



## Evidence for Implicit Learning in Syntactic Comprehension

Alex B. Fine,<sup>a,b</sup> T. Florian Jaeger<sup>a,c</sup>

<sup>a</sup>*Department of Brain and Cognitive Sciences, University of Rochester*

<sup>b</sup>*Department of Linguistics, University of Rochester*

<sup>c</sup>*Department of Computer Science, University of Rochester*

Received 10 December 2010; received in revised form 5 July 2012; accepted 10 July 2012

---

### Abstract

This study provides evidence for implicit learning in syntactic comprehension. By reanalyzing data from a syntactic priming experiment (Thothathiri & Snedeker, 2008), we find that the error signal associated with a syntactic prime influences comprehenders' subsequent syntactic expectations. This follows directly from error-based implicit learning accounts of syntactic priming, but it is unexpected under accounts that consider syntactic priming a consequence of temporary increases in base-level activation. More generally, the results raise questions about the principles underlying the maintenance of implicit statistical knowledge relevant to language processing, and about possible functional motivations for syntactic priming.

*Keywords:* Language processing; Syntactic priming; Learning; Adaptation

---

To comprehend language, we must infer intended messages based on incremental, noisy input. Psycholinguistic research suggests that humans accomplish this by exploiting statistical information in the linguistic signal. Recent work suggests that knowledge of this information is *malleable* in the face of new evidence, in that recent exposure can change linguistic representations in adults, ranging from phonetic representations (e.g., Norris, McQueen, & Cutler, 2003) to syntactic structures (Farmer, Monaghan, Misyak, & Christiansen, 2011b; Fine, Jaeger, Farmer, & Qian, 2012; Kaschak & Glenberg, 2004). This raises the question of what *mechanism* underlies this flexibility. We test the hypothesis that *implicit learning*, which plays an important role in human skill acquisition (e.g., Plunkett & Juola, 1999; Toscano & McMurray, 2010 on language acquisition; Botvinick & Plaut, 2004; on sequential motor skill acquisition), also operates during language processing in

adults (Chang, Dell, & Bock, 2006; Chang, Dell, Bock, & Griffin, 2000; Jaeger & Snider, in press; Kaschak, Kutta, & Coyle, 2012; Reitter, Keller, & Moore, 2011).

In this article, we articulate the hypothesis that implicit learning is operative in adult language processing in terms of *error-based* implicit learning. The central idea of error-based learning, which has successfully accounted for behavioral data across a wide array of cognitive and perceptual domains, is that behavior at a given time point,  $t$ , is influenced by an error signal relevant to that behavior at time  $t-1$ . For instance, the amplitude of a saccade to a visual target varies as a function of error signals from recent saccades, where error is defined as how far the fixation on a recent trial was from the intended location (Wallman & Fuchs, 1998). Similarly, during reaching movements, humans rapidly modulate motor plans according to “errors” introduced by experimentally controlled perturbations of prior reaching movements (e.g., Shadmehr & Mussa-Ivaldi, 1994; Shadmehr, Smith, & Krakauer, 2010). In the perceptual work cited above, these phenomena are often referred to as instances of *adaptation*, which can be generally thought of as any change in an organism’s behavior in response to a change in the statistics of the environment in which that behavior takes place. For the purposes of our discussion, we make a distinction between adaptation and learning, defining adaptation as the *outcome* of a learning *mechanism*.

Previous studies have found evidence that language processing is sensitive to prediction error (or “expectation violation”). Expectation violations are associated with signature patterns in event-related potentials (Kutas & Hillyard, 1980, 1984) and magnetoencephalography experiments (e.g., Dikker, Rabagliati, & Pylkänen, 2009). In reading experiments, comprehenders take longer to process unexpected words or structures (e.g., Jurafsky, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Kello, 1993). These effects are not limited to violations of strong expectations, as in so-called *garden paths* (Frazier, 1987): Word-by-word reading times in self-paced reading experiments and in corpora of eye movements in reading are correlated with word predictability (e.g., Demberg & Keller, 2008; Hale, 2001; Levy, 2008; McDonald & Shillcock, 2003). In short, a large body of evidence suggests that prediction error has immediate effects on how difficult it is to process words and phrases.

Error-based implicit learning accounts predict that the prediction error from *recently processed* material, which can be interpreted as a gradient error signal, affects expectations about *upcoming* material (Chang et al., 2000, 2006). This is the prediction we test here. A well-studied phenomenon in language processing whereby recently processed material affects how subsequent material is processed is syntactic priming. In production, the term syntactic priming refers to the tendency of speakers to reuse recently encountered structures; in comprehension, *syntactic priming* refers to the facilitated comprehension of a structure after it has been recently processed. Chang et al. (2006) present a connectionist model of syntactic acquisition that, without additional assumptions, is capable of capturing syntactic priming in production and comprehension. In their model, syntax is acquired via error-based learning (implemented as back-propagation), and the same mechanism predicts that syntactic priming should be stronger the more

*unexpected* the prime is. This prediction has been tested experimentally in production (Bernolet & Hartsuiker, 2010; Jaeger & Snider, in press) but not in comprehension, which is the focus of this study.

Consonant with work on adaptation in non-linguistic perceptual tasks, discussed above, we predict that expectations for a syntactic structure at time  $t$  will vary as a function of prediction error at time  $t-1$ . We quantify prediction error as the surprisal of the syntactic structure processed at  $t-1$ .

To address this question, we conduct a reanalysis of a visual world eye-tracking experiment on syntactic priming in comprehension conducted by Thothathiri and Snedeker (2008)(Experiment 3). Thothathiri and Snedeker (2008) show that comprehension of a syntactic structure is facilitated after exposure to that structure. They focus on the ditransitive alternation, where speakers of English can encode the same meaning using the double object (DO) or prepositional object (PO) structure:

1. DO: John gave his daughter a book.
2. PO: John gave a book to his daughter.

In their experiment, during each trial, subjects first heard a context story, followed by two primes related to the story. Both the story and the primes were spoken by the same speaker. On critical trials, either both primes were PO constructions or both were DO constructions. Immediately following the primes, subjects carried out a simple instruction as their eye movements to objects in a visual display (e.g., Fig. 1) were recorded. This target instruction was unrelated to the context story and the primes and was spoken by a different speaker. On critical trials, the instruction involved a ditransitive verb. An example trial with context story, two primes (both in the DO structure), and target instruction is given for the display in Fig. 1.

- (3) **Speaker A:** John's 2-year-old daughter's birthday was coming up. His secretary went to the bookstore to look for children's books. [context story]  
There, a nice bookstore clerk sold the secretary a book. [prime 1]  
That night, John read his daughter a story [prime 2]
- (4) **Speaker B:** Now you can give the horn to the dog. [target instruction]

Crucially, the post-verbal noun (e.g., "horn" in (4)) in the instruction was temporarily ambiguous between a name for an inanimate object in the display (e.g., "horn") and an animate object in the display (e.g., "horse"). For prime and target sentences, themes were inanimate and recipients were animate. Thus, anticipatory looks to the animal during the ambiguous region of the *target* sentence (e.g., looks to the horse during "hor...") indicate the subject's expectation for a DO structure (as in that structure, the recipient would follow the verb), while looks to the inanimate object (e.g., looks to the horn during "hor...") indicate expectation for a PO structure.

Here, priming consists of an increase in the subjective probability the listener assigns to hearing one structure versus another in the target trial (e.g., (4) above). If processing the two preceding prime structures affects expectations in the target sentence, an increase

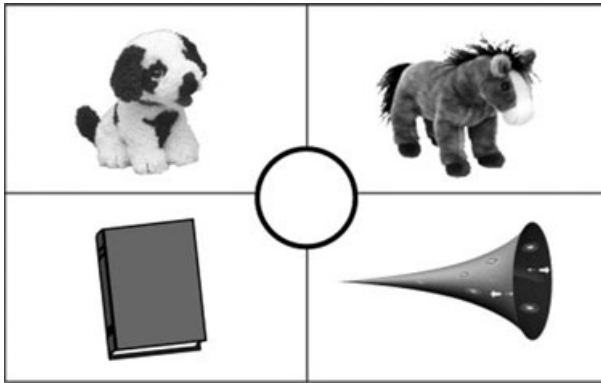


Fig. 1. An illustration of the design of Thothathiri and Snedeker (2008). The display contains two referents (horse and horn) whose labels share their onset (“hor...”).

in fixations during the ambiguous region to the animal following two DO primes and to the object following two PO primes should be observed. This is what Thothathiri and Snedeker (2008) find.

If syntactic priming in comprehension is due to error-based implicit learning (Chang et al., 2000, 2006; Jaeger & Snider, in press), we should observe the hallmark of such learning—*sensitivity to prediction error*. This prediction has not been previously tested. Specifically, we predict that the strength of the syntactic priming effect observed by Thothathiri and Snedeker will be positively correlated with the prediction error associated with comprehending the two primes. As we discuss below, this prediction is *not* made by competing accounts of syntactic priming (Malhotra, 2009; Pickering & Branigan, 1998; Pickering & Garrod, 2004; Reitter et al., 2011). A prime’s surprisal, defined as  $-\log p(\text{structure} \mid \text{preceding context})$ , provides an intuitive way to quantify the prediction error associated with processing a prime structure. Surprisal is high whenever the prime structure is unexpectedly given, for example, the preceding words in the sentence. Surprisal is also a reasonable metric for prediction error as it is known to be a good predictor of processing difficulty (Hale, 2001; Smith & Levy, 2008). Higher prime surprisal (i.e., higher prediction error) is expected to correlate with a larger priming effect (i.e., strengthened expectation for the same structure in the target sentence). Next, we describe how prime surprisal was estimated for the stimuli in Thothathiri and Snedeker’s experiment. Then we test the prediction that changes in linguistic expectations are a function of the prediction error of recently processed material.

## 1. Quantifying the prediction error

The surprisal of each prime structure was computed as  $-\log p(\text{structure} \mid \text{context story, Subject, Verb})$ . The conditional probability  $p(\text{structure} \mid \text{context story, Subject,$

*Verb*) was estimated using sentence-completion norms. One hundred completions per stimulus were collected via Amazon's online platform Mechanical Turk and annotated for DO/PO completions. Only subjects with U.S. IP addresses were allowed to provide norming data. In addition, instructions clearly indicated that subjects were required to be native speakers of English, and only subjects with at least a 95% approval rating from previous jobs were included. Each of the 171 subjects was permitted to complete each item only once, but the number of items completed by each subject ranged from 1 item to all 16 items (mean = 9, SD = 6). Mechanical Turk norming experiments (including sentence completion, forced choice, rating, and self-paced reading tasks) have been shown to reliably replicate laboratory-based experiments (Melnick, 2011; Munro et al., 2010).

Subjects in the norming study were presented with the context story (see (3) above) and were then asked to provide completions for two NP + Dative verb pairs. For the item presented in (3), for instance, subjects would read the context story and complete the strings "There, a nice bookstore clerk sold ..." and "That night, John read..." Each completion was then hand coded as PO (e.g., for Prime 1, "...a Winnie the Pooh book to the secretary."), DO ("...the woman several books"), or other ("...all the books before she could get any"). Thus, for a prime in the PO or DO structure in the Thot-hathiri and Snedeker (2008) experiment, the surprisal of that prime was computed as the negative log of the proportion of all completions for that prime in the norming study, respectively.

### 1.1. Predictions

If priming is sensitive to the error signal associated with the prime, the surprisal of the prime structure should *interact* with the main effect of prime structure. Although we test this prediction for both primes, the predictions are clearest for the first prime. If subjects' syntactic expectations are sensitive to the fact that the two primes always had the same structure, the structure of the second prime should become increasingly predictable throughout the experiment, given the structure of the first prime. In other words, if subjects' expectations are based not only lexical cues (the verb's subcategorization frame in the second prime) but also on the first prime's structure, the average surprisal of the second prime should at the very least be very low relative to the first prime. Evidence that comprehenders do, in fact, adapt their syntactic expectations to context-specific syntactic statistics comes from a series of recent self-paced reading experiments (Farmer, Fine, & Jaeger, 2011a; Fine, Qian, Jaeger, & Jacobs, 2010; Kaschak & Glenberg, 2004). Crucially, the design of the experiment is, if anything, expected to have the opposite effect on the surprisal of the first prime, as the structure of the first prime becomes increasingly unpredictable throughout the experiment. In terms of the discussion on error-based learning above, we thus primarily expect subjects' behavior at the target instruction to vary according to the error signal at the *first* prime. In addition, it is possible that a similar, but weaker, effect will be observed for the second prime.

## 1.2. Analysis

Following Thothathiri and Snedeker (2008), our dependent variable is the proportion of fixations, during the ambiguous region, to the animal (the potential recipient, e.g., the horse) *minus* the proportion of fixations to the object (the potential theme, e.g., the horn). This captures the degree to which subjects expect the recipient *relative to* the theme. When the prime structure is a DO, this difference score should be greater than when the prime structure is a PO. Following Barr (2008), proportions of fixations were first empirical logit-transformed before computing this difference score. Two trials containing primes with very large surprisal values (values that exceeded 6 bits; mean surprisal value = 2.25, SD = 1.4) were removed. The results below do not depend on this removal.

Main effects of prime structure, the surprisal of the first and second primes, target structure, and prime structure–target structure interaction, and the bias of the target verb (probability that the target verb occurs in the DO version of the dative alternation) were included in the analysis. In addition, the interaction between the surprisal of the first prime and prime structure as well as the interaction between the surprisal of the second prime and prime structure were included. The model included the maximal random-effect structure justified by the data. Collinearity was observed between prime structure and the surprisal of the second prime ( $r = -.59$ ; all other fixed-effect correlations  $r < .2$ ). Leave-one-out model comparison confirmed that collinearity did not affect any of the significant effects reported below. A standard ANCOVA over the difference scores yields the same results as those reported below.

## 2. Results

The main effect of prime structure remained only marginally significant when prime surprisal and the prime structure–prime surprisal interactions were included in the model ( $\beta = .40$ , SE = .26,  $p = .1$ ), but it was statistically significant when these terms were left out ( $\beta = .43$ , SE = .21,  $p < .05$ ), replicating Thothathiri and Snedeker (2008). The reason for the reduced significance of the main effect of priming is that the effect of prime structure is carried by the high-surprisal primes, discussed below.

As expected, no main effect of the surprisal of either the first or the second prime was observed ( $ps > .5$ ). Crucially, we found the predicted two-way interaction between the surprisal of the first prime and prime structure ( $\beta = .53$ , SE = .24,  $p < .05$ )—for DO primes, as prime surprisal increased, fixations to the animal relative to the object *increased*; for PO primes, as prime surprisal increased, fixations to the animal relative to the object *decreased*. The interaction between the surprisal of the second prime and prime structure was not significant ( $\beta = -.02$ , SE = .17,  $p = .9$ ). The significant interaction of prime structure and prime surprisal for prime 1 is shown in Fig. 2.<sup>1</sup>

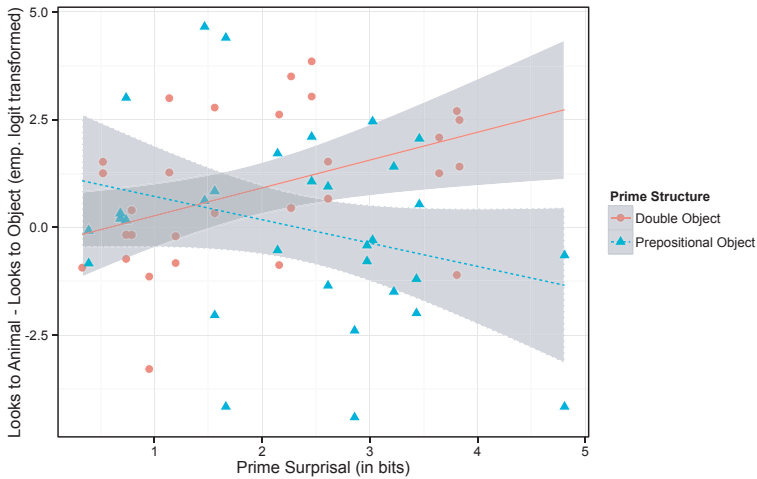


Fig. 2. The interaction between prime structure and prime surprisal. The scatterplot indicates observed empirical logit-transformed looks to the animal (i.e., looks to the potential recipient) minus empirical logit-transformed looks to the object (potential theme). Gray shaded areas indicate 95% confidence intervals around regression lines relating prime surprisal to the difference score for the two different prime types.

### 3. Discussion

Chang et al. (2006) propose a connectionist model that employs error-based learning to account for syntactic acquisition in infants. This model predicts error-based implicit learning in adults without further stipulations: The same error-based learning algorithm assumed to operate during infancy continues to operate throughout adult life. In connectionist architectures like the one proposed by Chang and colleagues, syntactic priming and its sensitivity to the prediction error associated with the prime arise as a consequence of learning (adjustments to connection weights in the network). Learning is achieved via back-propagation, whereby connection weights are adjusted as a function of error signals (the difference between the expected and the observed outcome, Chang et al., 2006: p. 234). As shown in Jaeger and Snider (in press), the error signal employed in the model by Chang and colleagues is closely related to the syntactic surprisal estimates employed here.

We found that comprehenders' expectations about upcoming input are shaped by the error signal associated with previous input, manifested in surprisal-sensitive priming. For the first of the two primes in each trial, prime strength increased with increasing prime surprisal—in line with the predictions of Chang et al. (2006). This effect replicates for comprehension what recent work has found for production: The strength of syntactic priming is sensitive to the prediction error of the prime (Benolet & Hartsuiker, 2010; Jaeger & Snider, in press; Kaschak, Kutta, & Jones, 2011).

The lack of a significant interaction between prime surprisal and prime structure for the second prime remains an issue for future work. One possible explanation is that our reanalysis failed to *detect* the effect, which—as we spelled out above—is expected to be

weaker than the effect of the first prime. Specifically, the design of the experiment does little to reduce subjects' uncertainty about the structure of the first prime, as this was counterbalanced within subjects; however, the experiment itself could greatly reduce the surprisal of the second prime, as the structure of the second prime was always the same as the first. The assumption that subjects implicitly take into account distributional information about the experiment is by no means *ad hoc*. Preliminary evidence supporting this assumption comes from recent experiments finding that subjects in syntactic processing experiments can rapidly adjust their expectations about the distribution of syntactic events in the input (Farmer et al., 2011a; Fine et al., 2010; Kaschak & Glenberg, 2004). Additional evidence comes from work on statistical learning in artificial languages (e.g., Braine et al., 1990; Wonnacott, Newport, & Tanenhaus, 2008). Still, future work, involving experiments explicitly designed to test the hypothesis of error-driven learning, rather than reanalyses such as the one reported here, will be needed to settle this issue.

Alternative accounts that attribute syntactic priming to short-term increases in the activation of syntactic representations after they have been processed (Pickering & Branigan, 1998; Pickering & Garrod, 2004) do not predict the sensitivity of prime strength to prime surprisal observed here. There are, however, extensions of these accounts that involve implicit learning. We return to these accounts below.

Previous assessments of implicit learning accounts have focused on syntactic priming in language production. Syntactic priming in production refers to the increased probability of reusing a syntactic structure after it has recently been processed (Bock, 1986). For example, in an alternation such as the ditransitive alternation, speakers are more likely to produce a PO structure if they have recently comprehended or produced a PO structure (and likewise for DO structures). Consistent with the hypothesis that syntactic priming is due to implicit learning, work on language production has provided evidence that the effect of syntactic priming is relatively long lasting (Bock & Griffin, 2000; Branigan, Pickering, Stewart, & McLean, 2000; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008) and, in particular, that it can persist beyond the most recently processed prime, leading to cumulative effects of recently processed primes (Jaeger & Snider, *in press*; Kaschak et al., 2011; Kaschak, Loney, & Borreggine, 2006; see also Fine et al., 2012; for evidence of cumulative priming in comprehension). Particularly germane to the current discussion is the so-called inverse frequency effect in syntactic priming in production: The boost in the probability of reusing a structure is higher for less frequent compared with more frequent primes (Bernolet & Hartsuiker, 2010; Jaeger & Snider, *in press*; Kaschak et al., 2011; Reitter et al., 2011; Scheepers, 2003). As the frequency of a prime is a very simple estimate of its average surprisal, the inverse frequency effect is compatible with an interpretation in terms of error-based implicit learning (cf. Chang et al., 2000, 2006; Ferreira, 2003). Error-based learning accounts attribute the inverse frequency effect to the cumulative effect of the error-signal associated with prime structures in its context. This predicts that the strength of syntactic priming is sensitive to the relative probability of a prime or its surprisal, as observed here for comprehension. Indeed, there also is mounting evidence from syntactic priming in production in support of this prediction (Bernolet & Hartsuiker, 2010; Jaeger & Snider, *in press*). The work presented



here extends the production findings to comprehension, suggesting that a single mechanism—implicit learning—might underlie priming phenomena in both language production and comprehension (cf. Tooley & Bock, 2011).

*Prima facie*, activation-based accounts of priming (Pickering & Branigan, 1998; Pickering & Garrod, 2004) predict the opposite of the “inverse frequency effect”: less frequent primes should prime *less* (lower frequency is generally assumed to be associated with lower resting activation, cf. Reitter et al., 2011: p. 24). Without additional assumptions, these accounts cannot capture the correlation between priming and the size of the error signal of recently processed structures.

So far, we have tested a specific instance of implicit learning, namely error-based learning. Since the error-based learning model by Chang and colleagues was first proposed, several alternative implicit learning accounts of syntactic priming have been developed, including both alternative supervised learning accounts without an explicit gradient error-signal (e.g., Bayesian belief update models, Fine et al., 2010) and even unsupervised learning accounts (Malhotra, 2009; Reitter et al., 2011). For example, recognizing the challenge to short-term activation accounts described above, Reitter and colleagues (Reitter et al., 2011) propose that the inverse frequency effect on syntactic priming is a result of so-called base-level learning (Anderson et al., 2004): Repeated retrieval of the same structure from memory is assumed to increase its base-level activation. Under plausible assumptions about activation boost and decay, this correctly predicts smaller proportional effects of temporary activation boosts due to recent retrieval (i.e., smaller effects of syntactic priming). The supervised and unsupervised implicit learning accounts make similar qualitative predictions for syntactic priming, but they differ in terms of the specific mechanisms assumed. It is therefore theoretically possible that unsupervised rather than supervised (error-based) implicit learning could underlie the effects observed here. To assess this possibility, we evaluated the predictions of a base-level learning account of syntactic priming against the data from Thotathiri and Snedeker (2008). We found no support for the hypothesis that base-level learning underlies the observed correlation between prime strength and prime surprisal. For details, we refer the reader to the Supplementary Information. Our efforts here can, however, only be seen as preliminary. Further work is required to convincingly adjudicate between the error-based account by Chang et al. (2006) and the account by Reitter et al. (2011).

Returning to the bigger picture, our findings, taken together with the findings from production discussed above, raise the possibility that continuous implicit learning is an essential property of the language processing system. Error-based implicit learning is endemic to certain connectionist architectures (Chang et al., 2006; Cleeremans & McClelland, 1991; Elman, 1990; Seger, 1994), and error-based models have been employed to successfully account for non-linguistic perceptual and cognitive learning (Botvinick & Plaut, 2004; Koedinger, Tenenbaum, & Shadmehr, 2007).

This study leaves open the question of what *purpose* implicit learning might serve. This is the computational counterpart to the mechanistic question addressed in this study (Marr, 1982). This question has two parts. First, why might continuous implicit learning

based on recently processed linguistic input operate throughout adulthood? Second, why would this learning be error based, in that larger deviation from what is expected leads to stronger priming?

To address the first question, consider recent work on perceptual adaptation, which suggests that comprehenders adapt to the pronunciation of specific speakers to more efficiently process their input (Bradlow & Bent, 2008; Kraljic & Samuel, 2005, 2006). For example, comprehenders adjust mappings from perceptual input to phonemes for specific speakers (Kraljic & Samuel, 2007). They are also capable of extracting characteristics of specific speaker groups (e.g., non-native Chinese speakers), improving comprehension for members of that group (Bradlow & Bent, 2008). Because these findings suggest that perceptual adaptation is speaker specific, they point to implicit learning mechanisms (possibly involving speaker-specific memory associations, cf. Horton, 2007), rather than to short-term boosts in transient activation (cf. Pickering & Garrod, 2004). The purpose of this implicit learning may be the facilitation of efficient communication. It is possible that what we, in line with previous research, have been referring to as syntactic priming serves rapid adaptation to syntactic preferences associated with different speakers or conversation topics. Such adaptation would enable comprehenders to adjust to changes in the statistics of the linguistic environment (due to shifts in topic, different speakers, contexts, etc.), thereby reducing comprehension difficulty that might otherwise result from inadequate assumptions about these statistics. Preliminary support for this hypothesis comes from recent evidence that the comprehension difficulty associated with garden-path effects decreases rapidly and cumulatively with repeated exposure to the relevant structures, suggesting that comprehenders are capable of adapting their expectations to specific linguistic environments (e.g., Fine et al., 2012; Kaschak & Glenberg, 2004).

It is possible that syntactic priming serves the adaptation to speaker-specific differences in syntactic productions (cf. Pickering & Garrod, 2004). Interestingly, although, in the experiment reported by Thothathiri and Snedeker (2008), prime sentences were produced by a *different* speaker than target sentences, but structural priming in comprehension, as well as error-based learning, is also observed. *Prima facie*, this is problematic for the hypothesis that comprehenders guide online inferences using speaker-specific statistics. However, it is possible that because both of the speakers that subjects heard in this task were part of the same situation (i.e., the experiment), listeners tacitly assumed that the two speakers' productions were being drawn from the same underlying distribution or process. Directly manipulating the influence of speaker identity in syntactic priming is a topic for future work.

Returning to the second question above, are there functional reasons why processing an unexpected syntactic structure should lead to a larger increase in expectations for the same structure? If the mechanisms underlying rapid linguistic adaptation serve to maintain efficient processing of noisy input, it is not immediately obvious why processing unexpected structures would lead to relatively large changes in expectations. An intriguing possibility to be explored in future work is that the less expected a comprehended structure, the more information it may contain about the probability of that structure being used in a specific context, in which case stronger priming by surprising primes would be a rational response

to the environment. It is also possible that error-based learning does not serve any functional reason other than guaranteeing convergence to the input statistics in the long run.

#### 4. Conclusion

We have presented evidence that syntactic priming in comprehension is due to implicit learning (Bock & Griffin, 2000). Specifically, we found that more surprising prime structures lead to stronger expectations that the same structure will be used in later sentences. This suggests that the implicit learning underlying syntactic priming is error-based (Chang et al., 2000, 2006), thereby underscoring the importance of error-based implicit learning beyond language acquisition (Chang et al., 2006). We have discussed how this effect may be part of an adaptive response facilitating efficient communication, as suggested by recent evidence from perceptual adaptation (Bradlow & Bent, 2008; Kraljic & Samuel, 2007; Kraljic, Samuel, & Brennan, 2008).

#### Acknowledgments

We are very grateful to Malathi Thothathiri and Jesse Snedeker for kindly providing us with the data reanalyzed here. We also thank Mike Kaschak and Camber Hansen-Karr for feedback on an earlier version of this manuscript. This work was partially supported by an NSF Graduate Research Fellowship to ABF and NSF BCS-0845059 as well as an Alfred P. Sloan Fellowship to TFJ.

#### Note

1. In addition to the difference score used as a DV here, two DVs available in this design are simply looks to the animal and looks to the object. We also conducted separate analyses for (a) looks to the animal and (b) looks to the object and obtained qualitatively identical results. For empirical logit-transformed looks to the animal, the two-way interaction between prime 1 surprisal and prime structure was significant ( $\beta = .25$ ,  $SE = .08$ ,  $p < .05$ ); for transformed looks to the object, the same interaction was marginally significant ( $\beta = -.23$ ,  $SE = .15$ ,  $p = .1$ ). In both of these analyses, the interaction showed the expected direction: larger priming for more surprising primes.

#### References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, *111*(4), 1036–1060.
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *Special Issue: Emerging Data Analysis*, *59*, 457–474.

- Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, 114(3), 455–461. doi: 10.1016/j.cognition.2009.11.005
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387. doi:10.1016/0010-0285(86)90004-6
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177–192. doi: 10.1037//0096-3445.129.2.177
- Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111, 395–429.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2), 707–729. doi: 10.1016/j.cognition.2007.04.005
- Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., & Fisch, S. M. (1990). Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language*, 29, 591–610.
- Branigan, H. P., Pickering, M. J., Stewart, A. J., & McLean, J. F. (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition*, 28, 1297–1302.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. doi: 10.1037/0033-295X.113.2.234
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29, 217–229.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120, 235–253.
- Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109, 193–210.
- Dikker, S., Rabagliati, H., & Pylkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, 110(3), 293–321.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Farmer, T. A., Fine, A. B., & Jaeger, T. F. (2011a). Implicit context-specific learning leads to rapid shifts in syntactic expectations. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 2055–2060.
- Farmer, T. A., Monaghan, P., Misyak, J. B., & Christiansen, M. H. (2011b). Phonological typicality influences sentence processing in predictive contexts: A reply to Staub et al. (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1318–1325.
- Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, 48, 379–398.
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2012). Rapid expectation adaptation during syntactic comprehension. Available at [http://www.bcs.rochester.edu/people/afine/Fine-et-al\\_SyntacticAdaptation\\_4-22-2012.pdf](http://www.bcs.rochester.edu/people/afine/Fine-et-al_SyntacticAdaptation_4-22-2012.pdf). Accessed on January 7, 2013.
- Fine, A. B., Qian, T., Jaeger, T. F., & Jacobs, R. A. (2010). Syntactic adaptation in language comprehension. *Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics* (pp. 18–26). Uppsala, Sweden: Association for Computational Linguistics.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 559–586). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hale, J. T. (2001). A probabilistic early parser as a psycholinguistic model. In ??? ?????????? (Ed), *Proceedings of the North American Association for Computational Linguistics* (pp. 159–166). Pittsburgh, PA: NAACL.
- Hartsuiker, R. J., Bernardolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, 58(2), 214–238. doi: 10.1016/j.jml.2007.07.003
- Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and Cognitive Processes*, 22, 1114–1139.

- Jaeger, T. F., & Snider, N. (in press). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194.
- Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General*, 133(3), 450–467. doi: 10.1037/0096-3445.133.3.450
- Kaschak, M. P., Kutta, T. J., & Coyle, J. M. (2012). Long and short term cumulative structural priming effects. *Language and Cognitive Processes*, 1–23
- Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin and Review*, 18, 1133–1139.
- Kaschak, M. P., Loney, R. A., & Borreggine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition*, 99, B73–B82.
- Koerding, K. P., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, 10(6), 779–786. doi: 10.1038/nn1901
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51(2), 141–178. doi: 10.1016/j.cogpsych.2005.05.001
- Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, 13(2), 262–268. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16892992>
- Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56(1), 1–15. doi: 10.1016/j.jml.2006.07.010
- Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, 19(4), 332–338. doi: 10.1111/j.1467-9280.2008.02090.x
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203–205.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161–163.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. doi: 10.1016/j.cognition.2007.05.006
- MacDonald, M. C., Pearlmutter, N., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- Malhotra, G. (2009). *Dynamics of structural priming*. PhD dissertation, University of Edinburgh.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of information*. San Francisco, CA: W. H. Freeman.
- McDonald, S., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14(6), 648–652. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14629701>
- Melnick, R. (2011). Judgment studies of information-driven syntactic reduction. Available at <http://www.stanford.edu/~rmelnick/files/Melnick2011.pdf>. Accessed on January 7, 2013.
- Munro, R., Bethard, S., Kuperman, V., Lai, V. T., Melnick, R., Potts, C., Schnoebelen, T., & Tily, H. (2010). Crowdsourcing and language studies: The new generation of linguistic data. Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, Los Angeles, CA.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47(2), 204–238. doi: 10.1016/S0010-0285(03)00006-9
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169–226.
- Plunkett, K., & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, 23(4), 463–490.

- Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4), 587–637. doi: 10.1111/j.1551-6709.2010.01165.x
- Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57, 348–379. doi: 10.1016/j.jml.2007.03.002
- Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, 89, 179–205.
- Seger, A. (1994). Implicit learning. *Psychological Bulletin*, 115(2), 163–196.
- Shadmehr, R., & Mussa-Ivaldi, A. (1994). Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, 14(5), 3208–3224.
- Shadmehr, R., Smith, M., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, 33, 89–108. doi: 10.1146/annurev-neuro-060909-153135
- Smith, N. J., & Levy, R. (2008). Optimal processing times in reading: A formal model and empirical investigation. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 595–600). Austin, TX: Cognitive Science Society.
- Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108(1), 51–68. doi: 10.1016/j.cognition.2007.12.012
- Tooley, K., & Bock, K. (2011). The Equivalence of Syntactic Priming in Comprehension and Production. The 24th Annual CUNY Conference on Human Sentence Processing. Palo Alto, CA.
- Toscano, J. C., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science*, 34(3), 434–464. doi: 10.1111/j.1551-6709.2009.01077.x
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19(3), 528–553.
- Wallman, J., & Fuchs, A. F. (1998). Saccadic gain modification: Visual error drives motor adaptation. *Journal of Neurophysiology*, 80(5), 2405–2416. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9819252>. Accessed May 25, 2010.
- Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, 56(3), 165–209. doi: 10.1016/j.cogpsych.2007.04.002

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Appendix.** Supplementary analyses.