Implicit Context-Specific Learning Leads to Rapid Shifts in Syntactic Expectations

Thomas A. Farmer (tfarmer@bcs.rochester.edu), Alex B. Fine (afine@bcs.rochester.edu), and T. Florian Jaeger (fjaeger@bcs.rochester.edu)
Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY 14627 USA

Abstract

During incremental language understanding, comprehenders draw on a rich base of probabilistic cues to efficiently process the noisy perceptual input they receive. One challenge listeners face in employing such cues is that most cues are context-dependent. Here, we present an experiment that investigates the extent to which listeners learn situation-specific adjustments in the information and/or weight of the lexical bias of a verb. Specifically, we ask to what extent comprehenders are able to rapidly change their interpretation of lexical cues to syntactic structure, where such behavior would be rational due to situation-specific statistics in the environment.

Keywords: Language Comprehension; Ambiguity Resolution; Learning Effects; Language Experience

Introduction

Across many different levels of linguistic representation, statistical regularities contained within the input have been shown to influence language processing and acquisition (Christiansen & Chater, 2001; Elman et al., 1996; Seidenberg & McClelland, 1989, to name only a few). These types of findings are typically captured by connectionist accounts in which learning and processing are often viewed as inseparable (Chang, Bock, & Dell, 2006). Every instance of exposure to linguistic input causes adjustments, however subtle, to the systems responsible for processing it. This process of gradual adjustment affords the ability to approximate, over time, the mathematical functions that underlie the relevant input (Elman, 1993). Within these experience-driven approaches, it follows naturally that learning about the regularities in language by processing them is continuous—it does not end at some discrete point of time in ontogenetic development, but instead may occur across the life-span.

This raises questions as to why it is the case that lifelong learning should be an advantageous strategy? At the level of speech perception, the utility of a flexible system is relatively easy to surmise. There exists a famous "lack of invariance" in the acoustic signal. Factors such as speech rate, linguistic and acoustic context, and speaker-based idiosyncrasies contribute to the presence of a wide range of variability in the signal. In light of this variability, a perceptual system designed to *adapt* to context-, situation, or even person-specific properties of a communicative context is a necessity during speech perception. Indeed, at the level of phonology, an emerging literature has started to demonstrate that listeners update their expectations about speech sound categories in a speaker-specific manner. For example, Kraljic and Samuel (2006, 2007) have demonstrated that speakers readily adapt phonetic category boundaries in order to accommodate the statistical properties of perceptual input (see Pardo & Remez, 2006, for an overview), but do so only when speaker adaptation can feasibly aid the communicative process (Krajlic, Brennan, & Samuel, 2008). Adult listeners also adapt to group-specific phonological properties: with sufficient exposure to non-native speakers of the same native language (e.g. English produced by native speakers of Chinese), listeners show improved comprehension of novel speakers of the same foreign accent (Bradlow and Bent, 2008). In other words, in the face of great inter-speaker variability, adaptation is functionally advantageous. The learning that leads to such adaptation is implicit, in that learners are never aware of the adjustments they make (and would not, for example, be able to describe them). In short, the mapping between acoustic percepts and phonological categories seems to be malleable throughout one's lifetime, allowing speaker-, group-, and context-specific adaptation, which in turn facilitates language understanding.

At the level of syntax, however, the degree to which the probabilistic knowledge of linguistic structure that is built up over a lifetime of exposure can be temporarily amended to accommodate context-dependent aspects of a communicative situation remains very much an open question. This is the question we seek to address in this paper. Specifically, we ask whether comprehenders rapidly adapt to context- or speaker-specific deviations from typical lexical statistics.

In the domain of on-line language comprehension, one well acknowledged finding is that people rely heavily upon syntactic expectations, which have been argued to contribute to the fast and robust nature of language understanding. The expectations that are responsible for on-line comprehension likely arise, at least in part, from multiple probabilistic cues to structure working in concert to constrain interpretation (MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998.). Indeed, probabilistic cues such as referential context (Altmann, Garnham, & Dennis, 1992), plausibility (Garnsey, Pearlnutter, Myers, & Lotocky, 1997), and prosody (Snedeker & Yuan, 2008) have all been shown to modulate the interpretation of an incoming sequence. Inherent to claims about multiple-cue integration in language comprehension is the notion that an individual cue,

in and of itself, is only probabilistic, such that multiple probabilistic cues are employed as soon as possible in pursuit of understanding.

Lexical biases of verbs, such as how often a verb like *warned* is likely to be used as a main verb (as in, The soldiers *warned* the townspeople) versus as the verb of a relative clause (as in, The soldiers *warned about the dangers* were nervous) constitute another cue that has been shown to reliably influence the ease with which incoming information can be processed (Garnsey et al., 1997; Trueswell, Tanenhaus, & Kello, 1993). The verb *warned* is over three times more likely to be used as a main verb in a sentence than it is to be used as the onset of a relative clause sentence (Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1994). Even young children are sensitive to the biases of verbs, such that their processing preferences mirror the probabilistic nature in which verbs are used in child-directed speech (Snedeker & Trueswell, 2004). Verb-bias, then, is a factor likely to drive the strength of expectations about which of a set of possible verb uses is most likely.

Although a lot is known about how learning gives rise to expectations, the literature on how those expectations can change in adulthood is substantially less clear. From the syntactic priming literature, we know that exposure to a prime structure can rapidly change syntactic expectations in later sentences (Tothathiri and Snedeker, 2008). According to some accounts of syntactic priming, these rapid changes are due to implicit learning (Bock and Griffin, 2000; Chang et al., 2006; Snider and Jaeger, submitted), although this issue is still under debate (Dubey, Keller, & Sturt, 2008). In line with the hypothesis that syntactic priming may be related to implicit learning, there is now preliminary evidence that recently processed stimuli can cumulatively affect expectations beyond the effect of the most recent prime (Farmer, Monaghan, Misyak, & Christiansen, in press; Fine, Qian, Jaeger, and Jacobs, 2010). On a longer time-scale, Wells et al. (2009) demonstrated that after a multi-day training regiment, subjects' reading patterns converged with the statistical regularities to which different groups of subjects were exposed over training. Common to all of these studies, subjects read a subset of trials in which some sort of cue-based expectation is violated, akin to *warned* being used to head a relative clause instead of being used as a main verb. And, in each experiment, there was evidence that subjects were progressively more able to adapt to the specific statistical patterns contained in the material.

The goal of the experiment reported here is to examine the degree to which people are able to learn about the statistical contingencies inherent to a set of sentence materials simply by reading them. By learning here, we do not mean simply short-term activation of a structure. Instead, we examine whether the learning effects are more long-lasting, or, extend beyond one item after the prime. Through statistical analyses that allow one to assess the continuous change in expectation strength across an experiment, we provide evidence that subjects adapt their syntactic expectations in response to repeated exposure to violations of verb bias. Moreover, the changes in expectations are not immediate; the change appears to reflect a more protracted adjusting of distributional knowledge about how verbs are used in a specific context.

Experiment

Subjects were presented with a sentence set originally utilized in MacDonald, Just, and Carpenter (1992) to elicit a garden-path effect.

1 (a) The experienced soldiers / warned about the dangers / **before the midnight** / raid.

(b) The experienced soldiers / spoke about the dangers / **before the midnight** / raid.

(c) The experienced soldiers / warned about the dangers / **conducted the midnight** / raid.

(d) The experienced soldiers / who were warned about the dangers / **conducted the midnight** / raid.

That is, in example (1), the sentences containing a temporary structural ambiguity (1a and 1c) become ambiguous at segment two. The verb "warned," for example, may be interpreted as either the main verb (MV) of the sentence (1a) or as the beginning of a reduced relative clause (RC) (1c). Segment three, the point of disambiguation, contains the information necessary to arrive at the ultimately correct interpretation of the ambiguity. People have a strong bias to interpret the verb "warned" as a main verb, such that when the ambiguity is resolved in accordance with that structural interpretation of the input, little to no evidence of processing difficulty is typically detected, relative to an unambiguous control sentence (1b, where the verb "spoke" cannot head a reduced relative clause, thus producing no ambiguity). When the ambiguity is resolved in accordance with the reduced relative clause interpretation, however, processing difficulty in the form of increased Reading Times (RTs) at the point of disambiguation is observed (i.e. the garden-path effect), relative to an unambiguous control (1d, where the inclusion of the relative pronoun plus the past tense form of the verb "to be" eradicates any ambiguity).

As noted above, the lexical bias for an MV reading of the verbs in this sentence set arises from how they are used in natural language. Tabossi et al. (1994) conducted a series of corpus analyses on the ambiguity-creating verbs (like *warned* in (1), above) used in the MacDonald et al. sentences. They found that there existed an overwhelming bias against relative clause resolution for each verb: the ambiguity-creating verbs appeared over three times more in main clauses than they did in relative clauses. This frequency-based bias against RC resolution of the ambiguity is likely to be a strong contributor to the presence of a garden-path effect (i. e. increased RTs at the point of disambiguation) when the ambiguity is resolved in accordance with an RC interpretation.

In accordance with their use in natural language, the ambiguity-producing verbs used here are likely to confer a strong expectation for an MV reading of the ambiguity.

Unlike natural language, however, reading through these sentences produces a situation in which the ambiguity-producing verbs, on 50% of the trials where an ambiguity is present, are used in an RC structure, thus strongly violating the expectations that subjects likely bring with them when arriving at the experiment.

If *adaptation* occurs at the syntactic level, then the biases of the ambiguity-producing verbs should weaken some as subjects see cumulatively more examples of a violation of the bias for MV usage. That is, because subjects are being exposed to a disproportionally large number of syntactic ambiguities that are resolved in a manner inconsistent with their strong bias against RC resolution, the verb-based biases should degrade over the course of reading the sentences. This weakening of verb-bias should become apparent in the magnitude of the garden-path effect on RC resolved ambiguous sentences. Over time, the magnitude of the garden-path effect on the RC-resolved ambiguous sentences should decrease, indicating that comprehenders have adjusted their expectations, as conferred by verb-bias, to accommodate new context-dependent distributional information.

To be precise, there are two reasons to expect such a change. First, the context-specific probability of an RC increases rapidly throughout the experiment. This holds both for each verb (the probability approaches .5 in the experiment) and for the overall probability of an RC. Second, the verb, which originally was a strongly informative cue to MV vs. RC structures will become less and less informative. This might lead listeners to overall rely less on the verb as cue.

Method

Participants Seventy-two native English speakers (*M*=18.89 years, *SD*=.994) enrolled at a medium-sized Mid-Atlantic university participated in this study for extra course credit.

Materials The sentence set contained 36 sentences frames, with four versions of each frame. The manipulation, as explained above, produces a 2 (Structure: Main Verb vs. Relative Clause) X 2 (Ambiguity Status: Ambiguous vs. Unambiguous) X (Region: Point of Ambiguity vs. Point of Disambiguation) design. The four versions of the 36 different sentence frames were counterbalanced across four presentation lists in a completely crossed repeated-measures design. As a result, each list was comprised of one sentence from each of the 36 sentence frames, such that each subject saw nine of each combination of Sentence Type X Ambiguity Status, but only one sentence from each sentence frame. Fifty filler items, along with eight unrelated practice items, were incorporated into each list.

Procedure Participants were randomly assigned to one of the four presentation lists, and the presentation order of all target and filler items was randomized per subject. All sentences were presented in a non-cumulative, word-by-word moving window format (Just, Carpener, & Woolley, 1982) using Psyscope version 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993).

Subjects initially viewed a tutorial designed to acquaint them with the task. Subjects were then instructed to press the 'GO' key to begin the task. The entire test item appeared on the center (left-justified) of the screen in such a way that dashes preserved the spatial layout of the sentence, but masked the actual characters of each word. As the subject pressed the 'GO' key, the word that was just read disappeared and the next one appeared. RTs (msec) were recorded for each word. Following each sentence, subjects responded to a Yes/No comprehension question, and upon another key press, the next item appeared.

Results and Discussion

RTs on each word were length-adjusted according to a procedure described by Ferreira and Clifton (1986). First, using the raw RTs on all words in both the experimental and filler items, we computed a regression equation predicting each subject's overall RT per word from the number of characters in each word. The equation was used in order to generate an expected RT on each word *given* its length. Expected RTs on each word were then subtracted from the observed RTs, and the resulting difference score was used for all analyses.

As per the segmentation of the sentences (denoted by "/") in example (1), ambiguous and unambiguous regions of the sentences were defined in accordance with the specifications used by MacDonald, Just, & Carpenter (1992; see p. 65 for a more in-depth discussion of the intricacies associated with this type of ambiguity). The point of ambiguity in the ambiguous (1a and 1c) sentences contained the ambiguity-producing verb (i.e. *warned*) in addition to the subsequently occurring words that follow it up until the point where information exists that can facilitate disambiguation. Although no ambiguity existed in the unambiguous condition (1b and 1d) conditions, a region consisting of words that eliminated the ambiguity (verbs that don't permit an ambiguity and must be interpreted as main verbs in the main verb, unambiguous condition, such as *spoke* in 1b, or the relative pronoun + past tense version of the verb *to be* in the relative clause, unambiguous sentences condition, such as *who were* in 1d) in addition to the same words that appeared in the point of ambiguity for the corresponding ambiguous-sentence versions was also specified. The point of disambiguation for all the ambiguous-sentence conditions (and corresponding unambiguous control sentences) started with the first word that could be used to strongly support one interpretation over the other, in addition to the subsequent words remaining in the sentence, but excluding the final word of the sentence. The final word of the sentences was excluded due to sentence "wrap-up" effects, where increases in RTs frequently occur due to extra processing before subjects progress to a new item. The length-adjusted RTs on each word within a region (either the point of ambiguity or disambiguation) were averaged, per region, to create a mean length-adjusted RT within each region, and these mean length-adjusted RTs comprised the dependent measure of interest.

Here, we employ a regression approach to investigate the degree to which the influence of sentence-level variables changes over the course of the experiment (see Baayen, Davidson, & Bates, 2008, for a discussion of the advantages of linear mixed effects models in understanding how effects change over the course of an experiment). To examine the extent to which the ambiguity effect (previously reported in MacDonald et al., 1992) changed over the course of the experiment, we regressed length-corrected RTs onto the full factorial design (i.e. all main effects and higher-order interactions) of Structure (RC vs. MV), Ambiguity (Unambiguous vs. Ambiguous), Region (Ambiguous region vs. Disambiguating region) and presentation order (simply the order in the experimental list at which the item was presented, coded 1-36). As an additional control, we included a main effect of log-transformed presentation order. This was done because previous work suggests that reading times in self-paced reading studies are subject to a general speed-up effect ("task adaptation", Fine et al. 2010), which surfaces as a log-linear relationship between reading times and presentation order, illustrated in Figure 1. We chose a linear rather than log-linear presentation order effect for the higher order interactions because the linear component of presentation order essentially captures the number of previous MV and RC trials, thereby starting with the assumption of a linear effect of recent exposure. Finally, the model included the maximum random effects structure justified by the data based on model comparison using log-likelihood ratio tests. All predictors were centered in order to reduce collinearity between higher-order interaction terms. Apart from an unsurprisingly high correlation between the main effects of presentation order and log-transformed presentation order (r = -.8), all correlations between fixed effects were <.2 (we note that collinearity only affects standard error estimates of predictors that are collinear).

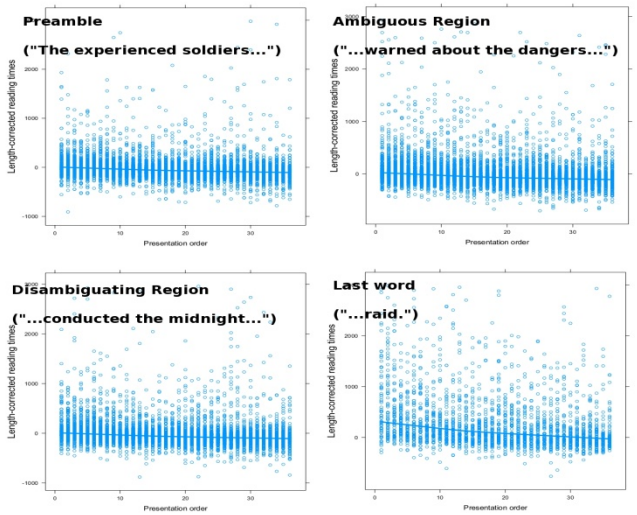

Figure 1. Collapsing across both Structure and Ambiguity Status, a main effect of Order is apparent at all regions of the sentence. RTs decrease in a protracted manner as the experiment proceeds.

Table 1. Beta coefficients associated with each main effect or interaction term in the model, along with corresponding t-values (terms with a corresponding t-value > 2 are significant at the .05 level).

	Coef β	SE (β)	t-value
Intercept	-9.28	4.66	-1.99
Structure (RC)	13.59	2.32	5.87
Ambiguity (Unamb.)	-11.56	2.32	-4.98
Region (Disamb. Region)	-0.84	3.03	-0.28
Presentation Order	-2.81	0.63	-4.43
Log Pres. Order	-26.38	6.70	-3.94
Struc. X Ambig.	-11.25	2.32	-4.85
Struc. X Region	15.57	2.31	6.73
Ambig. X Region	-12.57	2.31	-5.43
Struc. X Order	-0.73	0.33	-2.19
Ambig. X Order	0.72	0.23	3.22
Region X Order	-0.04	0.25	-0.14
Struc. X Ambig. X Region	-10.42	2.31	-4.50
Struc. X Ambig. X Order	0.73	0.22	3.26
Struc. X Region X Order	-0.76	0.22	-3.39
Ambig. X Region X Order	0.35	0.22	1.56
Struc. X Ambig. X Region X Order	0.44	0.22	2.00

Coefficients, standard errors, and t-values for each of the predictors in the model are given in Table 1. For data sets of the size used in this study, t-values with an absolute value of roughly 2 or greater are significant at $\alpha = .05$ (p-values based on MCMC-sampling in a model with only random intercepts confirm the conclusions given below). As illustrated in Table 1, the three-way Structure X Ambiguity X Region interaction was significant. Because the coefficient of this 3-way interaction is negative, a learning effect of the kind proposed here—i.e. where the ambiguity effect is weakened over the course of the experiment—would surface as an interaction between this 3-way interaction and presentation order with a *positive* coefficient. This four-way interaction is indeed significant and in the predicted direction, and is shown on the last line of Table 1. Put another way, the magnitude of the

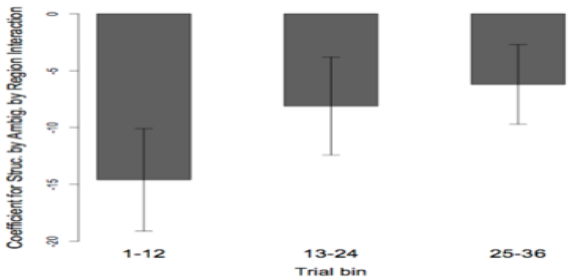

Figure 2.Coefficient size of the Structure X Ambiguity X Region interaction in "early", "middle", and "late" trial bins. Notice that as the experiment proceeds the coefficient of the 3-way interaction gets weaker and weaker (closer to 0).

3-way interaction is diminished as a result of experience (i.e. the items in the experiment), thus moving the coefficient closer to zero. This is illustrated in Figure 2, which shows the predicted effect of the 3-way interaction within the first, middle, and final third of the experiment.

The significance of the four-way interaction provides evidence that the nature of the Structure X Ambiguity X Region interaction is dependent on order. Figures 3 and 4, below, aid in decomposing the four-way interaction. In Figure 3, there is a strong *decrease* in the magnitude of the garden-path effect (i.e. a decrease in RTs at the point of disambiguation for the RC resolved ambiguous sentences, relative to the baseline unambiguous control sentences), as would be predicted if subjects were continuously modifying their expectations about the likelihood of the ambiguity-producing verb (such as *warned*) being used in a relative clause structure instead of being the main verb of the sentence. For the MV sentences, there exists no relationship between presentation order and the amount of difficulty observed when the ambiguity is resolved in accordance with the MV interpretation of the ambiguity (relative to a baseline unambiguous control sentence).

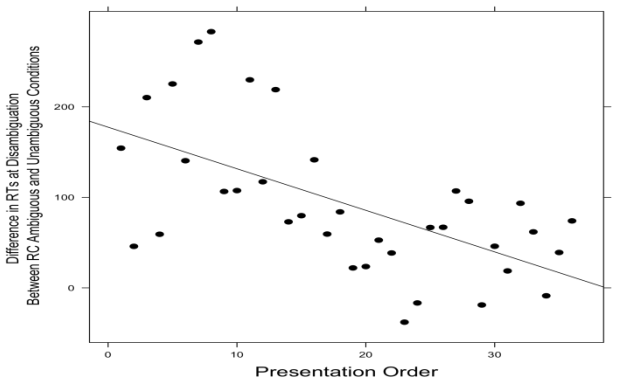

Figure 3. The ambiguity effect at the point of disambiguation for the Reduced Relative Clause sentences (RTs in Ambiguous condition minus Unambiguous condition) as a function of presentation order. The garden-path effect typically associated with RC resolution decreases continuously across the experiment.

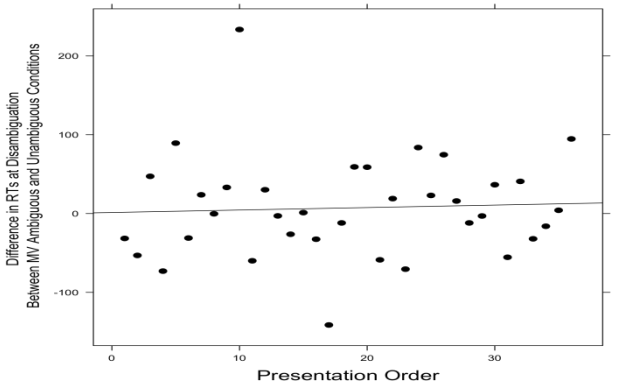

Figure 4. Ambiguity effect at the point of disambiguation for the Matrix Clause sentences as a function of presentation order. No relationship is observed.

General Discussion

Upon arriving at the experiment, subjects had a strong bias for the ambiguity-producing verbs that were used in this experiment to be used as main verbs of sentences and not as the onset of a reduced relative clause. The results reported here demonstrate that those expectations shifted over the course of the experiment as subjects encountered a cumulatively larger number of instances in which the verbs were used in the much less frequent relative clause structure. It should be noted that although the four-way interaction was significant, it was driven by RTs at the point of disambiguation when the ambiguity was resolved in accordance with the RC interpretation.

Consistent with the proposal of Fine et al. (2010), these results highlight a rarely-mentioned property of the language comprehension system. Namely, they raise the possibility that comprehenders can shift their distributional knowledge about the syntactic biases of verb behavior, as typically used in natural language, in order to adapt to the task demands conferred by a linguistic context. Such a result is consistent with input-driven accounts of language comprehension (e. g. Chang et al., 2006; MacDonald & Christiansen, 2002), which argue for a continuous updating of linguistic knowledge based on the processing of new input.

A large amount of work exists demonstrating that both infants and adults can easily learn various types of statistical regularities imbued within an artificial language (Saffran, 2003), constituting the plausibility of mechanisms for the induction of grammar. In these types of experiments, artificial languages are used instead of one's native language to control for the very same thing we study here. That is, artificial languages are used so that any pre-existing knowledge about the structure of a language doesn't influence the ease with which subjects can learn the statistical regularities under investigation. In the experiment reported here, we demonstrate that the adaptive nature of the comprehension system facilitates the ability to adjust one's prior beliefs about the structure of language (conferred, in this case, by the lexical biases of verbs), by repeated exposure to uses of a verb that are inconsistent with the biases in natural language use.

The degree to which results such as the ones reported here are generalizable to a naturalistic uncontrolled communicative context is, at this point, unclear. Unlike the phonological level, where variability in the input is the norm, the amount of variability in syntactic preferences during naturalistic conversation is less well-established. Indeed, we know little about how variability in personal history can create individual differences in how people use lexical cues to syntactic structure during language production. Work on individual differences in language comprehension provides some evidence that syntactic variability may exist. Dabrowska (1997), for example, found that a group of highly educated individuals were significantly more apt to answer comprehension questions about syntactically complex sentences in a manner that was commensurate with the grammatically correct parse of the sentence than were

groups of individuals with lower levels of formal education. Such a result suggests that language use may differ across various socio-cultural and socio-economic stratifications. And, it is situations when one is faced with such variability that syntactic adaptation from one speaker to another is likely to occur. Work is currently under way to better understand in what situations and under what conditions one would expect to find substantial syntactic variability, and whether or not adaptation occurs in these types of settings. And, even if it is the case that syntactic adaptation does occur in more naturalistic situations where one might expect adaptation based on syntactic variability in a communicative setting, it is also unclear at this point just how speaker- or context-dependent syntactic adaptation is. Experiments are currently underway to investigate whether the type of adaptation reported here stops once a subject leaves the testing room, or whether lingering effects of the adaptation process generalize to different linguistic or otherwise communicative contexts. Nonetheless, the results reported here, along with those reported in Fine et al. (2010) and Farmer et al. (in press), highlight the flexible nature of the systems responsible for language comprehension when faced with contexts in which language unfolds in unexpected ways.

References

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685-712.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed effects modeling with crossed random factors for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bradlow, A. R., & Brent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*, 707-729.

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

Christiansen, M. H., & Chater, N. (Eds.) (2001). Connectionist psycholinguistics. Westport, CT: Ablex.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1992). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257-271.

Dabrowska, E. (1997). The LAD goes to school: A cautionary tale for nativists. *Linguistics, 35*, 735-766.

Dubey, A., Keller, F., & Sturt, P. (2008). A Probabilistic Corpus-based Model of Syntactic Parallelism. *Cognition, 109*, 326-344.

Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). Rethinking innateness: A connectionist perspective on development. Cambridge, MA: The MIT Press.

Farmer, T. A., Monaghan, P., Misyak, J. B., & Christiansen, M. H. (in press). Phonological typicality influences sentence processing in predictive contexts: A reply to Staub et al. (2009*). Journal of Experimental Psychology: Learning, Memory, and Cognition.*

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

Fine, A. B., Qian, T., Jaeger, T. F., & Jacobs, R. (2010). Syntactic adaptation in language comprehension. In Proceedings of ACL Workshop on Cognitive Modeling and Computationa Linguistics. Uppsala, Sweden.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58-93.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms ad processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Kraljic, T., Brennan, S.E., & Samuel, A.G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition, 107*, 51-81

Kraljic, T. & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin and Review, 13*, 262-268.

Kraljic, T. & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language, 56,* 1-15.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35-54.

MacDonald, M. C., Just, M., A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognition, 24*, 56-98.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the effects of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 37*, 283–312.

Pardo, J. S. and Remez, R. E. (2006). 'The Perception of Speech', in M. Traxler and M. A. Gernsbacher (eds.), The Handbook of Psycholinguistics, 2nd edn. New York: Elsevier.Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*, 110-114.

Sidenberg, M. S., & McClellannd, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238-299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language, 58*, 574-608.

Snider, N. and Jaeger, T.F. (submitted) Syntax in Flux: Structural priming maintains probabilistic rep

Tabossi, P., Spivey-Knowlton, M., McRae, K., & Tanenhaus, M. (1994). Semantic effects in syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umilta & M. Moscovitch (Eds.) Attention and Performance XV: Unconscious processes. (pp. 589-616). Cambridge, MA: MIT Press.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition, 108*, 51-68

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*, 528-553.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58*, 250-271.