred message-to-syntax mappings. In addition to supporting a learning-based account of syntactic persistence generally, these results illuminate how less preferred options remain viable members of a cognitive mechanism’s vocabulary. We repeat what we learn, but what we learn is determined not only by what we are commonly exposed to, but also by what we are less commonly exposed to.
References


Ferreira, V. S. (in press). The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language.*


Appendix

Stimuli from all experiments. First column indicates stimulus number. Second column indicates sentence category (SC = sentence complement; OR = object relative; OE = object-extracted sentence-complement interrogative; DT = declarative target in Experiments 2 and 3). Third column indicates sentence subtype (f/r = full/reduced; t/t = that-trace; w/s = wh-in-situ; Sim = simplex). Fourth column reports stimulus.

1  SC  f/r   The massage therapist felt (that) the client was too tense.
1  OR  f/r   The massage therapist felt the muscle (that) the patient complained about.
1  OE  f/r   What did the massage therapist feel (that) the client had strained?
1  SE  t/t   Who did the massage therapist feel was too tense?
1  SE  w/s   The massage therapist felt that who was too tense?
1  SE  Sim  Who did the massage therapist feel the shoulder muscle of?
1  DT  f/r   The massage therapist felt (that) the client had strained a muscle.
2  SC  f/r   The psychic predicted (that) the desperate woman would win the state lottery.
2  OR  f/r   The psychic predicted the set of winning numbers (that) the woman needed.
2  OE  f/r   What did the psychic predict (that) the desperate woman would win?
2  SE  t/t   Who did the psychic predict would win the lottery?
2  SE  w/s   The psychic predicted that who would win the lottery?
2  SE  Sim  Who did the psychic predict the winning numbers for?
2  DT  f/r   The psychic predicted (that) the desperate woman would win the lottery.
3  SC  f/r   The prosecutor concluded (that) the detectives should gather more evidence.
3  OR  f/r   The prosecutor concluded the case (that) the public was watching closely.
3  OE  f/r   What did the prosecutor conclude (that) the detectives should gather?
3  SE  t/t   Who did the prosecutor conclude should gather the evidence?
3  SE  w/s   The prosecutor concluded that who should gather the evidence?
3  SE  Sim  Who did the prosecutor conclude the plea bargain with?
3  DT  f/r   The prosecutor concluded (that) the detectives should gather the evidence.
4  SC  f/r   The reporter heard (that) the rich tycoon made a large donation.
4  OR  f/r   The reporter heard the story (that) the White House source leaked.
4  OE  f/r   What did the reporter hear (that) the rich tycoon donated?
4  SE  t/t   Who did the reporter hear made a large donation?
4  SE  w/s   The reporter heard that who made a large donation?
4  SE  Sim  Who did the reporter hear the juicy new story from?
4  DT  f/r   The reporter heard (that) the rich tycoon donated a million dollars.
5  SC  f/r   The salesman guaranteed (that) his customers would be completely satisfied.
5  OR  f/r   The salesman guaranteed the workmanship (that) the expert had provided.
5  OE  f/r   What did the salesman guarantee (that) his customers would get?
5  SE  t/t   Who did the salesman guarantee would get the best price?
5  SE  w/s   The salesman guaranteed that who would get the best price?
5  SE  Sim  Who did the salesman guarantee the perfectly running vehicle to?
5  DT  f/r   The salesman guaranteed (that) his customers would get the best price.
6  SC  f/r   The professor realized (that) only some students studied the material.
6  OR  f/r   The professor realized the dream (that) the whole university knew about.
6  OE  f/r   What did the professor realize (that) his students hated?
Who did the professor realize hated his class?
The professor realized that who hated his class.
Who did the professor realize the life-long dream with?
The professor realized (that) his students hated his class.
The police officer warned (that) the aggressive driver would get a ticket.
The police officer warned the woman (that) the sniper was aiming at.
What did the police officer warn (that) the aggressive driver would get?
The police officer warned that who would get a ticket.
Who did the police officer warn about the dangerous fugitive?
The police officer warned (that) the company was about to report.
What did the accountant estimate (that) the wealthy widow would pay?
The accountant estimated that who would pay a lot of taxes.
Who did the accountant estimate the lucrative profits for?
The accountant estimated (that) the company was about to report.
What did the border patrol indicate (that) the agent should search?
The border patrol indicated that who should search the car.
Who did the border patrol indicate the suspicious activity to?
The border patrol indicated (that) the agent should search the car.
What did the attorney argue (that) the defendant had stolen?
The attorney argued (that) the defendant had done nothing wrong.
What did the attorney argue with about the insanity defense?
The attorney argued (that) the book recommended.
Who did the general declare had destroyed the bridge?
The general declared (that) the troops had destroyed the bridge.
The chiropractor observed (that) the suffering patient could not stand up straight.
The chiropractor observed the painful damage (that) the patient was complaining about.
What did the chiropractor observe (that) the patient had damaged?
Who did the chiropractor observe had damaged their back?
The attorney observed that who had damaged their back?
Who did the chiropractor observe the spinal damage of?
Who did the chiropractor observe had damaged their back.
The security guard demanded (that) the known shoplifter leave the store immediately.
The security guard demanded the weapon (that) the criminal was holding.
What did the security guard demand (that) the shoplifter return?
The security guard demanded that who return the clothing?
Who did the security guard demand the stolen clothing from?
The security guard demanded (that) the shoplifter return the clothing.
The secretary recalled (that) the annoying client had been there before.
The secretary recalled the annoying client (that) the security guard had warned her about.
What did the secretary recall (that) the boss had canceled?
Who did the secretary recall had canceled the appointment.
The secretary recalled that who had canceled the appointment.
Who did the secretary recall from the telephone conversation?
The secretary recalled (that) the boss had canceled the appointment.
The gambler guessed (that) the lucky newcomer had bluffed on the previous hand.
The gambler guessed the winning horse (that) the trainer had tipped him off about.
What did the gambler guess (that) his opponent was holding?
Who did the gambler guess was holding an ace?
The gambler guessed that who was holding an ace?
Who did the gambler guess the winning hand of?
The paramedic explained (that) the old woman had fainted.
The paramedic explained the possibility (that) the family was hoping for.
What did the paramedic explain (that) the computer was monitoring?
What did the paramedic explain was monitoring the patient's heartbeat?
The paramedic explained that what was monitoring the patient's heartbeat?
What did the paramedic explain about the regular heartbeat?
The paramedic explained (that) the computer was monitoring the patient's heartbeat.
The teacher noticed (that) the failing student skipped class.
The teacher noticed that who was visiting the school nurse was sick.
Who did the teacher notice (that) the student was visiting?
Who did the teacher notice was visiting the school nurse?
The teacher noticed who was visiting the school nurse?
Who did the teacher notice visiting the school nurse?
The broker advised (that) low risk investors sell their stock.
The broker advised the investor (that) the IRS had previously audited.
What did the broker advise (that) low risk investors sell?
Who did the broker advise sell their stock?
The broker advised that who sell their stock?
Who did the broker advise to sell their stock?
The receptionist denied (that) the pushy patient had an appointment.
The receptionist denied the accusation (that) the investigator levelled against her.
What did the receptionist deny (that) the pushy patient was late for?
Who did the receptionist argue was late for their appointment?
The receptionist denied that who was late for their appointment?
The receptionist denied the late-morning appointment to?
The receptionist denied (that) the pushy patient was late for their appointment.
The wedding planner proposed (that) the bride's family visit the chapel.
The wedding planner proposed the pattern (that) the bride's parents wanted.
What did the wedding planner propose (that) the bride's family visit?
The wedding planner proposed who visit the chapel?
The wedding planner proposed the crazy idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the bride's family visit the chapel.
The wedding planner proposed the idea the scout had raved about.
What did the wedding planner propose (that) the bride's family visit?
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the crazy idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The wedding planner proposed (that) the pushy patient was late for their appointment.
The wedding planner proposed that who visit the chapel?
The wedding planner proposed the idea to?
The dentist asserted the privilege (that) the lawyer had advised him to take.

What did the dentist assert (that) the child would need?

Who did the dentist assert would need a filling?

The dentist asserted that who would need a filling?

Who did the dentist assert the necessity of the procedure to?

The dentist asserted (that) the child would need a filling.

The surgeon promised (that) the injured goalie would heal soon.

The surgeon promised the procedure (that) the patient needed.

What did the surgeon promise (that) the surgery would cure?

What did the surgeon promise would cure the disease?

The surgeon promised that what would cure the disease?

The surgeon promised (that) the surgery would cure the disease.

The x-ray proved (that) the hurt child had broken a bone.

The x-ray proved the theory (that) the doctor had guessed at.

What did the x-ray prove (that) the child had broken?

What did the x-ray prove had broken a bone?

The coach proved that who had broken a bone?

The contest official pronounced (that) the original winner had been disqualified.

The contest official pronounced the winner (that) the judges had voted for.

What did the contest official pronounce (that) the winner had won?

Who did the contest official pronounce had won a million dollars?

The contest official pronounced that who had won a million dollars?

Who did the contest official pronounce to be the million-dollar winner?

The x-ray proved (that) the child had broken a bone.

The x-ray proved (that) the tenant had a cat.

The dog sensed (that) his owner might be in danger.

The dog sensed the danger (that) the owner was unaware of.

What did the dog sense (that) his owner was holding?

Who did the dog sense was holding some food?

The dog sensed that who was holding some food?

Who did the dog sense when he started growling?

The mechanic mentioned (that) the antique car could use a tune up.

The mechanic mentioned the price estimate (that) the customer had requested.

What did the mechanic mention (that) the antique car could use?

Which car did the mechanic mention could use a tune up?
Which car did the mechanic mention the price estimate for?
The mechanic mentioned (that) the antique car could use a tune up.
The librarian noted (that) the frequent visitor had an overdue book.
The librarian noted the book (that) the child tried to hide.
What did the librarian note (that) the visitor had lost?
Who did the librarian note had lost a book?
The librarian noted that who had lost a book?
Who did the librarian note stealing the lost book?
The librarian noted (that) the visitor had lost a book.
The jury believed (that) the young witness told the truth.
The jury believed the witness (that) the lawyer had accused of lying.
What did the jury believe (that) the young witness had lied about?
Who did the jury believe had lied about the alibi?
The jury believed that who had lied about the alibi?
Who did the jury believe when he cried on the stand?
The ticket agent feared (that) the flight would be cancelled.
The ticket agent feared the new state law (that) the airline companies had pushed for.
What did the ticket agent fear (that) the customer would complain about?
The ticket agent feared that who would complain about the flight.
Who did the ticket agent fear when he complained about the flight?
The plumber confessed (that) some customers had been overcharged.
The plumber confessed the fraudulent behavior (that) the police had discovered.
Who did the plumber confess had overcharged some customers?
The plumber confessed that who had overcharged some customers?
Who did the plumber confess to about the overcharge?
The applicant surmised (that) the company had already hired someone.
The applicant surmised the decision (that) the employer was trying to hide.
Who did the applicant surmise that the company had hired?
The agent surmised had hired his client?
Who did the agent surmise the theory about after the investigation?
The agent surmised (that) the company had hired his client.
The director announced (that) Hollywood's hottest actor would be playing the part.
The director announced the casting decision (that) the critics hoped for.
Who did the director announce (that) the actor would play?
Who did the director announce would play the part?
The director announced that who would play the part?
Who did the director announce the casting decision to?
The director announced (that) the actor would play the part.
The newspaper reported (that) three elderly couples had been robbed.
The newspaper reported the news item (that) everybody wanted to know about.
Who did the newspaper report (that) the criminal had mugged?
Who did the newspaper report had mugged three people?
The newspaper reported that who had mugged three people?
Who did the newspaper report about in the special issue?
The newspaper reported (that) the criminal had mugged three people.
The handwriting analysis revealed (that) the mafia boss had signed the contract.
The handwriting analysis revealed the forged signature (that) the detectives had identified.
What did the handwriting analysis reveal (that) the mafia boss had signed?
Who did the handwriting analysis reveal had signed the contract?
The handwriting analysis reveal as having signed the contract?
Who did the handwriting analysis reveal about in the special issue?
The handwriting analysis revealed (that) the mafia boss had signed the contract.
The hobbyist maintained (that) the antique cars were one of a kind.
What did the hobbyist maintain (that) the stamp was worth?
What did the hobbyist maintain was worth a thousand dollars?
The hobbyist maintained that what was worth a thousand dollars?
The hobbyist maintained as if it was worth a thousand dollars?
The travel agent confirmed (that) the honeymooning couple had a hotel reservation.
The travel agent confirmed the flight (that) the traveler wanted.
What did the travel agent confirm (that) the couple had reserved?
Who did the travel agent confirm had reserved the hotel room?
The travel agent confirmed that who had reserved the hotel room?
The travel agent confirm the hotel reservation for?
The travel agent confirmed (that) the couple had reserved a hotel room.
The pirate suspected (that) the stow-away had hidden the treasure.
The pirate suspected the stowaway (that) the crew had discovered.
What did the pirate suspect (that) the stow-away had hidden?
Who did the pirate suspect had hidden the treasure?
The pirate suspected that who had hidden the treasure?
The pirate suspect when he saw the empty treasure chest?
The pirate suspected (that) the stow-away had hidden the treasure.
Author Note

Portions of this work were presented at the Fourteenth Annual CUNY Conference on Human Sentence Processing in Philadelphia, PA (March, 2001). This research was supported by National Institute of Health grant R01 MH-64733. I thank Kay Bock, Fernanda Ferreira, Susan Garnsey, Neal Pearlmutter, the members of the psycholinguistics group at UCSD, and most especially Gary Dell for comments on this manuscript, and Carla Firato, Janet Lee, Melanie Hudson, and Hiromi Yoshita for data-collection assistance. Please address correspondence to Vic Ferreira at the Department of Psychology 0109, University of California, San Diego, La Jolla, CA, 92093-0109. Email: vferreira@ucsd.edu.
Table 1.

*Overall production performance on prime and target sentences in Experiment 1.*

<table>
<thead>
<tr>
<th>Prime type</th>
<th>Sentence complement</th>
<th></th>
<th></th>
<th>Object relative</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Reduced</td>
<td></td>
<td>Full</td>
<td>Reduced</td>
<td>Other</td>
<td>Forgot</td>
<td></td>
</tr>
<tr>
<td>Sentence complement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>334</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>151</td>
<td>219</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>0</td>
<td>0</td>
<td>286</td>
<td>41</td>
<td>51</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>222</td>
<td>44</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Production of targets presented with full embedded clauses

<table>
<thead>
<tr>
<th>Sentence complement</th>
<th>Full</th>
<th>Reduced</th>
<th></th>
<th>Object relative</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>148</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>122</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>124</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>111</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Production of targets presented with reduced embedded clauses

<table>
<thead>
<tr>
<th>Sentence complement</th>
<th>Full</th>
<th>Reduced</th>
<th></th>
<th>Object relative</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>102</td>
<td>62</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>85</td>
<td>76</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>86</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>78</td>
<td>78</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.

*Overall production performance on prime sentences in Experiment 2.*

<table>
<thead>
<tr>
<th>Prime type</th>
<th>Object extracted</th>
<th>Subject extracted</th>
<th>Other</th>
<th>Forgot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Reduced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object extracted</td>
<td>318</td>
<td>56</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Reduced</td>
<td>100</td>
<td>275</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Subject extracted</td>
<td>0</td>
<td>0</td>
<td>358</td>
<td>19</td>
</tr>
<tr>
<td>Full</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>346</td>
</tr>
<tr>
<td>Reduced</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3.
*Overall production performance on target sentences in Experiment 2.*

<table>
<thead>
<tr>
<th>Prime Type</th>
<th>Full sentence-complement</th>
<th>Reduced sentence-complement</th>
<th>Other</th>
<th>Forgot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target presented with full embedded clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object extracted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>145</td>
<td>18</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>Reduced</td>
<td>130</td>
<td>30</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Subject extracted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>134</td>
<td>25</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>Reduced</td>
<td>138</td>
<td>24</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Sentence complement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>105</td>
<td>57</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>Reduced</td>
<td>69</td>
<td>88</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>98</td>
<td>45</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>Reduced</td>
<td>79</td>
<td>76</td>
<td>25</td>
<td>12</td>
</tr>
</tbody>
</table>
Table 4.  
*Overall production performance on prime sentences in Experiment 3.*

<table>
<thead>
<tr>
<th>Prime type</th>
<th>Unreduced</th>
<th>Reduced</th>
<th>Unreduced</th>
<th>Reduced</th>
<th>Other</th>
<th>Forgot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object extracted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>319</td>
<td>56</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Reduced</td>
<td>83</td>
<td>285</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Subject extracted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>0</td>
<td>0</td>
<td>354</td>
<td>8</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Reduced</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>356</td>
<td>15</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 5. *Overall production performance on target sentences in Experiment 3.*

<table>
<thead>
<tr>
<th>Prime Type</th>
<th>Full sentence-complement</th>
<th>Reduced sentence-complement</th>
<th>Other</th>
<th>Forgot</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Object extracted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>142</td>
<td>21</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Reduced</td>
<td>131</td>
<td>25</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td><strong>Subject extracted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>137</td>
<td>28</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Reduced</td>
<td>141</td>
<td>30</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td><strong>Sentence complement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>107</td>
<td>54</td>
<td>21</td>
<td>10</td>
</tr>
<tr>
<td>Reduced</td>
<td>64</td>
<td>85</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td><strong>Object relative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreduced</td>
<td>79</td>
<td>70</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td>Reduced</td>
<td>67</td>
<td>94</td>
<td>21</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 6. 
*Feature decomposition of model’s message structures.*

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Example</th>
<th>Mood</th>
<th>Element Type</th>
<th>Subordinating element</th>
<th>Bound element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative sentence-complement</td>
<td>The jury believed (that) the witness told the truth.</td>
<td>Declarative</td>
<td>Proposition</td>
<td>Verb</td>
<td>Neither</td>
</tr>
<tr>
<td>Noun complement</td>
<td>The theory that penguins built the igloos was false.</td>
<td>Declarative</td>
<td>Proposition</td>
<td>Noun</td>
<td>Neither</td>
</tr>
<tr>
<td>Object relative</td>
<td>The jury believed the witness (that) the lawyer accused of lying.</td>
<td>Modifier</td>
<td>Proposition</td>
<td>Noun</td>
<td>Patient</td>
</tr>
<tr>
<td>Object-extracted sentence-complement</td>
<td>Who did the teacher notice (that) the student was visiting?</td>
<td>Interrogative</td>
<td>Proposition</td>
<td>Verb</td>
<td>Patient</td>
</tr>
<tr>
<td>Interrogative</td>
<td>Who did the teacher notice was visiting the student?</td>
<td>Interrogative</td>
<td>Proposition</td>
<td>Verb</td>
<td>Agent</td>
</tr>
<tr>
<td>Transitive</td>
<td>The teacher scowled at the failing student at the back of the class.</td>
<td>Declarative</td>
<td>Argument</td>
<td>Verb</td>
<td>Neither</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1. Percentage of full embedded clauses produced in sentence-complement targets after priming by sentence complement structures versus noun-complement structures with or without full embedded clauses. Data from Ferreira (in press), Experiment 2. Error bars indicate 95% confidence-interval halfwidths by speakers.

Figure 2. Experiment 1 results: Percentage of full embedded clauses produced in sentence-complement targets after priming by sentence complement structures versus object-relative structures with full versus reduced embedded clauses. Error bars indicate 95% confidence-interval halfwidths by speakers.

Figure 3. Experiment 2 results: Percentage of full embedded clauses produced in sentence-complement targets after priming by object-extracted versus subject-extracted sentence-complement interrogatives with full versus reduced embedded clauses. Error bars indicate 95% confidence-interval halfwidths by speakers.

Figure 4. Experiment 3 results: Percentage of full embedded clauses produced in sentence-complement targets after priming by object-extracted versus subject-extracted sentence-complement interrogatives with full versus reduced embedded clauses. Error bars indicate 95% confidence-interval halfwidths by speakers.

Figure 5. Results of Experiments 2 and 3 combined: Percentage of full embedded clauses produced in sentence-complement targets after priming by object-extracted versus subject-extracted sentence-complement interrogatives with or without reduced embedded clauses. Error bars indicate 95% confidence-interval halfwidths by speakers.

Figure 6. Learning-based model of syntactic persistence.

Figure 7. Input-output patterns that the model was trained on. Filled in squares represent a
value of 1, open squares a value of 0. Each pattern was trained multiple times per epoch, as described in the text.

Figure 8. Overall performance of the model: Accuracy (in terms of the squared difference between each output unit’s actual activity and its target activity) for the patterns illustrated in Figure 7 after 2000 epochs of training. Results averaged across 5 runs.

Figure 9. Syntactic persistence in the model. Ratio of activation of full embedded clause syntactic schema to the summed activation of all syntactic schemas as a function of being trained for a single sweep on different structures. Results from after 2000 epochs of training and averaged across 5 runs. Dashed line indicates persistence results from transitive (neutral) primes.
The company insured that the farm was covered for two million dollars.

The company insured the farm was covered for two million dollars.

The theory that penguins built the igloos was completely false.

The theory of the melting igloos was completely false.
The teacher noticed that the failing student skipped class.

The teacher noticed the failing student skipped class.

The teacher noticed the student that the school nurse said was sick.

The teacher noticed the student the school nurse said was sick.
Object-Extracted Sentence-Complement Interrogatives

Who did the teacher notice that the student was visiting? 78.7%

Who did the teacher notice the student was visiting? 60.5%

Subject-Extracted Sentence-Complement Interrogatives

The teacher noticed that who was visiting the school nurse? 77.7%

Who did the teacher notice was visiting the school nurse? 66.9%
Who did the teacher notice visiting the school nurse?

Object-Extracted Sentence-Complement Interrogatives

Who did the teacher notice the student was visiting?

Subject-Extracted Sentence-Complement Interrogatives

Who did the teacher notice visiting the school nurse?

Who did the teacher notice was visiting the school nurse?

% Full Embedded Clauses Produced in Sentence Complement Targets
Object-Extracted Sentence-Complement Interrogatives

Who did the teacher notice that the student was visiting? [78.9% Not Reduced, 60.7% Reduced]

Subject-Extracted Sentence-Complement Interrogatives

The teacher noticed that who was visiting the school nurse? / Who did the teacher notice visiting the school nurse? [72.9% Not Reduced, 64.7% Reduced]
Distributed representation of message-structure features

Input (50)

Hidden (25)

Output (3)

Full embedded clause schema

Reduced embedded clause schema

Noun phrase schema
<table>
<thead>
<tr>
<th>Distributed message-structure</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence complement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence complement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noun complement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object relative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object-extracted sentence-complement interrogative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object-extracted sentence-complement interrogative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject-extracted sentence-complement interrogative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mean Squared Error

Sentence complements

Noun complements

Object relatives

Object-extracted sentence-complement interrogatives

Subject-extracted sentence complement interrogatives

Transitives

More accurate | Less accurate

Full

Reduced

Full

Reduced

Full

Reduced

Full

Reduced

More accurate | Less accurate

0.00 0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00

0.03

0.08

1.30

0.09

1.28

0.22

0.93

0.23

0.01
Relative Activation of Full Embedded Clause Representation

Sentence complements

Full: 0.786
Reduced: 0.748

Noun complements

Full: 0.781
Reduced: 0.757

Object relatives

Full: 0.784
Reduced: 0.750

Object-extracted sentence-complement interrogatives

Full: 0.792
Reduced: 0.766

Subject-extracted sentence-complement interrogatives

Noun phrase: 0.779

Transitives