Error-Driven Adaptation of Higher-Level Expectations During Reading

Thomas A. Farmer[1] (thomas-farmer@uiowa.edu), Alex B. Fine[2] (abfine@illinois.edu), [1]Shaorong Yan (shaorong-yan@uiowa.edu), [3]Spyridoula Cheimariou (spyridoula-cheimariou@uiowa.edu) &
[4]T. Florian Jaeger (fjaeger@bcs.rochester.edu)

[1]Department of Psychology, University of Iowa, Iowa City, IA 52242 USA
[2]Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820 USA
[3]Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA 52242 USA
[4]Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY 14627 USA

Abstract

Fine et al. (2013) recently demonstrated that readers continually adapt their syntactic expectations in order to accurately approximate the distributions of syntactic structures in a given communicative context. Here, we examine patterns of eye movements as subjects read sentences that contain an atypical distribution of syntactic structures to gain more fine-grained insight into the time-course and nature of this adaptive process. An adaptation effect was only elicited on a late measure—second-pass reading times—consistent with the claim that expectation adaptation to an atypical distribution of syntactic structures occurs at a higher level that is abstracted away from the physical properties of the visual input.

Keywords: Language Comprehension; Eye Movements; Prediction; Adaptation; Learning; Reading

Introduction

Based on a lifetime of experience, adults possess a wealth of knowledge about the structure of language and about how linguistic events are typically distributed within a specific communicative context. Language comprehension arises by synthesizing this stored knowledge with the information available in the linguistic stimulus as it unfolds over time. Stated alternatively, prior knowledge about contextualized language facilitates the generation of expectations about what types of information are likely to be contained in the to-be-encountered portion of an unfolding linguistic signal.

Consistent with such a knowledge-driven model, it is widely recognized that readers and listeners use linguistic, visual, and social contexts to generate predictions for many aspects of the incoming linguistic input during on-line comprehension, from grammatical properties of the signal (e.g. Arai & Keller, 2013) to semantic properties of a word (e.g. DeLong, Urbach, & Kutas, 2005; Kamide, Altmann, & Haywood, 2003), all the way down to very low-level form-based perceptual properties of the physical input (e.g. Dikker, Rabagliati, Farmer, & Pylkkanen, 2010).

For example, upon encountering a single decontextualized sentence such as, *The child saved the ...* a reader isn't likely to accurately anticipate the precise word-form that they are about to next encounter (the context provided by this isolated sentence is not constraining enough), but can generate expectations for, and thus pre-activate representations for, syntactic information (a noun is likely to appear), semantic features (savable things), and physical form-based properties (physical information

probabilistically associated with words from an expected category) of an upcoming word. As the incoming signal flows from lower-level sensory processes to progressively higher levels of analysis (more abstracted away from the physical input), it is assessed with respect to progressively higher-level expectations that are generated via context.

Any mis-match between the properties of the signal and the prediction (at any level) will generate a prediction error. These "error signals" contain information about the difference between what was predicted and the structure of the input contained in the arriving signal. The error signal feeds forward to higher levels of representation, potentially facilitating an adjustment of higher-level expectations such that the predictions generated in the future may be more precise (see also Chang, Dell, & Bock, 2006; Fine & Jaeger 2013; Jaeger & Snider, 2013).[1]

Syntactic Expectation Adaptation

One consequence of this error-driven fine-tuning process is that it affords communicators the ability to rapidly adapt their expectations about properties of the multi-dimensional linguistic signal. Why is this important? Language use is highly variable, from properties of the speech signal, to lexical use, to the distribution of syntactic structures across, for example, different genres and regions. Error-driven fine-tuning of expectations affords readers and listeners the ability to deal with variability in the input. It fosters *adaptation* by providing a means through which to continuously update the expectations that are generated within a communicative context, such that comprehenders can more accurately anticipate aspects of the signal across multiple levels of analysis over the course of experience.

Expectation adaptation has been demonstrated across a wide range of perceptual and motoric tasks (Koerding & Wolpert, 2004; Kohn, 2007). From the hierarchical predictive processing account detailed above, however, expectations, and prediction errors associated with violations of them, should also occur at higher levels of analysis. As a result, we should be able to observe evidence of adaptation to those higher-level error signals as well. Indeed, Fine, Jaeger, Farmer, & Qian (2013) recently

---

[1] See Clark (2013) for a detailed perspective on "hierarchical predictive processing" (see also Farmer, Brown, & Tanenhaus, 2013, for a discussion of on-line language comprehension within such a framework).

demonstrated that a subject's *a priori* knowledge about how syntactic structures are typically distributed in the real world—knowledge that subjects carry into the testing room—can be adapted upon encountering language that contains an atypical distribution of linguistic events.

In Fine et al.'s Experiment 1, subjects were presented with a sentence set adopted from MacDonald, Just, and Carpenter (1992), as illustrated by (1a) – (1d):

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.

For sentences (1a) and (1c), the syntactic role of the verb *warned* is ambiguous. It could either act as the main verb (MV) of the sentence, or as the beginning of a reduced relative clause (RC) that modifies the subject. Although readers cannot resolve the ambiguity before encountering the disambiguating region (**bolded** in example 1), they exhibit a strong bias in favor of the MV reading. This bias stems from the fact that, in natural language, the probability of an MV/RC ambiguity-producing verb being used in an MV structure is .7 and the probability of the verb being used as the beginning of the RC is less than .01 (estimated form a large-scale corpus analysis of English, Roland, Dick, & Elman, 2007). 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 an MV, such that when the ambiguity is resolved in accordance with the MV interpretation, little to no evidence of processing difficulty is typically detected, relative to an unambiguous control sentence (1b, where the verb *spoke* cannot head an RC, thus producing no ambiguity). When the ambiguity is resolved in accordance with the RC interpretation (1c), and thus in a manner that is inconsistent with the reader's expectations, 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 "who were" eradicates the ambiguity) [2].

In a word-by-word self-paced reading (SPR) experiment, Fine et al. (2013, Experiment 1) had subjects read a set of sentences containing 9 items from each of the 4 conditions in (1), coupled with 50 filler items. Note here that the

probability of encountering an MV/RC ambiguity that is resolved in accordance with the RC interpretation increases substantially relative to subjects' *a priori* distributional knowledge, from less than .01 (based on estimates of the statistics of written English, see Fine et al., 2013) to .5. As such, the distribution of syntactic structures in the experimental setting is shifted relative to the subjects' experience prior to the task. Over the course of the experiment, the difference in RTs between RC-Ambiguous and RC-Unambiguous sentences at the disambiguation point decreased until there was no longer a statistically detectable difference in RTs between the two conditions at disambiguation.

The decrease in the magnitude of the garden-path effect was incremental and progressive (see Fine et al., Figure 5). Fine et al. interpreted this as evidence that subjects were continuously updating their expectations about the probability with which they would encounter an RC sentence. That is, subjects *adapted* their expectations about the distribution of syntactic structures in the novel context by integrating the statistical regularities of the local environment with their prior knowledge.

Experiment

The results of Fine et al. (2013) provided evidence that rapid adaptation can occur at high levels of representation. One potential problem with such an interpretation, however, is that syntactic structure is necessarily correlated with lower-level aspects of the stimulus. For example, RC disambiguation can be achieved with past participle verbs (1c), and MV resolution can be achieved with adverbs (1a), but in reality, disambiguation can be achieved by words from a host of categories (although more so in the MV case than the RC case, where many fewer disambiguation possibilities exist). A syntactic expectation is thus probabilistically yoked to lower-level lexico-syntactic category information (of, for example, a word that is intended to disambiguate a structural ambiguity). Complicating matters, probabilistic relationships exist between the lexical category of a word and its physical form-based features (for example, nouns have physical properties that differ, on average, from verbs, see Farmer, Christiansen, & Monaghan, 2006).

Word-by-word RTs obtained in SPR experiments provide only a coarse-grained index of the underlying processes that drive on-line comprehension. Visual word recognition, syntactic processes, and semantic interpretation are all a necessary part of incremental interpretation, and all of these processes are projected onto a uni-dimensional index of processing difficulty, i.e. reading time. Acknowledging the inter-correlations that exist between multiple levels of the stimulus hierarchy, one question arises with respect to the Fine et al. (2013) data: *Do changes in patterns of RTs directly reflect shifts in high-level syntactic representations?*

The presence of an atypical distribution of syntactic structures within a linguistic context will often entail context-specific shifts in distributions of correlated semantic,

---

[2] This ambiguity is disambiguated by the syntactic category of the first word at "disambiguation". In reality, however, disambiguation is never guaranteed (for example, *before the midnight raid* in (1a) could be part of a larger RC structure). We use this item set because it reliably elicits a strong garden-path effect on RTs upon RC disambiguation, and thus allows us to examine adaptation to atypical distributions of syntactic structures.

category-, and form-based information. In principle, incremental changes in patterns of RTs over the course of experience with the experiment (context) could be caused by summative learning-based shifts in expectations at all affected levels, or across a subset of them.

One implication of the hierarchical predictive processing account, as sketched above, is that different sources of information are likely to be assessed at different points of time along the continuum running from the concrete physical properties of the stimulus up through its higher-level abstract properties. Monitoring eye-movements during reading provides a variety of more fine-grained measures that can shed light on what processes are affected by encountering atypical distributions of linguistic events, and relatedly, when in the stream of processing expectation adaptation occurs (see Clifton, Staub, and Rayner, 2007 for an overview of the differential types of information that tend to influence different EM measures). Eye-movement measures have historically been categorized based on whether they are early (first-fixation duration, first-pass RT) versus later (regressions into a previously read segment, second-pass reading times, and total reading times) measures (see Clifton et al., 2007, for a more detailed discussion). Historically, "earlier" measures are more likely to be significantly influenced by factors that are associated with visual word recognition, such as frequency, length, and word predictability (e.g. Rayner, Ashby, Pollatsek, & Reichle, 2004), and less likely to be influenced by higher-level syntactic factors. Late measures have often been argued to index a reanalysis process (e.g. Frazier & Rayner, 1982, but see Bicknell and Levy, 2010 & von der Malsburg & Vasishth, 2011).

Here, we monitored eye-movements as subjects read a modified version of the materials from Fine et al (2013). Although we would not advocate for the explicit mapping of a specific EM measure to a specific psycholinguistic process, we rely on this "early versus late" distinction to provide some insight into what types of information subjects adapt to or otherwise learn about in this specific experimental context. If adaptation were to show up on early measures (especially first-fixation duration, which demonstrates marked sensitivity to form-based aspects of a visual word, and which tend to be less frequently influenced by higher level processes), such a result might indicate that the RT change was driven by shifts in low-level aspects of the stimuli. Adaptation in a late measure would, however, provide support for the notion that higher-level (potentially syntactic) expectations resulted in the patterns of RT change demonstrated by Fine et al.

Method

Participants Ninety-three native English speaking undergraduates from the University of Iowa participated in the experiment.

Materials We implemented the experimental design described in Fine et al. (2013). For each item frame (1), four

sentences were created (1a-d), with one version corresponding to each of the 4 conditions created by the 2 x 2 manipulation (Structure: MV vs. RC) x Ambiguity Status (Ambiguous vs. Unambiguous). Sentences were counterbalanced across four presentation lists such that each subject only saw 1 version of each item, but an equal number of trials per condition. We reduced the number of items from 36 to 24 in order to eliminate any overlap in exposure to the ambiguity-producing verb (thus reducing any effects of "lexical boost" associated by repetition of the ambiguity-producing verb (Pickering & Branigan, 2008), which was minimally present in the materials used in Fine et al.'s Experiment 1. Subjects were also presented with 72 filler items that did not contain relative clause structures.

Procedure Eye movements were recorded with an EyeLink 1000 eye-tracker at a sampling rate of 1,000 Hz. Viewing was binocular but data were only recorded from the right eye. Stimuli were presented with a 19-inch ViewSonic CRT monitor with a 1024*768 pixel resolution. Subjects were seated ~60 cm from the screen, with their head positioned on a chin rest.

Participants were randomly assigned to one presentation list, and presentation order was randomized per each subject. Each sentence remained on the screen until the subject pressed a button on a hand-held controller to proceed to a Y/N comprehension question.

Results and Discussion

In order to provide to provide an analysis of these data that best parallels Fine et al.'s SPR data, we used the same segment delineations as both Fine et al. and MacDonald et al. (1992), denoted by "/" in example (1). The primary region of interest, the disambiguation region (bolded in (1)), includes the first word of disambiguation followed by all subsequent words, but excluding the final word of the sentence. In order to further investigate where the prediction error was most strongly experienced, we also conducted the analyses at disambiguation when each of the first three words of disambiguation were delineated into different segments.

Both first-pass measures (First fixation times, First-pass times, Go-past times, Probability of regressing out leftward during first-pass) and later measures (Total reading times, Second-pass times, & Probability of regressing back into a region) were computed for each segment. Fixations less than 80 ms in duration and less than one character away from the closest fixation were incorporated into the previous fixation. Fixations longer than 800 ms or less than 80 ms were then excluded, leaving over 99% of observations for analysis. Linear mixed-effects models and mixed logit models (for regression probabilities) were adopted in analyzing the measures, and the analyses were implemented with the lme4 package (Bates & Maechler, 2010) in the R environment. Each independent variable, Structure (MV vs. RC), Ambiguity Status (Ambiguous vs. Unambiguous), Item order (the presentation order of an item among critical items) was entered into the model as a fixed effect with a full factorial

design. Log stimulus order (the presentation order of an item among all items including fillers) was also included to control for general practice effects (Fine et al. 2010). Word length and frequency were added to the model when modeling the single word analyses. All predictors were centered to reduce collinearity with higher-order interaction terms. The maximum random effect structure by subjects and items was identified for every model detailed in this paper based on model comparison using log-likelihood ratio tests. T-values (for linear mixed-effect models) and z-values (for mixed logit models) for the analyses on the disambiguating region that are larger than 1.96 were interpreted as significant at the .05 level.

Garden Path Effect Given that syntactic expectation adaptation is defined as the updating of syntactic expectations (in this case, the probability of encountering an RC relative to an MV structure), as a result of an error signal produced by an higher-level expectation violation (in this case, RC resolution of an ambiguity produced by a verb that is strongly biased against being used at the beginning of an RC), we first examine the data for evidence of a garden-path effect. Garden-path effects, in the form of longer RTs or more regressive reading on RC ambiguous sentences, were observed in various measures across the eye-movement record. This effect was represented in the model by the positive interaction of Structure and Ambiguity (see Table 1).

Similar to many of the previous EM Experiments summarized by Clifton et al., when the full disambiguation segment was analyzed as a whole, garden-path effects appeared on second-pass reading times, a late measure, but unlike some (but not all) of the previous experiments, no garden-path effect was observed on measures that are traditionally considered to be "first-pass" in nature. There was, however, a garden-path effect on first-pass regressions out. In the word-by-word analysis of the disambiguation segment, the garden path effect was observed for go-past time, and again on second-pass RTs and regressions out.

Table 1. Garden-path effect in Ambiguous/Disambiguating region (represented by t/z-values for the coefficient for Structure*Ambiguity interaction).

Measurements	Disambiguating Region	First Word	Second Word	Third Word	Ambiguous Region
First Fixation Times	0.06	-0.44	0.28	-0.18	1.04
First-pass Times	-1.82	0.97	-0.08	-0.22	1.29
Go-past Times	0.44	2.83**	2.34**	0.66	3.88**
Total Reading Times	0.06	1.36	2.91**	2.02**	2.63**
Second-pass Times	3.86**	5.86**	3.44**	1.13	6.57**
First-pass Regressions Out	4.99**	3.33**	1.02	1.02	0.12
Regressions-in	1.14	1.24	1.73	-0.74	6.24**

Adaptation Effects The garden-path effect associated with RC resolution does not arise in the earliest eye movement measures reported here, is present in many of the later measures, and is strongest in second-pass RTs. We inspected the data for evidence of RC adaptation by examining the three-way interaction between Structure, Ambiguity Status, and Item order. No adaptation effect was

found for traditional first-pass measures, either when examining the combined disambiguation segment, or when examining each measure on each separate word of disambiguation (Table 2), but see below for a discussion of a marginal adaptation effect on regressions out.

Table 2. Estimated coefficients for first-pass eye-movement measurements at disambiguating region.

Variables	First Fixation Times		First-pass Times		First-pass Regressions Out		Go-past Times	
	Coef β	t-value	Coef β	t-value	Coef β	z-value	Coef β	z-value
Structure(RC)	10.145	4.25**	29.780	2.63**	0.072	0.91	42.528	0.75
Ambiguity status(Ambig.)	4.480	1.68	-1.199	-0.17	0.206	3.75**	36.636	1.24
Item order	0.151	0.27	-4.297	-1.98	-0.002	-0.13	-6.224	-2.27**
Log stimulus order	-2.224	-0.23	33.453	0.90	-0.090	-0.32	6.472	0.14
Struc. : Ambig	0.104	0.06	-13.132	-1.82	0.274	4.99**	27.122	0.44
Struc. : Item order	0.328	1.19	-0.842	-0.80	0.008	0.97	0.255	0.20
Ambig. : Item order	0.417	1.52	0.181	0.17	0.007	0.92	2.416	1.88
Struc. : Ambig : Item order	0.019	0.07	-0.830	-0.79	0.000	0.01	-0.757	-0.59

For the late measures, there was a near-significant adaptation effect for second-pass times (β=-1.967, t=-1.90, p = .057), and no other reliable adaptation effects were identified (see Table 3). As illustrated in Figure 1, this interaction was in the predicted direction, such that the difference in second-pass RTs at disambiguation between the Ambiguous and Unambiguous sentences decreased over time, while no statistically distinguishable RT differences existed between ambiguous and unambiguous MV sentences over time. As subjects encountered more instances of RC-resolution, they spent less time re-reading the disambiguating region, commensurate with the SPR data in Fine et al., which demonstrated a decreasing difference in RTs at disambiguation between the ambiguous and unambiguous sentences in the RC condition over time.

Table 3. Estimated coefficients for examined eye-movement measurements in the disambiguating region (later measures).

Variables	Total Reading Times		Second-pass Times		Regressions-in	
	Coef β	t-value	Coef β	t-value	Coef β	z-value
Structure(RC)	-17.692	-0.24	34.659	2.79**	-1.434	2.04**
Ambiguity status(Ambig.)	-16.409	-0.83	27.136	3.63**	0.112	0.42
Item order	-0.540	-0.16	-3.118	-1.49	0.023	-0.39
Log stimulus order	-34.622	-0.59	-41.625	-1.15	-0.006	-1.52
Struc. : Ambig	4.931	0.06	27.072	3.86**	-0.402	1.14
Struc. : Item order	2.592	1.58	-1.233	-1.19	0.063	1.28
Ambig. : Item order	1.409	0.85	-0.781	-0.76	0.010	-0.13
Struc. : Ambig : Item order	-0.260	-0.16	-1.967	-1.90	-0.001	-0.70

Re-analysis of SPR data from Fine et al. (Exp. 1) To compare effects across the EM and SPR data, we re-analyzed the original Fine et al. data with the disambiguation region segmented in the same 5 ways used in the analysis of EMs above. First, we reproduce the analysis reported by Fine et al. (2013), in which we analyze length-corrected RTs *averaged* across all three words of the disambiguating region. This analysis replicated the key findings in the eye-tracking experiment, producing a significant Structure x Ambiguity x Item Order interaction, β=-0.5, p < .05.
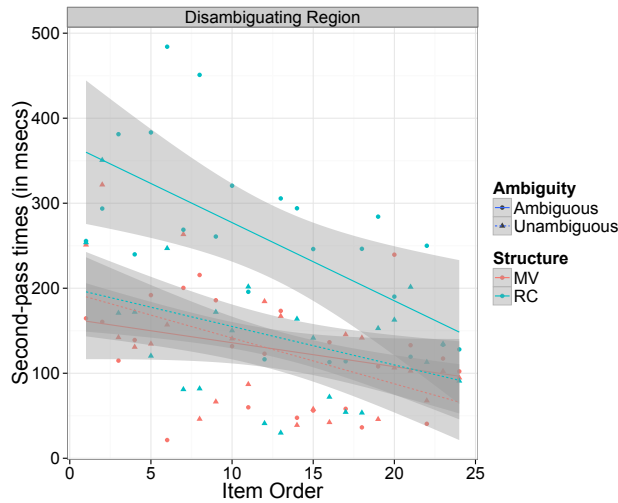
Figure 1. Adaptation effect: The garden path effect (longer reading time for RC ambiguous sentences than the other three) decreased relative to Item order for Second-pass times.

In order to determine when exactly in the disambiguating region the adaptation effect occurs, we then analyzed length-in the disambiguating region, as well as for the RTs averaged across the first and second words in this region. Because averaging across the entire disambiguating region reduces noise attributable to idiosyncratic lexical differences at each individual word position, the results of the three models fitted to individual words were slightly less statistically reliable than those of the model conducted over the whole-region averages. Nevertheless, a key point about the locus of the adaptation effect can be discerned, which is that the effect trends towards significance or is marginally significant and in the predicted direction at both the first word in the disambiguating region (β = -0.7, p = .16), the second word in the disambiguating region (β = -0.7, p = .09), and in the combined data from these two words (β = -0.7, p = .06);
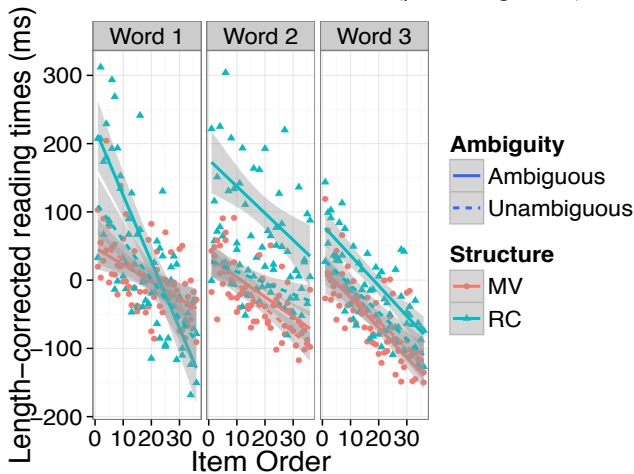

Figure 2. Adaptation effect from the SPR experiment plotted separately for words 1, 2, and 3 in the disambiguating region.

however, the effect does not approach statistical significance at the third word in the disambiguating region (β = 0.1, p = .8). In sum, the adaptation effect seems to be occurring early in the disambiguating region.

General Discussion

Upon arriving at the experiment, subjects had a strong bias in favor of MV resolution of the MV/RC ambiguous sentences. The results reported here, and in Fine et al. (2013), demonstrate that those expectations shifted over the course of the experiment such that the MV bias faded incrementally and progressively as the subject encountered progressively more instances of RC resolution. But, as evident through the new data presented here, this adaptation effect appears to occur only on a later measure (second pass reading times). In both the EM and SPR data, the adaptation effect seems to be most robust when RTs elicited by multiple words of a disambiguating segment are averaged together, and on only the first word of disambiguation.

A syntactic expectation adaptation effect on a late measure is generally expected under a hierarchical predictive processing framework—expectations for abstract aspects of a stimulus should be evaluated later in the stream of processing a stimulus than expectations that have been generated for more concrete physical properties of the stimulus. The "early" versus "late" distinction is a simplifying heuristic that does not necessarily entail early versus late processing (see Clifton et al. for discussion). But, the fact that there was no evidence of adaptation on the traditional first-pass measures, but only on a late measure, supports the claim that changes in patterns of RTs on RC disambiguated sentences do reflect shifts in higher-level syntactic expectations, as argued by Fine et al.

It is important to note, however, that there was some evidence of adaptation, in the form of a three-way interaction with a p-value of .095, on the first-pass regressions out measure. The probability of regressing leftward upon fixating the first word of disambiguation on first-pass increased on the RC Ambiguous sentences, and decreased on the RC Unambiguous sentences, over the course of the experiment. Depending on one's view of first-pass regressions out (e.g. Altmann, 1994; Rayner & Sereno, 1994), such a result may be indicative of some lower level learning. It is important to note that no other near-significant trend was identified in the analyses conducted per-word (or first two words combined) in the EM record (not reported above due to space constraints).

One deviation from the original Fine et al. results involves removing the presence of verb overlap from the materials. It is possible that verb overlap provided a boost in the adaptation effects observed in Fine et al. If so, then the removal of verb overlap may explain the reason that the only adaptation effects observed here were weak. Recent work by Fine & Jaeger (in prep), however, has not been able to detect evidence that error-driven adaptation on similar sets of syntactically ambiguous materials is dependent upon verb overlap. Additionally, Fine et al. demonstrated that the

adaptation effect described here was still robust after all items that contained the same verb were removed, except for the first item that contained the verb.

Hierarchical predictive processing frameworks (e.g. Clark, 2013) of on-line language comprehension may prove viable in terms of facilitating an understanding of how prior knowledge and top-down contextual information modulate the perception and interpretation of a physical signal during the lower-level, and even perceptual, processing of a linguistic signal. More globally, we believe that such an account has the ability to help unify prediction-based accounts of contextualized language processing by guiding work on questions related to what's being predicted, and when in the chain of processing those predictions are generated and assessed.

Acknowledgments

References

Altmann, G. T. M. (1994). Regression-contingent analyses of eye-movements during sentence processing: Reply to Rayner and Sereno. *Memory and Cognition, 22*, 286-290.

Arai, M. & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28*, 525-560.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes.

Bicknell, K., & Levy, R. (2010). Rational eye movements in reading combining uncertainty about previous words with contextual probability. *In Proc. 23rd annual conference of the Cognitive Science Society.*

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

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181-204.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In van Gompel, Fischer, Murray, & Hill (Eds). *Eye movements: A Window on the mind and brain* (pp. 341-372). Amsterdam: Elsevier.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience, 8*, 1117-1121.

Dikker, S., Rabagliati, H., Farmer, T. A., Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science, 21*(5), 629-634.

Farmer, T. A., Brown, M., & Tanenhaus, M. K. (2013). Prediction, explanation, and the role of generative model in language processing. *Behavioral and Brain Sciences, 36*,

211-121.

Farmer, T. A., Monaghan, M. C., & Christiansen, M. H. (2006). Phonological typicality influences sentence processing. *Proceedings of the National Academy of Sciences of the USA, 103*, 12203-12208.

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.

Fine, A.B. and Jaeger, T. F. (2013). Evidence for error-based implicit learning in adult language processing. *Cognitive Science,* 578-591.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. PLoS ONE, 8: e77661.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology,* 178-210.

Jaeger, T. F., & Snider, N. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given bothprior and recent experience. *Cognition, 127*, 57-83.

Kamide, Y., Altmann, G. Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133-156,

Koerding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*, 244–247.

Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, *97*(5), 3155–64.

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

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

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader Model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720-732.

Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences*, *26*, 518–526.

Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition,* 22, 281-285.

von der Malsburg, T. and Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language*, 65,109-127

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects in eye movements and E-Z Reader. *Cognition, 111*, 132-137.