



The time course of anticipatory constraint integration

Anuenu Kukona^{a,b,*}, Shin-Yi Fang^{a,b}, Karen A. Aicher^{a,b}, Helen Chen^b, James S. Magnuson^{a,b}

^a Department of Psychology, University of Connecticut, United States

^b Haskins Laboratories, New Haven, CT, United States

ARTICLE INFO

Article history:

Received 4 February 2008

Revised 27 November 2010

Accepted 1 December 2010

Available online 14 January 2011

Keywords:

Sentence processing

Thematic roles

Anticipation

Eye tracking

ABSTRACT

Several studies have demonstrated that as listeners hear sentences describing events in a scene, their eye movements anticipate upcoming linguistic items predicted by the unfolding relationship between scene and sentence. While this may reflect active prediction based on structural or contextual expectations, the influence of local thematic priming between words has not been fully examined. In Experiment 1, we presented verbs (e.g., *arrest*) in active (Subject–Verb–Object) sentences with displays containing verb-related patients (e.g., *crook*) and agents (e.g., *policeman*). We examined patient and agent fixations following the verb, after the agent role had been filled by another entity, but prior to bottom-up specification of the object. Participants were nearly as likely to fixate agents “anticipatorily” as patients, even though the agent role was already filled. However, the patient advantage suggested simultaneous influences of both local priming and active prediction. In Experiment 2, using passive sentences (Object–Verb–Subject), we found stronger, but still graded influences of role prediction when more time elapsed between verb and target, and more syntactic cues were available. We interpret anticipatory fixations as emerging from constraint-based processes that involve both non-predictive thematic priming and active prediction.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

A listener's interpretation of language is shaped not simply by the words and phrases in an utterance, but also by the potentially rich context in which the utterance is embedded. The question of *when* contextual information impacts a listener's interpretation is a point of serious contention among theories of sentence processing (e.g., Frazier & Fodor, 1978; MacDonald, Pearlmutter, & Seidenberg, 1994), and has been the focus of considerable research in psycholinguistics (e.g., Duffy, Morris, & Rayner, 1988; Ferreira & Clifton, 1986; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Recent studies have demonstrated that language users are able to use linguistic and non-linguistic context to

anticipate upcoming information (e.g., Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Ferretti, McRae, & Hatherell, 2001). These results suggest that various constraints quickly conspire to afford robust prediction of upcoming syntactic constituents and even specific words (for related results, see Pearlmutter & MacDonald, 1995), and have motivated theories that posit active forecasting mechanisms in sentence processing (e.g., Altmann & Kamide, 2007; Knoeferle & Crocker, 2007). On such views, the full complexity of bottom-up, top-down, and contextual information sources is analyzed with the explicit goal of predicting upcoming input. Our aim in the current work is to ask what role simpler mechanisms, like priming, might play in anticipation.

In the following section, we briefly review relevant findings concerned with anticipatory eye movements in language. While a handful of these findings are only consistent with active forecasting, many of these results could be explained in whole or in part by simple mechanisms like priming. In the subsequent sections, we describe

* Corresponding author. Address: Department of Psychology, University of Connecticut, Storrs, CT 06269, United States.

E-mail address: anuenu.kukona@uconn.edu (A. Kukona).

a framework for anticipatory effects based in part on local, lexical-level structure, and describe related evidence.

1.1. Evidence for anticipation

In 1995, Tanenhaus et al. launched what is now known as the *visual world paradigm*. Listeners' eye movements are monitored as they view (and often interact with) a visual scene and listen to speech (for a related precursor, see Cooper, 1974). Subsequent studies employing this paradigm have demonstrated that listeners literally “look ahead” when the conjunction of linguistic information and visual context affords strong expectations for upcoming referents. Altmann and Kamide (1999) first demonstrated this by presenting visual scenes that included one edible object (e.g., *cake*) and various other inedible objects (e.g., *ball*, *train*, etc.) while listeners heard a spoken utterance that contained either the verb *eat* or *move* (e.g., *The boy will eat/move the cake*). Listeners were able to constrain their interpretations to plausible direct objects immediately based on the selectional constraints of the verb *eat* (after which they immediately began looking at the *cake*), but had to wait for full specification of the direct object in the case of the verb *move*. Altmann and Kamide (1999) argued that the mechanisms underlying sentence processing use the fit between scene (e.g., the items in the visual world) and sentence (e.g., thematic and event knowledge associated with the verb) to restrict the possibilities for subsequent reference, and to direct eye movements based on these restrictions. They suggested that thematic fit was relevant in the case of *eat*, but proposed more broadly that “any and all information that is available is recruited to the task of predicting subsequent output” (p. 262).

Kamide, Altmann, and Haywood (2003) reported further evidence for predictive context integration in the visual world paradigm based on linguistic constraints arising more globally across multiple words. Their displays depicted two possible agents (e.g., *man* and *girl*) and two ride-able objects (e.g., *motorbike* and *carousel*). Their sentences described who would ride what (e.g., *The girl/man will ride the carousel/motorbike*). With the subject *girl* and the verb *ride*, eye movements anticipatorily favored the object best fitting the combined semantic constraints of the subject and verb (e.g., *carousel*, not *motorbike*). While Kamide, Altmann, et al. (2003) discuss the possibility that anticipatory fixations could result from a system that does not actively predict upcoming referents, they argue that their results are most consistent with “an incremental processor that establishes the fullest possible interpretation at each moment in time” (p. 153). This suggests a processor that predicts upcoming referents based on constraints that include syntactic, semantic, and environmental context, and the explicit tracking of already specified roles.

Chambers and San Juan (2008) have also demonstrated that anticipatory eye movements are driven by dynamic, “situation-specific” knowledge concerning the relationship between scene and sentence. They had participants follow a sequence of instructions in a visual world task. The first instruction involved moving an object (e.g., *Move the chair to area 2.*), and the second instruction used either the verb *move* again (e.g., *Now move the chair to area 5.*) or the verb

return (*Now return the chair to area 5.*). They found reliably more anticipatory fixations to the chair during the second instruction with the verb *return* as compared to *move*, suggesting that participants used “situation-specific” knowledge about prior locations of objects to anticipate upcoming linguistic information.

Several visual world studies, including Chambers and San Juan (2008), provide clear support for active forecasting (see also Altmann & Kamide, 2007; Knoeferle & Crocker, 2006, 2007). Studies using other methods are also consistent with active forecasting. For example, readers exhibit an N400 (an ERP component that is sensitive to how surprising a word is) upon encountering a determiner that does not fit with an expected upcoming noun (DeLong et al., 2005; see also Wicha, Moreno, & Kutas, 2004, and for related results in Dutch, see Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). However, this does not rule out a role for priming in many other cases where anticipatory effects have been reported. In the next section, we review evidence that suggests that priming may play an important role in sentence processing.

1.2. Is there a role for priming?

The conclusions of Altmann and Kamide (1999) and Kamide, Altmann, et al. (2003) support an explicit forecasting mechanism that attempts to complete syntactic constructions via active, anticipatory search. However, another body of visual world paradigm findings suggests that seemingly anticipatory fixations could be driven at least in part by simpler processes of semantic priming, which could include priming based on thematic relations. Huettig and Altmann (2005) and Yee and Sedivy (2006) found that fixations were readily directed toward items related to a spoken word by semantic association, category, or function. For instance, Yee and Sedivy (2006) presented a visual array to listeners that included a semantically related pair of objects (e.g., *lock* and *key*) and other unrelated distractor objects. When participants heard *key*, they fixated the picture of the semantic competitor *lock* reliably more than other semantically unrelated distractors in the scene.

As Altmann and Kamide (1999) emphasize, verbs provide considerable constraints that may largely drive anticipatory fixations. Verbs have been shown to prime related nouns without visual world (or any other) context. Using single-word priming, Ferretti et al. (2001) found that verbs primed noun targets that were typically involved in the situations described by the verb. For example, when noun targets followed related verb primes (e.g., *arrest* – *cop*), judgments of animacy were speeded. Nouns with varying thematic roles were primed, including typical performers (agents), and typical recipients (patients). The converse also holds: nouns prime event-relevant verbs (McRae, Hare, Elman, & Ferretti, 2005).¹

Kuperberg (2007) reviews several ERP findings that suggest a role for local priming. She argues that explaining the

¹ Ferretti et al. (2001) also reported results that seem inconsistent with simple local priming: following sentence fragments strongly cuing a specific role (*she arrested the...*), naming was only facilitated for role-appropriate items. We return to this point in Section 4.

interplay of influences of lexical-semantic memory, morphosyntactic processing, and thematic fit requires at least two distinct and weakly interacting processing streams: lexical-semantic memory, which proceeds in a local manner (i.e., as a priming mechanism), and a combinatorial mechanism that builds sentence structure via morphosyntactic rules and evaluation of thematic fit. One key piece of evidence supporting this view is the fact that ERP components putatively linked distinctly to semantic-level expectation (the N400) and syntactic processing (the P600) can be modulated by either information source, consistent with the hypothesis that lexical priming can play an important role in syntactic processing. For example, Kim and Osterhout (2005) pitted syntactic cues against semantic cues to thematic role in sentences like *The meal was devouring...*: while syntactic features of *meal* suggest that the verb is semantically anomalous (e.g., *cooking* would be more appropriate), semantic cues suggest the verb should be in past participle form (e.g., *devoured*) and therefore the sentence contains a syntactic anomaly. Semantic cues won: a P600 (associated with syntactic anomaly) was found rather than an enhanced N400 (associated with semantic anomaly); for similar results, see Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). Hoeks, Stowe, and Doedens (2004) explicitly pitted “lexico-semantic” relatedness against the predictiveness of sentence contexts in an ERP paradigm. They reported main effects of sentence predictiveness (lower N400 amplitude for highly predictive contexts) and lexico-semantic fit of a target word with other words in the sentence (lower N400 for high fit), but there was also an interaction: sentence predictiveness did not observably mitigate the impact of poor lexico-semantic fit, suggesting larger scale (sentence level) constraints may not override smaller scale constraints (word–word semantic fit).

These results from the ERP literature add further credence to the possibility that in situations like those used by Altmann and Kamide (1999), “anticipation” of cake given *The boy will eat...* could be driven at least in part by local priming of thematic (or event or situation) knowledge at the verb (that is, substantial priming to both good agents and objects of *eat*). Ruling out a significant contribution from priming would require displaying items that are semantically related to the verb in the utterance, but inconsistent with predictions that would follow from active forecasting given the full, global context of the sentence and the scene. Kamide, Scheepers, and Altmann (2003) used a visual world design which included such items: their displays depicted sets like a *hare*, *fox*, *cabbage*, and *tree*, and their sentences described one of the items (e.g., *hare*) in differing thematic roles in an active (e.g., *The hare will eat the cabbage*) vs. passive (e.g., *The hare will be eaten by the fox*) sentences, with thematically appropriate nouns in the other role (e.g., *cabbage* or *fox*). Consistent with active forecasting, there were more anticipatory looks at the verb in the active case to the thematically appropriate *cabbage* (30% of trials) as compared to the inappropriate *fox* (13% of trials), and there were more anticipatory looks in the passive case to the thematically appropriate *fox* (27%) as compared to the inappropriate *cabbage* (20%). However, this result does not rule out priming: there were still substantial looks to the thematically

inappropriate items in each case, and these looks were not compared with looks to an unrelated distractor (e.g., *tree*).

But what about anticipatory looks in Kamide, Altmann, et al. (2003) to *carousel* given *The girl will ride the...*, as described above? These results are not inconsistent with local priming; fixations to *carousel* could be driven by conjoint priming of event knowledge by *girl* and *ride*² (indeed, this is not incompatible with the explanation offered by the authors). All the same, there was a moderate (though statistically unevaluated) likelihood of fixating the other potential rider at the verb, and other examples of non-predictive looking later (e.g., looks back to the named agent). However, an appeal to conjoint priming reveals a difficulty in evaluating our hypothesis. Given that nouns can be primed by related nouns (e.g., Huettig & Altmann, 2005; Yee & Sedivy, 2006) or verbs (e.g., Ferretti et al., 2001), if we hypothesize that both types of priming should have additive effects (as in *girl-ride-carousel*), then a clear test will require us to use materials that differ from those of Kamide et al. Specifically, we will need materials in which the agent described in the sentence is not thematically related to the patient described in the sentence, in order to avoid thematic priming between agents and patients.

1.3. A constraint-based approach

As we have just reviewed, a handful of studies (e.g., Altmann & Kamide, 2007; Chambers & San Juan, 2008; DeLong et al., 2005; Ferretti et al., 2001; Knoeferle & Crocker, 2006, 2007) provide strong evidence for anticipation based on forecasting. However, such results do not rule out the possibility of important contributions of local thematic priming in addition (cf. Kuperberg, 2007). The complexity required of a forecasting mechanism would be reduced if some of the work could be done in an emergent fashion based on interactive activation among lexical representations. The phenomenon of “local coherence” also suggests that local- and global- (sentence) scale constraints simultaneously impact sentence processing. For example, Tabor, Galantucci, and Richardson (2004) report evidence that the active interpretation of the locally coherent phrase, *the player tossed a frisbee*, competes with the global interpretation of *the coach smiled at the player tossed a frisbee*.

Rather than pitting active forecasting mechanisms (based on global, linguistic information at the level of a sentence) against lexical-semantic priming (based on local, context-free lexical information at the level of a word), one might assume that global/contextual and local, lexical sources of constraint operate simultaneously (consistent with constraint-based theories; e.g., MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). This view is consistent with the recent position of Altmann and colleagues (e.g., Altmann & Kamide, 2007; Altmann & Mirkovic, 2009), who propose that anticipation may be governed by the feature-based (e.g., phonetic, semantic, or thematic) fit of items in the visual context with representations in “an unfolding (mental) world,” which is jointly constrained by language

² Conjoint priming from subtle featural overlap is plausible given the strong association between role/filler typicality and role/filler featural similarity reported by McRae, Ferretti, and Amyote (1997).

and context. It is also consistent with the “coordinated interplay account” of Knoeferle and Crocker (2006), which explicitly posits that the balance between basic comprehension processes (driven by stored linguistic knowledge) and active forecasting (based on predictions computed from scene and other context) shifts according to the relative availability or validity of the two information sources.

These variants of the constraint-based framework differ in important ways, but both emphasize the role of prediction, and explain anticipation as the result of a forward-looking processor that continuously predicts upcoming linguistic elements by integrating visual context, sentence context, and verb-specific thematic information, among other constraints. However, as we discuss in Section 4, there are other theories that maximize prediction (e.g., Hale, 2001; Levy, 2008) to the point that they are not compatible with, for example, local coherence effects (see Levy, Bicknell, Slattery, & Rayner, 2009, for discussion and an alternative view of the phenomena). Such views would not predict activation of representations that are compatible with already-filled roles. In the following two experiments, we asked whether anticipatory fixations were *fully* forward-directed (biased towards unfilled roles *only*), or if fixations were also directed toward (strongly) verb-related items in the visual context whose typical roles have already been filled. We carefully controlled the relationship between verbs and potential agents and patients, in order to provide a clear test of the influence of priming.

2. Experiment 1

We examined anticipatory fixations in the visual world paradigm, with active sentences like *Toby arrests the crook*, to verb-related items (good agents and patients) and verb-unrelated items, following the offset of the verb but prior to the onset of the direct object noun. “Toby” was a recurring, thematically neutral character who was the subject of each sentence. Prior to the experiment, we made Toby a highly salient subject by showing participants a picture of him and telling them that all sentences would be about things Toby does. The picture of Toby was used as the central fixation point in the display, which was visible throughout every trial (see Fig. 1). This allowed us to avoid possible agent–patient (noun–noun) priming, and to focus instead on priming from verbs to agents and patients.

For each verb, we selected a good agent and a good patient. For *arrest*, for example, the animate nouns *policeman* and *crook* play prominent but distinct roles in the situations typically described by the verb *arrest*. While *policeman* is a typical agent of *arrest* (i.e., a typical performer of the verb), *crook* is a typical patient of the verb (i.e., a typical recipient of the action of the verb). In the sentences we presented, the typical patients occupied their usual roles as the direct object of the verbs, and the typical agents were not mentioned, since the agent role was explicitly filled by Toby. In our critical condition, participants listened to these sentences while they viewed a visual display that



Fig. 1. Example competitor present display from Experiments 1 and 2. In Experiment 1, participants heard *Toby arrests the crook* in the predictive condition, and *Toby notices the crook* in the non-predictive condition. In Experiment 2, participants heard *Toby was arrested by the policeman* in the predictive condition, and *Toby was noticed by the policeman* in the non-predictive condition. Both the good patient (*crook*) and good agent (*policeman*) of the predictive verb (*arrest*) are pictured in this display, and the character Toby is pictured at the center.

included an image of both the good patient (i.e., the target) and good agent (i.e., the competitor).

The presence in the display of two semantic associates of the verb with differing thematic roles allowed us to assess lexical-semantic priming in the following way: if predictive fixations are the result of lexical priming without respect to sentence context, we expect post-verbal looking of equal proportions to both the target patient (e.g., *crook*) and competitor agent (e.g., *policeman*) of the verb (e.g., *arrest*) until there is bottom-up support for the correct post-verbal referent (i.e., the target patient is named in the sentence). However, if predictive fixations are instead driven by an anticipatory process that evaluates the event-based thematic fit of a verb with respect to a sentence context, we expect post-verbal anticipatory fixations above baseline (distractor) levels *only* to the target patient (e.g., *crook*): because the subject position has already been filled by *Toby*, an optimal processor should restrict the subsequent domain of reference at the verb to only those semantic associates that share a patient relationship with the verb. Alternatively, if processing is constrained by both local lexical semantic-priming and sentence structure, we expect graded anticipatory fixations to the target patient and to the competitor agent, such that while most fixations will be directed to the sentence-appropriate target patient, more fixations will be directed to the sentence-inappropriate competitor agent than to other unrelated distractors.

2.1. Methods

2.1.1. Participants

Sixteen students from the University of Connecticut participated for partial course credit. All participants were native speakers of English with normal or corrected-to-normal vision.

Materials. With a separate group of participants, we normed a set of 40 predictive verbs, with both a strongly associated agent and patient (e.g., *arrest* with *policeman* and *crook*), as well as a set of 24 non-predictive verb controls (see [Appendix A](#) for details). Agents and patients were animate people and animals. The mean agenthood of agents ($M = 6.64$, $SD = 0.43$, on a 7-point scale) was robustly higher than that of patients ($M = 2.78$, $SD = 1.16$; $t(39) = 18.59$, $p < .001$), and the mean patienthood of agents ($M = 3.04$, $SD = 1.07$) was robustly lower than that of patients ($M = 6.36$, $SD = 0.58$; $t(39) = 17.36$, $p < .001$). We compared the relative strength of agents and patients by calculating difference scores between the agenthood of agents and the patienthood of patients with predictive ($M = .28$, $SD = 0.67$) and non-predictive verbs ($M = .49$, $SD = 1.50$). While agents tended to be slightly stronger than patients overall (i.e., difference scores > 0), this bias was not reliably different between predictive and non-predictive verbs, $t(39) = 0.85$, $p = .40$. Agents and patients were matched on number of phonemes (agent $M = 5.68$, $SD = 1.86$; patient $M = 5.10$, $SD = 2.01$), $t(39) = 1.49$, $p = .15$, and Kučera–Francis (KF) frequency per million (Kučera & Francis, 1967) (agent $M = 29.58$, $SD = 45.98$; patient $M = 26.13$, $SD = 45.44$), $t(39) = 0.34$, $p = .74$. Predictive and non-predictive verbs were also balanced on number of phonemes (predictive $M = 5.95$, $SD = 1.54$; non-predictive

$M = 6.40$, $SD = 1.63$), $t(39) = 1.36$, $p = .18$, and KF frequency (predictive $M = 5.28$, $SD = 8.35$; non-predictive $M = 5.98$, $SD = 7.52$), $t(39) = 0.66$, $p = .51$. Finally, with non-predictive verbs, good agents and patients were matched on both agenthood (agent $M = 5.09$, $SD = 1.32$; patient $M = 5.10$, $SD = 1.44$), $t(39) = 0.06$, $p = .95$, and patienthood (agent $M = 4.72$, $SD = 1.17$, patient $M = 4.60$, $SD = 1.35$), $t(39) = 0.55$, $p = .58$.

We assembled 80 additional distractor items, which were also animate humans and animals (e.g., *surfer* and *gardener*). Our distractors were chosen such that they were plausible but not particularly likely agents or patients of the predictive verbs (compared to the predictive verbs' "good" agents and patients), and such that they were equally plausible agents or patients of non-predictive verbs. Our non-predictive condition used verbs that might easily apply to all items in the display (i.e., target patients, competitor agents, or distractors).

The predictive verbs and their patients were fashioned into forty predictive sentences (e.g., *Toby arrests the crook*). *Toby* was the subject of each sentence. The good patients filled their typical thematic roles as direct objects of the verb. The good agents were never mentioned. A non-predictive verb was substituted into each sentence to create forty non-predictive sentences (e.g., *Toby notices the crook*). We recorded each predictive/non-predictive sentence pair using Praat software with a sampling rate of 44.1 kHz and 16-bit resolution for use as our auditory stimuli. Our visual stimuli were color photographs of the various agents, patients, and distractors, set against a white background. The full item list is presented in [Appendix A](#).

2.1.2. Design

We used a 2×2 design with verb type (predictive or non-predictive) and competitor presence (present or absent) as factors. In predictive conditions, participants heard sentences that contained a verb that was thematically related to the critical item(s) in the display: thus, thematic information from the verb could be used to anticipate the post-verbal referent. In non-predictive conditions, the sentence contained a verb that was thematically unrelated to the critical item(s) in the display: all items in the display were equally probable as a post-verbal referent. In competitor present trials, the critical items in the visual display included the target patient and competitor agent of the predictive verb. In competitor absent trials, the target patient was the only critical item, and the competitor agent was absent. The competitor absent condition is analogous to the design of [Altmann and Kamide \(1999\)](#), while the competitor present condition allows us to test whether anticipation is directed exclusively to items appropriate for upcoming, unfilled roles.

To construct four counterbalanced lists, we divided the 40 predictive/non-predictive sentence pairs into four groups and rotated them through each of the four conditions in a Latin Square. Twenty predictive and twenty non-predictive sentences appeared on each list, equally divided between competitor present and absent conditions. A given predictive/non-predictive sentence pair appeared in only one condition on each list. Participants were

randomly assigned to a list, and the order of sentences was randomized for each participant.

Two visual displays were associated with each predictive/non-predictive sentence pair. These displays were presented on a computer monitor, and they included a photograph in each corner of the screen, and a recurring image of the character Toby as the central fixation point, which remained at the center of the display throughout every trial. The display for the competitor present condition included a photograph of the target patient and competitor agent of the verb in the predictive sentence, and two unrelated distractors (see Fig. 1). The display for the competitor absent condition included the target patient and three distractors, two of which were semantically related to one another, but not to the verb or to the target patient. Thus, the presence of a semantically related pair in the display could not be used as a cue in determining the target. The locations of images were randomized on every trial.

2.1.3. Procedure

Participants listened to the recorded sentences over headphones while they viewed the corresponding visual display. A 500 ms preview of the display preceded the presentation of each sentence. Participants were told that they would be hearing short sentences about a character named Toby. In order to make Toby a plausible and salient agent, we presented a photograph of Toby along with a description about him prior to the experiment, which described him as an adventurous fellow who participates in many kinds of activities. The description made it clear that the sentences participants would hear would *all* describe Toby's actions. Participants were instructed to use a computer mouse to click on the image of the person or animal that Toby had interacted with in each sentence (i.e., the target patient). The trial ended when they clicked on an image.

Participants completed five practice trials at the beginning of the experiment, during which they received feedback on their performance (no feedback was provided throughout the rest of the experiment). Practice trials used non-predictive verbs. Although there were no additional filler trials, the competitor absent conditions served as a baseline for replicating Altmann and Kamide (1999) (i.e., for predictive vs. non-predictive verbs), and for counterbalancing strategies related to the presence of related pairs of images in the display. An ASL R6 remote optics eye tracker with a head-tracking device recorded participants' eye movements (Applied Scientific Laboratories, MA, USA). Materials were presented using E-Prime (Version 1.0, Psychology Software Tools, Inc., Pittsburgh, PA.), and the full session lasted under 20 min.

2.2. Results

To more carefully balance the patienthood of target patients with the agenthood of competitor agents, we removed one patient with a patienthood in the good agent range (*deer*, patienthood = 4.63), and two patients with agenthoods in the good agent range (*suspect*, agenthood = 5.69; *rock star*, agenthood = 5.63), from our

Table 1

Mean (standard deviation) accuracies and reaction times in Experiment 1.

	Accuracy		Reaction time	
	Verb type		Verb type	
	Predictive	Non-predictive	Predictive	Non-predictive
Competitor present	.95 (.08)	.99 (.03)	2955 (306)	3180 (260)
Competitor absent	.99 (.03)	.99 (.03)	2949 (348)	3143 (358)

analyses. Accuracies and reaction times were submitted to an ANOVA with verb type (predictive and non-predictive) and competitor presence (present and absent) as factors. Means and standard deviations are reported in Table 1. The analysis of accuracies revealed a reliable main effect of verb type in the participants analysis, $F_1(1, 15) = 6.74$, $p < .05$, $\eta_p^2 = .310$, $F_2(1, 36) = 1.99$, $p = .17$, $\eta_p^2 = .052$, such that accuracies were higher with non-predictive verbs; a marginal main effect of competitor presence, $F_1(1, 15) = 3.12$, $p = .09$, $\eta_p^2 = .172$, $F_2(1, 36) = 2.94$, $p = .09$, $\eta_p^2 = .076$, such that accuracy was marginally higher with competitors absent; and a non-reliable interaction of verb type and competitor presence, $F_1(1, 15) = 2.48$, $p = .14$, $\eta_p^2 = .142$, $F_2(1, 36) = 1.99$, $p = .17$, $\eta_p^2 = .052$.

Inaccurate trials were removed from the analysis of reaction times, which revealed a reliable main effect of verb type, $F_1(1, 15) = 38.04$, $p < .001$, $\eta_p^2 = .717$, $F_2(1, 36) = 16.12$, $p < .001$, $\eta_p^2 = .309$, such that reaction times were faster with predictive verbs. Neither the main effect of competitor presence, nor the interaction of verb type and competitor presence was significant ($F_s < 1$).

The mean proportions of fixations to the target patient, competitor agent (when present), and distractors are plotted by condition in Fig. 2. Inaccurate trials were removed from eye movement analyses, and fixations were time locked to the direct object (target patient) noun onset in each trial. The plotted window extends from verb onset to 1000 ms following the onset of the direct object noun. The mean onset and offset of the verb and noun are indicated. Note that although *Toby* was the agent of each sentence, when we refer to the target patients and competitor agents in the display, we mean the good patients (the targets, such as *crook*) and good agents (the competitors, such as *policeman*) of the predictive verb (e.g., *arrest*). Fixations to the character *Toby* are not plotted in the figures and are not included in the analyses.

We defined an *anticipatory window* of interest (shaded gray in Fig. 2) that extended from the mean offset of the verb to 200 ms following the onset of the noun (duration $M = 383$ ms). The 200 ms buffer following noun onset was motivated both by the mean time required to plan and launch an eye movement, and the typical lag observed between eye movements and fine-grained phonetic detail in the speech stream (Alloppenna, Magnuson, & Tanenhaus, 1998)³. Thus, even within the 200 ms buffer, listeners' eye movements were based on bottom-up information about the verb, but not yet the direct object. Given, however, that

³ Note that 200 ms is not assumed to be the minimum time for planning and launching, but an estimate of the average lag. See Altmann and Kamide (2004) for distributions of launch times in the visual world paradigm.

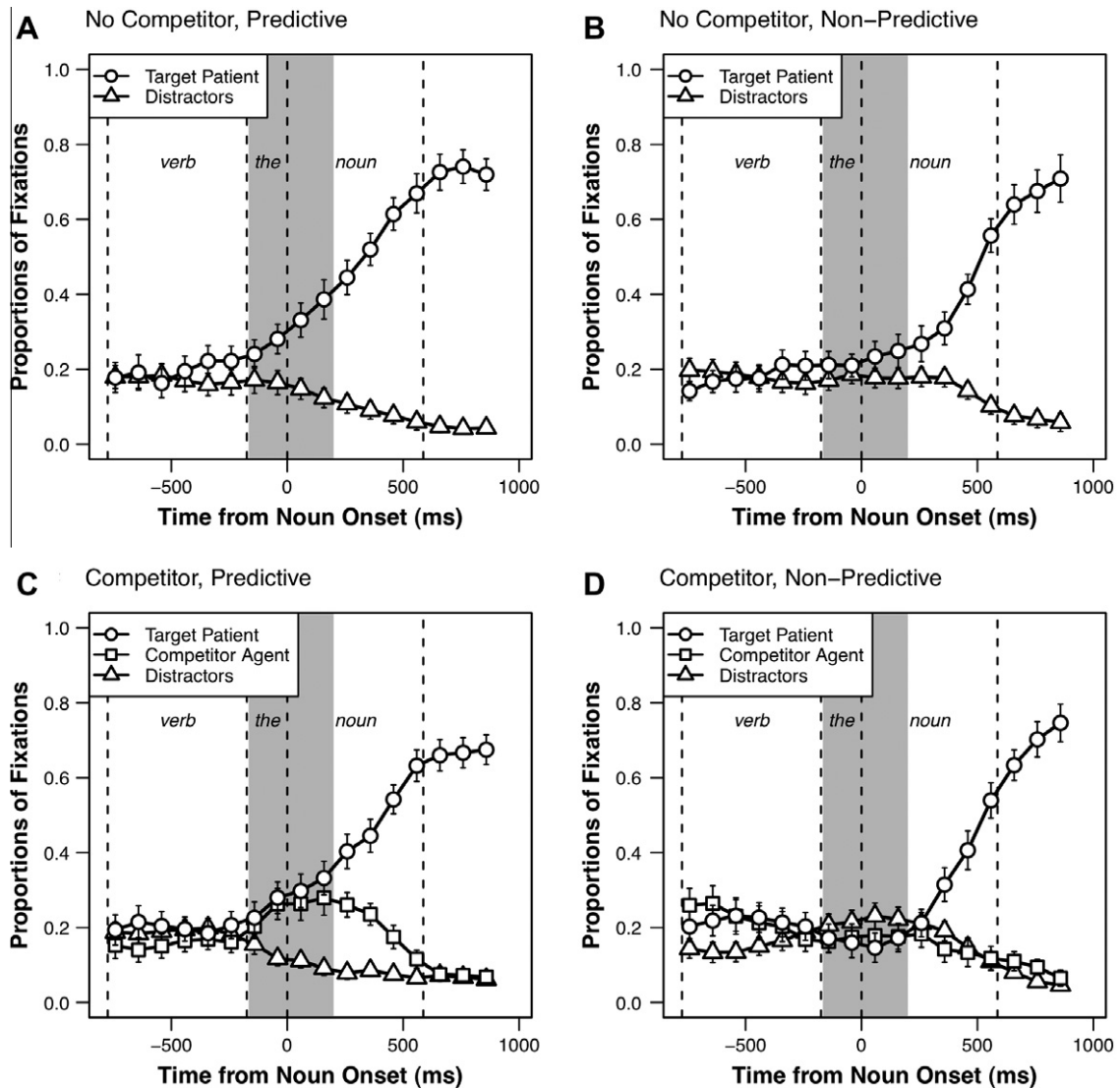


Fig. 2. Mean proportions of fixations (with SE bars) to the target patient (e.g., *crook*), competitor agent (e.g., *policeman*), and distractors in Experiment 1 by condition. The anticipatory window is shaded gray. Looks to Toby are not plotted.

Altmann and Kamide (2004) have demonstrated that saccades can be launched within 200 ms in similar but simpler tasks, it is likely that a small number of eye movements within our anticipatory window were launched in response to information about the named target patient. Additionally, some fixations within the anticipatory window could have begun prior to the naming of the verb, when no information about the target was available. However, because our analyses involve comparisons between predictive and non-predictive contexts, seepage of fixations into the anticipatory window, based on the naming of the target patient, or due to eye movements launched prior to the verb, should be equally likely (though rare) in both predictive and non-predictive conditions.

First, we consider the competitor absent conditions (Fig. 2A and B), which provide a replication of Altmann and Kamide (1999). Qualitatively, there is clear

anticipation of the target patient in the predictive condition relative to the distractors (see Fig. 2A), particularly in the shaded anticipatory window. However, because the target patients and distractors represent different items, there is a possibility that this difference is due to differences in salience, rather than linguistic predictiveness. Thus, as in Altmann and Kamide (1999), we test for anticipation by comparing fixation proportions to specific target patients with predictive vs. non-predictive verbs, as plotted in Fig. 3A.

Second, we consider the competitor present conditions (Fig. 2C and D), to test for effects related to the competitor agent. The crucial question is whether there is also anticipatory fixation to the competitor agent in the predictive condition (Fig. 2C). Qualitatively, there is a clear advantage for the competitor agent compared to the distractors in the predictive condition, which even extends beyond the

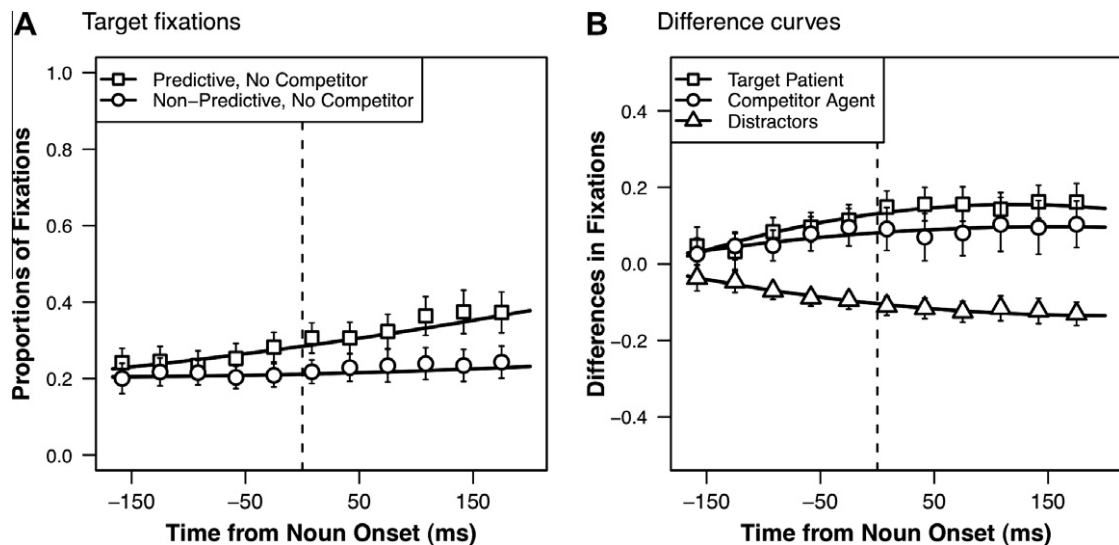


Fig. 3. Target patient (e.g., *crook*) fixations with predictive and non-predictive verbs in the competitor absent conditions (A), and difference curves reflecting predictive minus non-predictive proportions of fixations in the competitor present conditions for the target patient (e.g., *crook*), competitor agent (e.g., *policeman*), and distractors (B) in the anticipatory window in Experiment 1. The symbols (with SE bars) indicate mean trajectories, and the lines indicate growth curve fits.

anticipatory window; in fact, the probability of fixating the target patient and competitor agent is nearly identical throughout the anticipatory window. However, because the target patients, competitor agents, and distractors represent different items, we again cannot exclude the possibility that differences in saliency are influencing fixations. Additionally, we cannot simply compare fixation proportions in the predictive condition to the non-predictive condition for each item (as in the target analysis above) because we are also interested in comparing fixation patterns between items (e.g., target patient vs. competitor agent). Thus, to control for potential saliency effects, we computed difference trajectories for each item by subtracting fixation probabilities in the non-predictive condition from fixations in the predictive condition (thus “subtracting out” saliency differences which would also be present in the non-predictive condition). Difference trajectories for the three item types across the anticipatory window are shown in Fig. 3B.

To assess the time course of the fixation curve differences, we used growth curve analysis (GCA) as adopted for visual world data by Magnuson, Dixon, Tanenhaus, and Aslin (2007; see Mirman, Dixon, & Magnuson, 2008, for a detailed introduction). GCA provides a multilevel statistical approach that explicitly assesses change over time. Conceptually, a polynomial curve is fit at the level of each subject–condition combination, and then statistical tests are used to assess whether curve parameters differ reliably between conditions. The Mirman et al. approach affords dynamic consistency: the mean of the subject–condition fits is the same as the fit of the condition means due to the use of orthogonal power polynomials, which ensure that each polynomial term is independent. One convenient outcome of this approach is that the intercept is recentered in the analysis window, making it analogous to mean fixation proportion. The linear term corresponds to mean

linear slope over the analysis window, and the quadratic term indicates degree of curvature. While higher order terms can of course be applied, linking them transparently to cognition and perception is challenging (see Mirman et al. for discussion). In our case, the time course within the analysis window is quite simple, and the lower terms will easily suffice. Our growth curve models tested for intercept, linear, and quadratic effects of time on fixation proportions, using orthogonal power polynomials, with fixed effects of subject on each time term, and random effects of intercept and slope.

For the analysis of target fixation proportions in the competitor absent conditions (Fig. 3A), the GCA model included the critical fixed effect of verb type (predictive and non-predictive) on each time term. Curve fits are plotted as lines in Fig. 3A. The model revealed a reliable effect of verb type on the intercept, $estimate = -.08$, $t(15) = 2.70$, $p < .05$, capturing the clear mean difference, as well as a significant effect on the linear term, $estimate = -.18$, $t(670) = 3.01$, $p < .01$; the quadratic effect was not significant, $estimate = -.01$, $t(670) = 0.40$, $p = .69$.

For the analysis of the competitor present conditions (Fig. 3B), we directly compared target patient, competitor agent, and distractor difference curve trajectories in GCA. The predictions are not as simple as items that are anticipated should show positive differences and items that are not anticipated should show zero difference. The former is obvious, but an item that is not anticipated may actually show a *negative* difference. This is because fixation proportions are not independent. An increase in target fixations should entail a decrease in fixations to distractors. The crucial question then is whether the difference curve for competitor agents is closer to the difference curve for target patients (if anticipatory looking depends only on simple, local priming, difference curves should be approximately the same for targets and competitors) or

the difference curve for distractors (if anticipatory looking depends only on prediction of upcoming roles and is not influenced by local priming, difference curves should be approximately the same for distractors and competitors).

Qualitatively, while the pattern for competitors is quite similar to that for targets, it also shows a trend toward being intermediate between predictions for full dependence on local priming and full dependence on prediction. We ran separate GCM models on fixation difference curves, with the critical fixed effect of item on the time terms, for each of the three pairwise comparisons (target vs. competitor, target vs. distractor, and competitor vs. distractor), and we report only the critical effect of item on each time term. Curve fits for a single model with a 3-level item factor (target patient, competitor agent, and distractors) are plotted as lines in Fig. 3B. These fits are virtually indistinguishable from the pairwise curve fits, and thus illustrate the critical pattern of results (note that the 3-level model does not test all pairwise comparisons).

The first question is whether the target and competitor differences were reliably different from the distractor difference curve. The model comparing target and distractor difference curves revealed a reliable effect of item on the intercept, $estimate = -.22$, $t(15) = 6.13$, $p < .001$, capturing the clear mean difference, as well as significant effects on the linear term, $estimate = -.33$, $t(670) = 3.83$, $p < .001$, and the quadratic term, $estimate = .13$, $t(670) = 6.05$, $p < .001$, capturing the complementary slopes and curvatures for these difference curves. The results for the model comparing competitor and distractor difference curves were similar, with a reliable effect on the intercept, $estimate = -.18$, $t(15) = 4.41$, $p < .001$, and quadratic term, $estimate = .07$, $t(670) = 3.31$, $p < .001$, and a marginal effect on the linear term, $estimate = -.25$, $t(670) = 1.68$, $p = .09$.

The next question is whether the subtle differences between the target and competitor difference curves were reliable. The model comparing target and competitor difference curves revealed a reliable effect of item on the quadratic term, $estimate = .05$, $t(670) = 2.00$, $p < .05$, capturing the subtle differences in curvature between the conditions, but no effect on the intercept, $estimate = -.04$, $t(15) = 0.89$, $p = .39$, or the linear term, $estimate = -.08$, $t(670) = 0.68$, $p = .49$. The quadratic effect captures the sharper peak in the target patient difference curve, which is consistent with greater anticipation for target patients than competitor agents, even though the difference is not large enough to show up as a reliable difference in intercept (again, analogous to mean fixation probability in the entire window).

Finally, we considered the possibility that variation in visual similarity of target patients and competitor agents might be influencing our results. Some pairs (e.g., *police-man-crook*) have intuitively higher visual similarity than others (e.g., *mailman-kitten*). We split the items into two subsets according to a rough similarity threshold: if both were (individual) people or both were animals, they were considered high-similarity (14 pairs), while if one was a person but the other was either an animal, child, or group, they were considered low-similarity (23 pairs). We then repeated the growth curve analyses with the subsets. While there was a trend for a slightly stronger competitor effect (i.e., more similar target and competitor curves) for

high-similarity items, the low-similarity items showed the same competitor effects observed with all items: the competitor difference curve was intermediate between target and distractor curves, but closer to the target curve (with a reliable difference on the quadratic term). Thus, similar trends were present for both similar and dissimilar pairs, suggesting that while visual similarity may have had some influence, it was not driving the effects observed in Fig. 3B.

2.3. Discussion

In part, these results replicated those of Altmann and Kamide (1999): we observed reliably more target patient fixations in the anticipatory window with predictive verbs as compared to non-predictive verbs. However, post-verbal fixations were directed in large part to *both* items that were thematically related to the predictive verb (i.e., the target patient and competitor agent) until the direct object was named (see Fig. 2C), even though *Toby* filled the subject role in each sentence and made the competitor agent unlikely as the post-verbal argument. Although there was still evidence for a slight anticipatory advantage for targets over competitors, which showed up as a slightly sharper peak in the item-specific growth curve analysis (see Fig. 3B), there was clear evidence for local thematic priming, with similarly robust differences between distractors and target patients and competitor agents.

Together, these results suggest that in contexts in which potential referents and verbs are highly related, local thematic priming may be an important factor in driving anticipatory fixations, with the potential to strongly modulate the impact of sentence context. However, given that context effects depend on strength of context and the amount of time over which the context might have an effect (e.g., Shillcock & Bard, 1993; Swinney, 1979), we increased both in Experiment 2.

3. Experiment 2

Experiment 2 probed the somewhat surprising implication of Experiment 1 that local thematic priming can play a highly robust, and potentially larger role in anticipation than sentence-level syntactic constraints (since the competitor agent difference curve, although intermediate, was closer to the target patient curve than the distractor curve) when nouns and verbs are strongly related thematically. In Experiment 2, we used passive sentences (*Toby was arrested by the policeman*), in which the agents of predictive verbs were the post-verbal nouns: therefore, we now refer to target *agents* and competitor *patients* of predictive verbs in the display. We hypothesized that passive sentences might be more likely to reveal sentence-level influences on anticipatory fixations for two reasons. First, they include additional syntactic information (*was... and by...)* that further specifies an upcoming role. Second, the word *by* also extends the anticipatory time window, allowing more time for role anticipation to have a detectable impact. Additionally, we fully balanced the agenthood of target agents with the patienthood of competitor

patients, which was slightly biased in favor of the sentence-inappropriate competitor agents in Experiment 1.

3.1. Methods

3.1.1. Participants

Sixteen students from the University of Connecticut participated for partial course credit. All participants were native speakers of English with normal vision. None had participated in Experiment 1 or the norming.

3.1.2. Materials

We constructed 24 predictive and non-predictive sentences with the materials normed for Experiment 1. These sentences used a passive frame (*Toby was arrested/noticed by the policeman*), with the agents of the predictive verbs (*policeman*) always performing the action as the post-verbal noun. Actions were always directed at the character *Toby*, and the patients were never mentioned. Some items from Experiment 1 were eliminated because they lacked pragmatic sensibility in this frame (e.g., *Toby was milked by the farmer*). Among the remaining items, the mean agenthood of agents ($M = 6.57$, $SD = 0.49$) was higher than that of patients ($M = 3.23$, $SD = 1.06$; $t(23) = 13.46$, $p < .001$), and the mean patienthood of agents ($M = 3.47$, $SD = 0.94$) was lower than that of patients ($M = 6.34$, $SD = 0.57$; $t(23) = 14.34$, $p < .001$). Difference scores between the agenthood of agents and the patienthood of patients did not differ reliably between predictive ($M = .23$, $SD = 0.67$) and non-predictive verbs ($M = .16$, $SD = 1.33$), $t(39) = 0.27$, $p = .79$. Agents and patients were also matched on number of phonemes (agent $M = 5.75$, $SD = 1.78$; patient $M = 5.79$, $SD = 2.15$), $t(23) = 0.09$, $p = .93$, and *KF* frequency (agent $M = 27.75$, $SD = 35.69$; patient $M = 31.92$, $SD = 52.26$), $t(23) = 0.33$, $p = .70$.

Predictive and non-predictive verbs were also balanced on number of phonemes (predictive $M = 6.67$, $SD = 1.79$; non-predictive $M = 6.54$, $SD = 2.15$), $t(23) = 0.22$, $p = .83$, although non-predictive verbs were reliably more frequent (*KF* predictive $M = 15.13$, $SD = 24.72$; *KF* non-predictive $M = 56.21$, $SD = 65.93$), $t(23) = 2.87$, $p < .01$. The frequency difference should not impact the hypothesis under test (note that the only effect would be to speed processing with non-predictive verbs, which is counter to the predictions). Finally, with non-predictive verbs, good agents and patients were matched on patienthood (agent $M = 4.96$, $SD = 1.29$, patient $M = 4.98$, $SD = 1.16$), $t(23) = 0.11$, $p = .92$, although the agenthood of good patients was marginally higher than that of good agents (agent $M = 5.14$, $SD = 1.47$; patient $M = 5.74$, $SD = 0.82$), $t(23) = 2.05$, $p = .05$.

3.1.3. Design

As in Experiment 1, we crossed verb type (predictive or non-predictive) and competitor presence (present or absent) in a 2×2 design. Visual arrays were adjusted so that the target agent of each sentence was always present, while the competitor patient only appeared in the competitor present conditions. Items were counterbalanced across four lists, with 24 items per list.

Table 2

Mean (standard deviation) accuracies and reaction times in Experiment 2.

	Accuracy		Reaction time	
	Verb type		Verb type	
	Predictive	Non-predictive	Predictive	Non-predictive
Competitor present	1.00 (.00)	1.00 (.00)	3171 (376)	3456 (384)
Competitor absent	1.00 (.00)	1.00 (.00)	3079 (408)	3350 (323)

3.1.4. Procedure

The procedure was identical to Experiment 1, except that participants heard a passive sentence with a good agent as the target. The description of the character *Toby* was also modified for the passive sentences, making it clear that each sentence would be about something that happened to him, and that the task was to click on the picture of the entity that did something to Toby.

3.2. Results and discussion

An editing error left an extended period of silence at the beginning of one audio file, and this item (*bounty hunter*) was removed from the analysis. As in Experiment 1, we removed the three outlier items from the analyses that were outside the typical agent or patient range. Accuracies were perfect (100%) across all four conditions, so we did not analyze them further. Reaction times (see Table 2) were submitted to an ANOVA with verb type and competitor presence as factors. The analysis revealed a reliable main effect of verb type, $F_1(1, 15) = 33.41$, $p < .001$, $\eta_p^2 = .690$; $F_2(1, 19) = 22.48$, $p < .001$, $\eta_p^2 = .542$, such that reaction times were faster with predictive verbs; a reliable main effect of competitor presence in the items analysis, $F_1(1, 15) = 6.33$, $p < .05$, $\eta_p^2 = .297$, $F_2(1, 19) = 5.51$, $p < .05$, $\eta_p^2 = .225$, such that reaction times were faster with competitors absent; and a non-significant interaction of verb type and competitor presence (both $F_s < 1$).

The mean proportions of fixations to target agent, competitor patient (when present), and distractors are plotted by condition in Fig. 4, with the same anticipatory window described for Experiment 1 (mean verb offset to 200 ms following target agent onset; duration $M = 583$ ms) shaded gray.

Qualitatively, there are clear effects of predictiveness in the competitor absent conditions (Fig. 4A and B), with more fixations to target agents than distractors in the anticipatory window (see Fig. 4A), replicating *Altmann and Kamide (1999)*. As in Experiment 1, we used GCA to compare target agent proportions of fixations in the anticipatory window with predictive vs. non-predictive verbs, as plotted in Fig. 5A. The model (see Fig. 5A for curve fits) revealed a reliable effect of verb type on the intercept, *estimate* = $-.21$, $t(15) = 4.69$, $p < .001$, capturing the clear mean difference, as well as a significant effect on the linear term, *estimate* = $-.36$, $t(1054) = 2.60$, $p < .01$, and quadratic term, *estimate* = $.11$, $t(1054) = 3.64$, $p < .001$.

Again qualitatively, clear effects of predictiveness are also apparent in the competitor present conditions

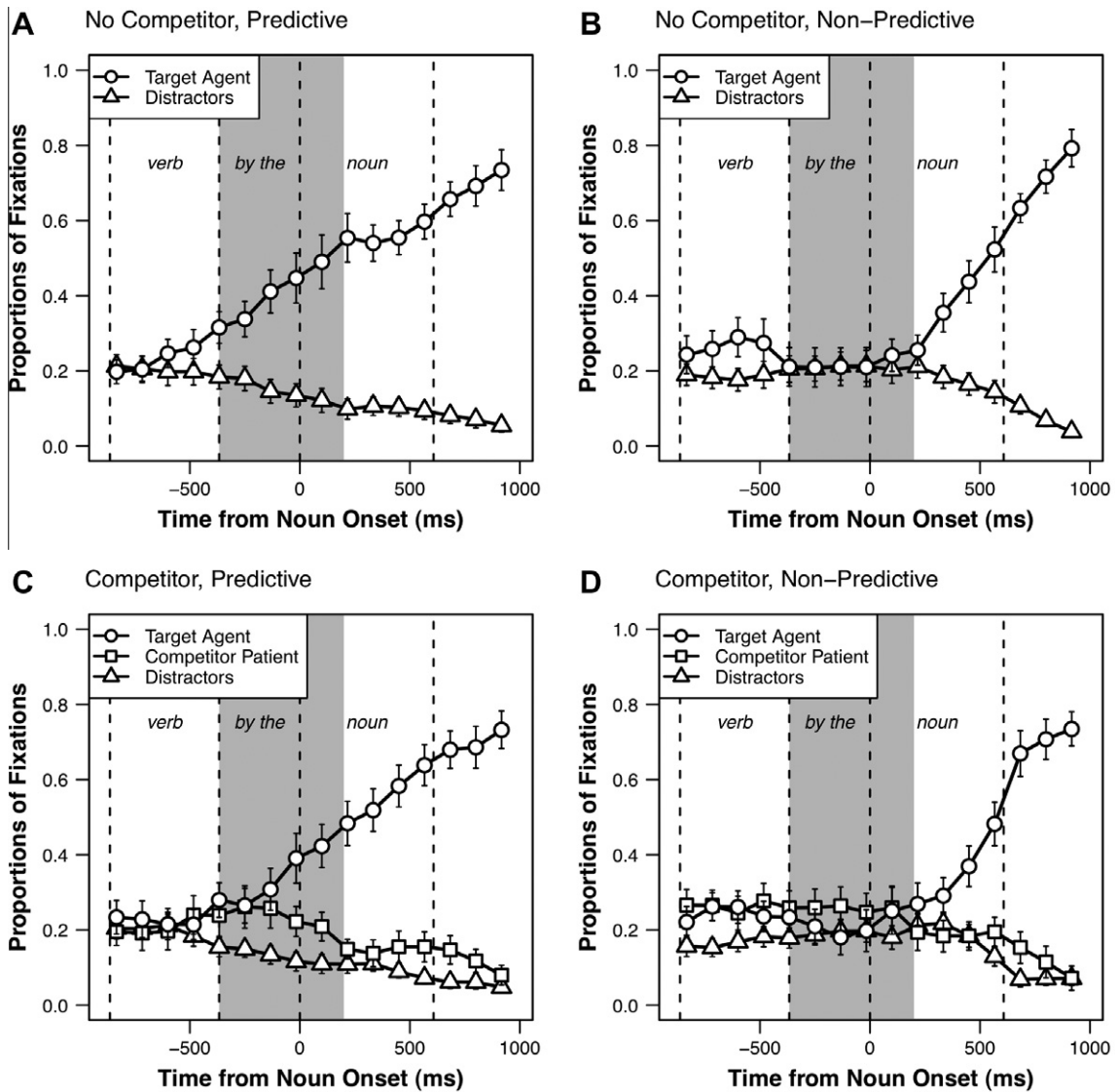


Fig. 4. Mean proportions of fixations (with SE bars) to the target agent (e.g., *policeman*), competitor patient (e.g., *crook*), and distractors in Experiment 2 by condition. The anticipatory window is shaded gray. Looks to Toby are not plotted. Note that the good agents were the targets of the passive sentences.

(Fig. 4C and D). In the early portion of the anticipatory window for the predictive condition (see Fig. 4C), there appears to be no difference between fixation probabilities for target agents and competitor patients, but a large target advantage quickly emerges. As in Experiment 1, however, we use growth curve analysis with item-specific difference curves (mitigating potential salience differences), which are plotted in Fig. 5B, to assess anticipation quantitatively.

Curve fits for a single model with a 3-level item factor (target agent, competitor patient, and distractors) are plotted in Fig. 5B, which are virtually indistinguishable from the patterns observed in the pairwise models described below. The model comparing target and distractor difference curves revealed a reliable effect of item on the intercept, $estimate = -.17$, $t(15) = 2.73$, $p < .05$, linear, $estimate = -.58$, $t(1054) = 3.03$, $p < .01$, and quadratic terms, $estimate = .11$, $t(1054) = 3.44$, $p < .001$. The model comparing competitor

and distractor difference curves revealed a reliable effect on the quadratic term, $estimate = .09$, $t(1054) = 3.08$, $p < .01$, reflecting opposite curvatures in the two difference curves, and non-significant curvatures on the intercept, $estimate = -0.04$, $t(15) = 1.33$, $p = .20$, and linear term, $estimate = -.20$, $t(1054) = 0.99$, $p = .32$. Thus, reliable advantages were found for the target agents and competitor patients relative to distractors, even in the item-wise analysis. The model comparing target and competitor difference curves revealed a marginal effect of item on the intercept, $estimate = -.12$, $t(15) = 2.06$, $p = .06$, a reliable effect on the linear term, $estimate = -.38$, $t(1054) = 2.73$, $p < .01$, and an unreliable effect on the quadratic term, $estimate = .02$, $t(1054) = 0.66$, $p = .51$. The intercept effect captures the larger mean difference for target agents apparent in Fig. 5B, and the linear term captures the clear differences in slope.

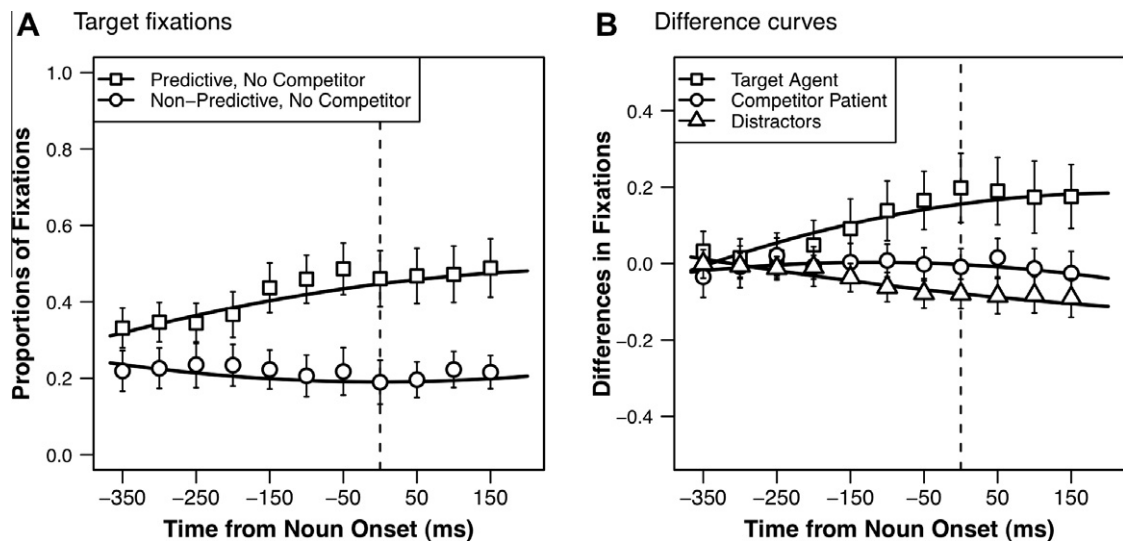


Fig. 5. Target agent (e.g., *policeman*) fixations with predictive and non-predictive verbs in the competitor absent conditions (A), and difference curves reflecting predictive minus non-predictive proportions of fixations in the competitor present conditions for the target agent (e.g., *policeman*), competitor patient (e.g., *crook*), and distractors (B) in the anticipatory window in Experiment 2. The symbols (with SE bars) indicate mean trajectories, and the lines indicate growth curve fits.

As in Experiment 1, we also performed subset analyses on the relatively more visually similar target-competitor pairs (those that were both animals or both people; 12 pairs) and the relatively dissimilar pairs (those where one item was a person and the other an animal, child, or group; eight pairs) as a check for effects of visual similarity. Given the small number of items, the difference curves tended to be noisier. As in the subset analysis for Experiment 1, there was a trend towards a stronger competitor effect (i.e., more similar target and competitor curves) with high-similarity items. However, even for the dissimilar pairs, the difference curves showed similar trends as in Fig. 5B: the competitor difference curve was intermediate between target and distractor curves, but closer to the distractor curve. Thus, the same trends seen in Fig. 5B remained even for the relatively dissimilar pairs.

Note, however, that the anticipatory window in Experiment 2 was longer than that in Experiment 1 by 200 ms, although it was defined in the same way (verb offset to 200 ms beyond noun onset). In Experiment 2, more words intervened in this interval, making the window longer. However, shortening the analysis window to parallel the anticipatory window in Experiment 1 (by eliminating the last 200 ms) does not materially change the outcome of the growth curve analysis. We attribute the reduced competitor “anticipation” compared to Experiment 1 to the fact that the passive sentence frames afforded greater constraint in favor of targets, due both to the greater syntactic cues (the modal, *was*, and *by*, although note that these items did not provide sufficient constraint to immediately wipe out local thematic priming effects) and greater time for constraints to have impact.

An alternative explanation for the post-verbal fixations to the competitor images in Experiments 1 and 2 is that participants may have been anticipating *other* sentence structures in which the competitor would be a predictable

post-verbal noun (e.g., *Toby arrested, with the policeman, the crook*; or, *Toby was arrested with the crook*). However, given the repetitiveness of the simpler active and passive sentence structures in Experiments 1 and 2, respectively, we find this explanation to be unlikely. Furthermore, these alternative sentence structures predict greater fixations to competitors than targets, which was not observed.

Finally, note that the longer passive sentences in Experiment 2 entail a potential confound regarding the display time of the visual scene. Specifically, through the onset of the target noun, listeners viewed the display for 1696 ms in Experiment 1, and 1937 ms in Experiment 2. The longer viewing times in Experiment 2 could potentially have allowed listeners to more fully settle on or access semantic representations for the items in the display. However, if greater preview time affords easier mapping of speech to visual objects, we should observe earlier divergence of targets from competitors and distractors in the non-predictive conditions in Experiment 2, as compared to Experiment 1. If one examines the early time course in the non-predictive conditions in the two experiments, however, it is clear that there is no such advantage in Experiment 2.

4. General discussion

There is considerable evidence from experiments using the visual world paradigm that listeners systematically fixate particular kinds of visual objects or images whose phonological forms have not occurred in the linguistic input. For example, listeners look to a *lock* when they hear *key* (Huettig & Altmann, 2005; Yee & Sedivy, 2006), they look to a *key* when they hear *logs* (via phonological mediation from *lock*; Yee & Sedivy, 2006), they look to a *cake* when they hear *eat* (Altmann & Kamide, 1999), they look to a

rope when they hear *snake* (i.e., to objects that share visual features with the target) (Dahan & Tanenhaus, 2005), and they look to a *typewriter* when they hear *piano* (i.e., to objects that are manipulated in ways similar to a target) (Myung, Blumstein, & Sedivy, 2006). In short, eye movements are readily directed toward things in the visual world that relate semantically, in a very general sense, with the words that we have heard, whether based on thematic role compatibility (Altmann & Kamide, 1999), visual features of objects (Dahan & Tanenhaus, 2005), depicted or implied event information (Altmann & Kamide, 2007; Knoeferle & Crocker, 2006, 2007), category (Huettig & Altmann, 2005), function (Yee & Sedivy, 2006), manipulation constraints (Myung et al., 2006), or affordances (Chambers, Tanenhaus, & Magnuson, 2004).

We have argued that findings of anticipation in many visual world studies do not unequivocally imply active forecasting of upcoming structure or elements based on the integration of global, sentence-level constraints. Rather, we suggested that anticipation might reflect, at least in part, local thematic priming independent of upcoming role expectations. Our experiments support this conjecture. In Experiments 1 and 2, we found considerable anticipatory post-verbal fixation (after a verb like *arrest*...) both to items appropriate for an upcoming, unfilled role (e.g., a typical thematic patient, such as *crook*, given *Toby arrests the...*) and to items that were typical for an already filled role (e.g., a typical thematic agent, such as *policeman*) that were matched in their relatedness to the verb.

To ensure that trends apparent in the mean data in each condition did not depend on variation in salience, we used item-specific difference analyses, which compared fixation proportions to specific items in predictive vs. non-predictive verb conditions. The predictive minus non-predictive difference was positive for targets and negative for distractors in the “anticipatory window” extending from verb offset to 200 ms after noun onset. In Experiment 1 (*Toby arrests the...*), the advantage of target patients over competitor agents was slight (see Fig. 3B). This suggests that under the specific conditions of this experiment (with competitors and targets with very strong associations with verbs), local priming may have been the primary driver of anticipatory fixations. In Experiment 2, where we expected passive sentence frames (*Toby was arrested by the...*) to potentially boost thematic expectations due to additional syntactic cues (function words *was* and *by*), and due to simply greater time for constraint impact, we again found a pattern consistent both with predictive forecasting (a stronger target advantage compared to competitors) and local thematic priming (an advantage for competitors over distractors, as quantified by comparing fixation proportions to both items types in predictive vs. non-predictive verb conditions; see Fig. 5B).

4.1. Relation to prior work

Our results extend a number of findings concerned with effects of global, sentence-level and local, word-level linguistic constraints on sentence processing, using both the visual world paradigm, as well as other methods. In part, though, our results appear to conflict with the fourth

experiment reported by Ferretti et al. (2001). They asked participants to name visual words displayed after auditory sentence fragments (e.g., *She arrested the...*; or, *She was arrested by the...*). In earlier experiments, Ferretti et al. found that single word presentation of verbs led to reliable priming of good agents, patients and locations. But given a sentence onset with the agent role filled (active) or patient role filled (passive), naming was facilitated only when the displayed word was appropriate for the unfilled role (i.e., an appropriate patient given an active onset, or an appropriate agent given a passive onset) compared to sentences where the verb was thematically unrelated to the agents and patients (e.g., *kissed*). Certainly, it seems a logical extension from this result to predict anticipation of the unfilled role, and no anticipation of the filled role, in Experiments 1 and 2.

However, there is a noteworthy wrinkle to the Ferretti et al. results: there appeared to be no penalty when a thematically inappropriate item was presented (e.g., *She arrested the – COP* was no slower than *She kissed the – COP*). If perceivers were actively anticipating likely patients, *cop* ought to be more unexpected following *She arrested the...* than *She kissed the...* (the relative likelihood of cops being arrested in different parts of the world notwithstanding; these are just example items). On the other hand, one might argue that our results predict facilitation for *cop* in the patient role (i.e., facilitation for *She arrested the – COP*, consistent with Experiment 1), from local thematic priming. Our interpretation is that in fact both occurred in the Ferretti et al. experiment: *cop* was unexpected, but the penalty for an unexpected patient was mitigated by local thematic priming (leading to the prediction that a penalty should be observed for an unexpected patient that is also thematically unrelated to *arrest*, like *saint* or *puppy*).

Our findings are complementary to those of Knoeferle and Crocker (2006, 2007). They found that depicted event information could override thematic knowledge; we found that *sufficiently strong* thematic connections can have early, strong effects on activation even when they are inconsistent with sentence level thematic expectations. How pervasive such effects are would depend on the relative weight of local (lexical–lexical) and global (sentential and above) constraints; effects would be weaker given verbs with weaker thematic role predictability or weaker semantic association with agents and patients (e.g., *put*, as in Chambers et al., 2004, or *return*, as in Chambers & San Juan, 2008). Let us emphasize again that we are not excluding a role for prediction based upon context-specific evaluation of verb and object properties. As Chambers et al. put it, highly specific and possibly never experienced non-linguistic scene and event features (e.g., the goal of moving objects with an instrument with a hook at the end) “contour” the domain of reference. Our results suggest that some of this “contouring” could emerge from simple mechanisms based on priming principles.

Taken together, it may be helpful to think of local priming influences on anticipatory eye movements along a continuum. At one end, we have findings of anticipatory fixations with verbs that have no thematic connection with the referents of interest, such as the Chambers and San Juan (2008) finding that participants anticipate reference

to a previously moved object on hearing *return* (several other studies also fall at this end of the continuum, e.g., Altmann & Kamide, 2007; Knoeferle & Crocker, 2006, 2007; Magnuson, Tanenhaus, & Aslin, 2008). Such findings require active forecasting and cannot be accounted for by priming. At the other end of the continuum, we have fixations to items in a display that are semantically related to a named target, but that have no other connection to the unfolding language or the task at hand (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2005; Myung et al., 2006; Yee & Sedivy, 2006). Such findings seem to require priming between individual concepts. Other instances of anticipatory fixations lie between these two extremes. Our own findings are perhaps intermediate between the extremes, but nearer the priming endpoint (because we used especially strong verb–referent relationships to maximize the possibility of detecting such effects), whereas the results of Kamide, Scheepers, et al. (2003) are further toward the active prediction end (with weaker verb–referent relationships, and perhaps stronger global context). Such graded influences of local, lexically-based information are generally consistent with a constraint-based approach.

4.2. Theoretical considerations

Experiments 1 and 2, and the various studies described above, support a rich interaction between local, word-level and global, sentence-level constraints, consistent with a number of theoretical proposals in the literature. This interpretation owes much to the theoretical perspective and empirical contributions of McRae and colleagues. While their studies have revealed that accessing a word triggers activation of rich thematic and event knowledge (Ferretti et al., 2001; McRae et al., 2005), they have also shown that there is a subtle and complex interplay between such priming, and mechanisms of ongoing sentence processing that they explain via constraint-based theory. For example, the sense of a verb (e.g., “let in” vs. “acknowledge” senses of “admit”) indicated by a sentence context in turn affects parsing preferences (Hare, McRae, & Elman, 2003, 2004), and verb aspect (“was skating” vs. “had skated”) modulates which aspects of event knowledge are activated by a word (Ferretti, Kutas, & McRae, 2007). On their view, these rich, iterative interactions among lexical knowledge and sentence processing afford reliable “expectancy generation” that facilitates fast, efficient language comprehension.

At first glance, our results may appear consistent with the classical view of context integration proposed by Swinney (1979) and Tanenhaus et al. (1979): initially, all forms matching the bottom-up input are activated (even those that have a strong association with an uttered verb for an already filled role), and context is applied at a later stage to constrain interpretation. However, the original interpretation of these results no longer holds; Shillcock and Bard (1993) demonstrated that immediate context-selective priming is obtained when constraints are substantially stronger than those used in the 1979 studies (see also Magnuson et al., 2008). Furthermore, since we are considering *anticipatory* fixations, we are looking at cases where effects *precede bottom-up specification* of the

direct object. We suggest that the temporal lag between bottom-up input and clear impact of thematic context (i.e., the brief period apparent in both experiments when verb–noun associations appear to at least partly drive fixation proportions before thematic expectations dominate) does not indicate separate access and (later) selection processes. Instead, our view is that we are observing the impact of constraints for which the relevant window of analysis or grain size varies in scale. The impact of lexical priming is observable even as a word is unfolding (Huettig & Altmann, 2005; Yee & Sedivy, 2006). A conjoint constraint – the combination of the fact that, e.g., two roles are required, and the fact that one role is already filled – will emerge over a larger (e.g., phrasal) window. This may make it easy for such constraints to be masked by constraints that emerge over a more local window, such as very strong lexical and/or thematic priming. In general, for effects to be detectable immediately, constraints must be overwhelmingly strong, whether they are based on lexical priming, or linguistic (Dahan & Tanenhaus, 2004), scene (Altmann & Kamide, 1999), or pragmatic (Magnuson et al., 2008) context.

This view is very much like that proposed by Kuperberg (2007) in considering the possibility that the N400 and P600 ERP components may reflect the operation of multiple processing streams that operate largely in parallel but with constant potential for interaction. She proposes that one stream “...is a semantic memory-based system that constantly compares lexical associative and categorical relationships between incoming groups of words with pre-existing information stored within semantic memory” (p. 44). This stream would be consistent with priming-driven anticipation, and activation of items that are not consistent with forward-directed prediction, as we have found in the current studies – with modification to allow for greater event-related/thematic knowledge to be associated with this stream. The second stream is “combinatorial” and “...is sensitive to morphosyntactic as well as to thematic-semantic constraints” (p. 44), and thus is similar to proposals for forward-directed prediction in the visual world paradigm (e.g., Altmann & Kamide, 2007; Altmann & Mirkovic, 2009; Knoeferle & Crocker, 2006, 2007).

What sorts of computational mechanisms might support the different components of Kuperberg’s dual stream theory? For the priming-based component, consider first a well-known example from spoken word recognition where anticipation emerges from unordered, parallel constraint interaction: *recognition point* effects (i.e., high probability of word “recognition” prior to word offset). In the original Cohort model (Marslen-Wilson & Welsh, 1978), such effects follow from “active,” explicit anticipation based on tracking the set of possible lexical completions as a word is heard relative to an explicitly identified word onset. However, a model like TRACE (McClelland & Elman, 1986) predicts such effects without explicit tracking of word onsets or offsets, or any explicit representation of a discrete set of possible completions. Instead, recognition point (and related) effects emerge based on the dynamics of lexical activation and competition. Promiscuous looking may reflect an underlying system in which multiple interpretations are kept active (or generated) based on information accruing on a large (sentence or discourse) temporal

scale, but in which strong local constraints can conspire to *additionally* activate representations that are incompatible with that larger context. This would be analogous to fixations to rhymes in spoken word recognition studies (e.g., looking at *speaker* given the unambiguous input *beaker*; Allopenna et al., 1998). Rhyme effects in TRACE result from activation based on local overlap, and *despite* clear mismatch with earlier information – that is, with larger (whole word) scale context. Thus, the lexical-semantic (plus thematic/event/situational knowledge) memory stream might be modeled with an interactive activation model akin to TRACE (and augmented with semantic representations), or in the attractor network framework of Cree, McRae and colleagues (e.g., Cree & McRae, 2003; Cree, McRae, & McNorgan, 1999; O'Connor, McRae, & Cree, 2006), or a related attractor framework tailored to sequential processing of spoken input (Plaut & Kello, 1999).

What about the forward-directed component of Kuperberg's dual stream theory? Several possibilities have been proposed in the literature. McRae, Spivey-Knowlton, and Tanenhaus (1998) modeled related effects using *normalized recurrence*, in which constraints most naturally enter into competition simultaneously (rather than separate staging of local and global constraints). Differences in the time course of lexical and sentential constraints might be modeled in the recurrence normalization framework by differential weighting, distance from input or integration nodes (which would allow for differences in the temporal impact of constraints), or explicit delays. Functionally, this approach would have much in common with the *Sentence Gestalt Model* (McClelland, St. John, & Taraban, 1989; St. John & McClelland, 1990), since both have been construed as models of role filling (though neither framework is limited to such scope), and both have logically assumed forward directionality. McRae et al. (1997) also suggest that the distributed memory model of McClelland and Rumelhart (1985) might provide a good framework for priming based on thematic fit. In this framework, representations are distributed featurally in a matrix memory. Given features appropriate for thematic roles and other expectations, a stimulus such as *arrest* would lead the system to compute representations corresponding to likely agents and patients. It is not immediately clear, however, how one would adapt the framework to the task of *successively* activating appropriate lexical items in the context of full sentences. Indeed, it not apparent that any of these mechanisms could be easily integrated with a full parsing model (and of course, none is meant to; each models specific phenomena at a scale smaller than whole sentences).

One possible alternative to Kuperberg's dual stream framework comes from "satisficing" (good enough, non-maximizing; Simon, 1956) parsing theories; the "good enough representations" of Ferreira and colleagues (e.g., Ferreira, Ferraro, & Bailey, 2002), and the SOPARSE self-organizing sentence processing model of Tabor and colleagues (Tabor & Hutchins, 2004; Tabor et al., 2004). Ferreira's theory is explicitly satisficing in that it does not assume that listeners or readers maintain a complete parse of the unfolding sentence in every instance. Especially when there are many words or complex structure, listeners seem to represent just enough to support maintenance of

the likely gist of the sentence. For example, participants often simultaneously interpreted the garden path sentence, *While Anna dressed the baby played in the crib*, as indicating both that Anna dressed the baby, and that the baby played. While Ferreira's theory does not include an implemented model, it is similar in spirit to SOPARSE. In SOPARSE, local information can compete quite powerfully with global (longer time scale) information related via a set of linked treelets activated by the input thus far. For example, given a phrase with high local coherence for an active interpretation, such as *the player tossed a frisbee*, within a larger structure that dictates that the phrase is actually a relative clause (e.g., *The coach smiled at the player tossed a frisbee*), Tabor and colleagues have shown that the locally coherent, active reading competes with the global, reduced relative reading. Thus, this model affords necessary aspects of a forward-directed predictor, but also aspects of semantic-memory-like activation driven at a more local (word-level) grain.

SOPARSE is motivated by assumptions of self-organization, and phenomena like attraction to "merely local" coherence (i.e., activation of structures that are incompatible with available context) emerge naturally in such a framework; so might the sorts of results we have reported here. One possibly fruitful question for future research is whether the SOPARSE architecture might provide a unifying framework encompassing both hypothesized streams (Kuperberg, 2007). Indeed, similar phenomena might emerge from a self-organizing system in any domain with similar informational characteristics.

But if this were so, why should it be a general property of self-organizing systems that they activate or prime representation or structure that is incompatible with predictions afforded by, for example, a word onset (in the case of rhyme activation), a role clearly having been filled (as in priming in the current studies), or global structure (as in cases of merely local coherence)? On some approaches to language processing, such activation might be viewed as wasteful, sub-optimal, or irrational (e.g., Hale, 2001; Levy, 2008). Would it not be more rational for the processor to maximize prediction accuracy by fully exploiting predictive information? Here, we appeal to the distinction between *exploitation* (maximization) vs. *exploration* (Movellan & McClelland, 2001). In many situations, humans and other species probability match rather than maximize (i.e., they select among choices probabilistically, e.g., proportionally to prior probability, rather than choosing the most likely alternative). As Movellan and McClelland discuss, maximizing is not optimal when the environment is non-stationary, that is, when prior probabilities change over time (as is the case in nearly all natural contexts). As long as an organism maximizes based on subjective probabilities estimated at time t , it loses the opportunity to update its subjective probabilities if they change after time t .

Kamide, Altmann, et al. (2003) suggest one more alternative: *simple recurrent networks* (SRNs; Elman, 1990). Such models operate by explicitly predicting upcoming information. Properly trained (with appropriate parameters and training corpora), SRNs show graded sensitivity to constraints at multiple time scales. SRNs in some ways provide

a simpler alternative to SOPARSE, in that one could easily build a system that would predict sentential-level phenomena from sublexical input. Altmann and Mirkovic (2009) have elaborated on this proposal, suggesting that linked SRNs with independent input and output layers (linguistic vs. non-linguistic, e.g., scene information) but shared hidden units (see Dienes, Altmann, & Gao, 1999) could learn linguistic and non-linguistic contingencies and provide a forward predictive model without explicit mechanisms for evaluating linguistic representations. Instead, consistent with Altmann and Kamide (2007, 2009), lexical, syntactic, thematic and other aspects of linguistic knowledge would constitute features at varying time scales, that would be simultaneously embedded statistically within the patterns of connectivity within the network without explicit levels of representation (see also Elman, 2004, 2009).⁴

Such a system might provide a way to account for forward and local bases for anticipation without postulating two mechanisms. Spivey (2007) visualizes the state of a recurrent network over time as points in a many dimensional space (e.g., one per node in the network). When visualized metaphorically in three dimensions, and points visited just prior to a word and just after it are marked, the result is a manifold shaped like a “wasp-waisted tube” spreading backwards in time to encompass points traversed in “contextual pasts” (along trajectories from prior words in utterances that led to the current word) and forward away from the current word, covering “anticipated futures” (e.g., p. 317). That is, there is a cloud of visited points preceding the attractor for a particular concept that funnels to the attractor region. These represent trajectories that have been visited as the system approaches the concept attractor. A similar cloud funnels out away from the attractor, representing trajectories that have followed a visit to the concept attractor. “Activation” of a representation from this perspective can be thought of as distance from the current state in multidimensional space; priming is a result of the priming word pulling the system to a state closer to the region associated with a probe. Thus, priming based on local thematic relatedness follows from proximity to trajectories in the contextual past. Forward facilitation for upcoming, unfilled roles follows from the same principle, but additional facilitation follows from the fact that system is following a *trajectory*. Thus, an advantage is predicted for forward prediction, although the current results suggest that whether and when such an advantage is observed depends on the relative strength of thematic relations and the probability of an upcoming role (among other constraints).

⁴ Concern must be taken in training SRNs, if they are to remain sensitive to local and larger time scales. When SRNs are over-trained on static corpora (that is, if they are effectively given stationary priors), they become less sensitive to local information and would cease activating representations inconsistent with preceding context (since they eventually acquire perfect priors for unrealistically long sequences, which would not be found in an unconstrained corpus; see Magnuson, Tanenhaus, Aslin, and Dahan (2003) and Magnuson, Tanenhaus, and Aslin (2000), for simulations showing that rhyme effects disappear from SRN models of spoken word recognition with sufficient, noise-free training).

5. Conclusions

Our findings are consistent with a constraint-based system in which the constraints allow partially inappropriate outcomes (words inappropriate for upcoming roles) to compete with more globally appropriate ones. The system may require two loosely interacting streams (lexical-semantic and morphosyntactic-thematic) as suggested by Kuperberg (2007; see also Kim & Osterhout, 2005), or it may be possible for both local priming and forward-directed prediction to emerge within a single self-organizing system, such as SOPARSE (Tabor & Hutchins, 2004; Tabor et al., 2004) or an SRN-based approach (e.g., Altmann & Mirkovic, 2009). There are potential advantages to these kinds of organization, where both local priming and forecasting conspire to afford powerful constraint-satisfaction. To the degree that some of the work can be done via “dumb” mechanisms like local priming, computational complexity is minimized. Potentially, such a system is maximally flexible: it is poised to “anticipate” any number of outcomes. Additionally, at the discourse level, it may prove especially useful to have the rich semantic content of a verb activated (typical arguments, etc.) even when that content may impede anticipation of upcoming items at the sentence level, as that content may facilitate understanding the intricate relations among referents in an unfolding discourse, and evoke connections between an ongoing discourse and related ideas and knowledge in memory.

Acknowledgements

We thank Ken McRae for sharing agent–patient–verb ratings with us. Preliminary results from Experiments 1 and 2 were presented at the 20th Annual CUNY Conference on Human Sentence Processing in San Diego, California (March, 2007). We thank Ken McRae, Mary Hare, and Gerry Altmann for helpful suggestions about that presentation. We also thank Ken McRae, Whit Tabor, Dan Mirman, Stephen Tobin, Craig Chambers, and Michael Spivey for comments and discussions that greatly improved this manuscript. We thank Navin Viswanathan for “playing” the role of Toby. We gratefully acknowledge support from NICHD Grants HD01994 and HD40353 to Haskins Laboratories, NIDCD Grant DC0055765 and NSF CAREER 0748684 to J.S.M., and NICHD Predoctoral NRSA HD060414 to A.K.

Appendix A. Verb-argument norming

For Experiments 1 and 2, we normed predictive verbs with good agents and patients. Additionally, we collected ratings on the thematic relationship between potential unrelated verbs (for use in non-predictive sentences) and the agents and patients of potential predictive verbs. Although photographs were not used with the norming, the data nevertheless constrained our choices for good agents and patients.

A.1. Methods

A.1.1. Participants

Twenty-eight native English-speaking undergraduates from the University of Connecticut participated for course credit. None had participated in Experiment 1.

A.1.2. Materials

We constructed a questionnaire that was modeled on McRae et al. (1997). As a measure of agenthood, we asked participants to judge how common it was for particular people or animals to perform the action described by a verb. As a measure of patienthood, we asked for judgments of how common it was for particular people or animals to be the recipient of the action described by a verb. We elicited ratings of both agenthood (e.g., *How common is it for a ___ to arrest someone/something?*) and patienthood (e.g., *How common is it for a ___ to be arrested by someone/something?*) for a given set of people and animals (e.g., *violinist, sheep, cashier, crook, cellist, truck driver*).

The stimuli included 45 verbs with potentially good agents and patients (predictive verb candidates). Some of our items were based on materials from McRae et al. (1997, 1998) and Binder, Duffy, and Rayner (2001). Each verb was grouped with seven animate nouns: some nouns were good agent or patient candidates, some were poor candidates, and some were included to fall in the middle of the typicality scale. These verbs were each paired with a non-predictive verb candidate, which was rated on the same noun set. We created eight lists, such that agenthood and patienthood were rated between-participants, and such that each list contained half of the predictive and non-predictive verb candidates. Lists were also structured

so that a participant did not see both members of a potential predictive/non-predictive verb pair. Ratings were on a seven-point scale, with a rating of 7 reflecting highly common on half of the lists, and highly uncommon on the remaining lists.

A.1.3. Procedure

The questionnaires were administered in small group settings using a paper and pencil format. Instructions were read to each group and printed on each questionnaire, and example ratings were provided. Each session lasted approximately 45 min.

A.2. Results and discussion

Among the 45 potential predictive verbs, 5 were rated as not having a good agent and/or patient, and they were removed from the analysis. Items and ratings are reported below. As in McRae et al. (1997), agents of predictive verbs were rated high on agenthood ($M = 6.64$, $range = 5.00–7.00$) but low on patienthood ($M = 3.04$, $range = 1.19–4.94$), with no overlap between agenthood and patienthood ratings. Patients of predictive verbs were rated low on agenthood ($M = 2.78$, $range = 1.06–5.69$) and high on patienthood ($M = 6.36$, $range = 4.63–7.00$). Each predictive verb was paired with one of 24 non-predictive verbs. The agent and patient of each predictive verb were also rated on their agenthood and patienthood with the non-predictive verb in each pair: the agenthood and patienthood ratings of agents (agenthood $M = 5.09$, $range = 1.46–6.94$; patienthood $M = 4.72$, $range = 2.19–6.77$) and patients (agenthood $M = 5.10$, $range = 1.31–6.88$; patienthood $M = 4.56$, $range = 1.50–6.75$) overlapped considerably.

Ratings of agenthood and patienthood for good agents and good patients of predictive verbs. Non-predictive verbs and distractors from Experiments 1 and 2 are also indicated.

Predictive verb	Non-predictive verb	Good agent		Good patient		Distractors		
		Agent-hood	Patient-hood	Agent-hood	Patient-hood			
Adopt	Touch	Couple	6.50	2.44	Puppy	2.31	6.13	Batter, clerk, pitcher
Arrest	Notice	Policeman	7.00	2.92	Crook	1.63	6.15	Cashier, cellist, violinist
Assist	Ignore	Meter maid	5.00	2.91	Elderly person	3.94	6.38	Journalist, photographer, rugby player
Astound	Annoy	Acrobat	6.44	4.13	Onlookers	3.13	6.19	Housekeeper, janitor, snake
Attack	Expect	Tiger	6.25	2.81	Mouse	2.38	5.31	Figure skater, hockey player, wizard
Beg	Notice	Bum	6.81	1.81	Tourist	4.00	5.19	Ant, grasshopper, punk rocker
Boil	Hide	Cook	7.00	1.19	Crab	1.19	6.63	Bride, groom, saxophone player
Captivate	Select	Magician	6.50	4.06	Crowd	4.19	6.56	Hula dancer, sailor, soldier
Capture	Telephone	Bounty hunter	6.69	2.47	Prisoner	2.63	6.13	Construction worker, contractor, nun

(continued on next page)

Appendix A (continued)

Predictive verb	Non-predictive verb	Good agent		Good patient		Distractors		
			Agent-hood	Patient-hood	Agent-hood		Patient-hood	
Carry	Touch	Mailman	7.00	2.88	Kitten	3.81	6.69	Astronaut, prince, queen
Carve	Annoy	Wood smith	6.94	1.80	Turkey	1.06	5.94	Businessman, skier, snowboarder
Catch	Bewilder	Fisherman	6.94	2.63	Salmon	3.38	6.56	Newscaster, weatherman, whale
Convict	Pester	Jury	6.88	1.75	Criminal	1.94	6.81	Bull, groomer, poodle
Cook	Hear	Chef	6.38	1.50	Lamb	1.79	6.00	Elf, fairy, guitarist
Cure	Telephone	Doctor	6.69	4.38	Patient	3.19	6.81	Banker, referee, weightlifter
Dazzle	Assess	Showgirl	6.88	4.56	Theatergoers	3.25	6.56	Eagle, sheriff, vulture
Defend	Hear	Lawyer	7.00	4.25	Suspect	5.69	6.44	Goat, sheep, woman
Deliver	Disturb	Pizza boy	7.00	2.47	Newborn	1.44	6.88	Goldfish, robin, woodpecker
Dissect	Prefer	Biologist	6.63	1.44	Frog	1.19	6.88	Kangaroo, koala, realtor
Excavate	Select	Archaeologist	6.31	2.81	Mummy	1.19	5.63	Parrot, salamander, toucan
Feed	Accompany	Farmhand	6.38	3.00	Llama	3.38	6.17	Nurse, preacher, surgeon
Frighten	Disregard	Monster	6.75	1.88	Toddler	3.38	6.69	Dolphin, knight, starfish
Idolize	Push	Teenager	6.88	4.31	Rock star	5.63	6.94	Beauty queen, rat, scientist
Imprison	Bewilder	Warden	6.19	3.20	Jailbird	2.56	6.73	Fencer, vampire, witch
Interrogate	Expect	Detective	6.69	4.06	Inmate	3.44	6.60	Badminton player, scorpion, squash player
Interview	Approach	Reporter	7.00	4.75	Soccer player	2.56	6.13	Bunny, pharmacist, squirrel
Lecture	Push	Coach	6.50	4.00	Boy	4.19	6.56	Electrician, mechanic, ninja
Milk	Watch	Farmer	6.94	1.19	Cow	1.75	6.94	Gorilla, monkey, pirate
Rescue	Awaken	Lifeguard	6.69	3.00	Hostage	2.06	6.75	Conductor, lizard, turtle
Ride	Educate	Cowboy	6.69	2.38	Horse	1.38	6.88	Ballerina, disco dancer, baton twirler
Scold	Hide	Principal	6.88	3.38	Delinquent	3.38	6.81	Clam, raccoon, shrimp
Scrub	Outwit	Maid	6.94	2.50	Elephant	2.56	5.06	Admiral, drummer, pianist
Serve	Approach	Waitress	7.00	3.75	Customer	2.44	6.63	Bat, gardener, owl
Shoot	Greet	Hunter	7.00	3.50	Duck	1.44	5.38	Cheerleader, football player, welder
Sketch	Describe	Artist	7.00	4.94	Model	3.06	6.38	Flutist, ice skater, trumpet player
Slaughter	Pester	Butcher	6.93	1.31	Pig	1.81	6.92	Fly, miner, moth
Stalk	Awaken	Lion	5.94	3.50	Deer	2.31	4.63	Graduate, sculptor, weaver
Summon	Watch	Monarch	5.56	3.93	Waiter	4.38	6.60	Cat, hippopotamus, zebra
Walk	Outwit	Trainer	6.31	3.06	Dog	3.63	7.00	Psychic, seagull, swan
Worship	Disregard	Minister	6.38	4.56	Goddess	2.50	6.67	Basketball player, tennis player, usher

References

- Allopena, P., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57, 502–518.
- Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264.
- Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition*, 111, 55–71.
- Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33, 583–609.
- Binder, K., Duffy, S., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language*, 44, 297–324.
- Chambers, C., & San Juan, V. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition*, 108, 26–50.
- Chambers, C., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 687–696.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 84–107.

- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.
- Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*, 371–414.
- Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin and Review*, *12*, 453–459.
- Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 498–513.
- DeLong, K., Urbach, T., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117–1121.
- Dienes, Z., Altmann, G. T. M., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science*, *23*, 53–82.
- Duffy, S., Morris, R., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*, 446–449.
- Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*, 1–36.
- Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science*, *7*, 301–306.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *25*, 348–368.
- Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.
- Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 182–196.
- Ferretti, T., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*, 516–547.
- Frazier, L., & Fodor, J. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.
- Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. In *Proceedings of NAACL* (Vol. 2, pp. 159–166). Pittsburgh, PA.
- Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, *48*, 281–303.
- Hare, M., McRae, K., & Elman, J. L. (2004). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*, *19*, 181–224.
- Hoeks, J. C. J., Stowe, L. A., & Doedens, L. H. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*, 59–73.
- Huetting, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*, 23–32.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156.
- Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*, 37–55.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.
- Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, *57*, 519–543.
- Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science*, *30*, 481–529.
- Kučera, H., & Francis, W. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.
- Kuperberg, G., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117–129.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.
- Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*, 21086–21090.
- MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.
- Magnuson, J. S., Dixon, J., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*, 133–156.
- Magnuson, J. S., Tanenhaus, M. K., & Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition*, *108*, 866–873.
- Magnuson, J. S., Tanenhaus, M. K., & Aslin, R. N. (2000). Simple recurrent networks and competition effects in spoken word recognition. *University of Rochester Working Papers in the Language Sciences*, *1*, 56–71.
- Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The microstructure of spoken word recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*, 202–227.
- Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions during word recognition in continuous speech. *Cognition*, *10*, 29–63.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.
- McClelland, J. L., St. John, M. F., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, *4*, 287–335.
- McRae, K., Ferretti, T., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*, 137–176.
- McRae, K., Hare, M., Elman, J. L., & Ferretti, T. R. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, *33*, 1174–1184.
- McRae, K., Spivey-Knowlton, M., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.
- Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*, 475–494.
- Movellan, J. R., & McClelland, J. L. (2001). The Morton–Massaro law of information integration: Implications for models of perception. *Psychological Review*, *108*, 113–148.
- Myung, J., Blumstein, S., & Sedivy, J. (2006). Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. *Cognition*, *98*, 223–243.
- O'Connor, C. M., McRae, K., & Cree, G. S. (2006). Conceptual hierarchies arise from the dynamics of learning and processing: Insights from a flat attractor network. In *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 2577). Mahwah, NJ: Erlbaum.
- Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, *34*, 521–542.
- Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahwah, NJ: Erlbaum.
- Shillcock, R., & Bard, E. G. (1993). Modularity and the processing of closed-class words. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 163–185). Mahwah, NJ: Lawrence Erlbaum Associates.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.
- Spivey, M. J. (2007). *The continuity of mind*. New York: Oxford University Press.
- St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, *46*, 217–257.
- Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.
- Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 431–450.

- Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50, 355–370.
- Tanenhaus, M. K., Leiman, J., & Seidenberg, M. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 18, 427–440.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives in sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16, 1272–1288.
- Yee, E., & Sedivy, J. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1–14.