Title: The interplay of local attraction, context and domain-general cognitive control in activation and suppression of semantic distractors during sentence comprehension

Authors: Nazbanou Nozari [1,2], John C. Trueswell[3], and Sharon L. Thompson-Schill[3]

[1]Department of Neurology, Johns Hopkins University School of medicine; [2]Department of Cognitive

Science, Johns Hopkins University; [3]Department of Psychology, University of Pennsylvania

Corresponding Author: Nazbanou Nozari

Email: nozari@jhu.edu

Address: Neurology Department, Johns Hopkins University School of Medicine

1629 Thames Street, Suite 350, Baltimore, MD 21231

Phone: 443-287-1712

Fax: 410-955-0188

Abstract

During sentence comprehension, real-time identification of a referent is driven both by local, context-independent lexical information and by more global sentential information related to the meaning of the utterance as a whole. This paper investigates the cognitive factors that limit the consideration of referents that are supported by local lexical information but not supported by more global sentential information. In an eye-tracking paradigm, participants heard sentences like "She will eat the red pear" while viewing four black-and-white (colorless) line-drawings. In the experimental condition, the display contained a "local attractor" (e.g., a heart), that was locally compatible with the adjective but incompatible with the context ("eat"). In the control condition, the local attractor was replaced by a picture that was incompatible with the adjective (e.g., "igloo"). A second factor manipulated contextual constraint, by using either a constraining verb (e.g., "eat"), or a non-constraining one (e.g., "see"). Results showed consideration of the local attractor, the magnitude of which was modulated by verb constraint, but also by each subject's cognitive control abilities, as measured in a separate Flanker task run on the same subjects. The findings are compatible with a processing model in which the interplay between local attraction, context, and domain-general control mechanisms determines the consideration of possible referents.

Keywords

Sentence comprehension, eye-tracking, cognitive control, local attraction, distractor activation

Introduction

Referent identification during sentence comprehension unfolds over time and is driven by multiple sources of linguistic and nonlinguistic information (see Barr & Keysar, 2006; Tanenhaus & Trueswell, 2006; and references therein). A focus on the dynamics of processing leads to a natural division of these information sources: reference in the moment is constrained by both local (lexical) and global (contextual) factors. Although it is obvious that the individual words being heard play a central role in what is considered as a referent, use of this information must be regulated when it clashes with prior context. Yet theories of sentence comprehension have disagreed on whether contextual information is consulted early or late during processing (see Dahan & Tanenhaus, 2004 for a review of competing accounts). In addition, in spite of growing evidence for the involvement of cognitive control in various aspects of sentence and discourse processing (Novick, Kan, Trueswell, & Thompson-Schill, 2009; Nozari, Arnold, & Thompson-Schill, 2014; Nozari, Mirman, & Thompson-Schill, 2016; Nozari & Thompson-Schill, 2015), the link between control processes and inhibition of context-incompatible information has received little attention (but see Brown-Schmidt, 2009; Nilsen & Graham, 2009). Critically, it is unclear whether inhibitory resources regulate the constraining effect of context, and whether such resources are shared between linguistic and non-linguistic domains or are domain-specific. This study answers these questions.

Sensitivity to context in sentence processing

Numerous studies have reported early context effects in lexical and sentential processing (e.g., Altmann & Kamide, 1999; Barr, 2008; Chambers & San Juan, 2008; Dahan & Tanenhaus, 2004; Magnuson, Tanenhaus, & Aslin, 2008; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). For example, the classic "cohort competitor" effect in the visual world paradigm (i.e., looks toward a buckle when hearing "bucket") can be eliminated in a constraining context ("Empty the…") as compared to an unconstraining one ("Click on the…") (Barr, 2008, Experiment 2). There are,

however, studies that show evidence of processing of related competitors even when rendered implausible by the context (e.g., Kukona, Fang, Aicher, Chen, & Magnuson, 2011; Swinney, 1979; Tabor, Galantucci, & Richardson, 2004; Tanenhaus, Leiman, & Seidenberg, 1979). Most recently, Kukona, Cho, Magnuson and Tabor (2014) showed that upon hearing a sentence such as "The boy will eat the brown cake", participants also considered a brown car, even though it was incompatible with the verb "eat" (see also Kukona & Tabor, 2011). While this study provides strong evidence for local, context-insensitive processing of information, this effect might be driven in part by salient bottom-up information in the scene: it is possible that the color brown activates the lexical semantic category for "brown" even before the word is heard, and when the time comes for choosing, there are two objects with the same salient feature competing for capturing visual attention. If attending to the brown car is truly due to bottom-up capture of attention by color (see Simons, 2000 for a review of the conditions where color is a pop-out feature), then there is no need for a self-organizing account.

The current design builds on Kukona et al. (2014), with one important difference: consideration of semantic competitors could not be explained by prior bottom-up activation of visual features. On the experimental trials, participants heard sentences like "She will eat the red pear" and looked at a scene with four black-and-white line drawings: a pear (target), a banana (a verb competitor), a heart (an adjective competitor = local attractor), and a fourth unrelated object (Figure 1; Table 1). On control trials the adjective competitor was replaced with a picture incompatible with the color adjective (igloo). The difference between fixation proportions to the adjective competitor (heart) and the control picture (igloo) indexed "local attraction", i.e., the context-insensitive influence of the local word.

This factor was crossed with a contextual manipulation: Constraining trials had a constraining verb like "eat", while non-constraining trials had verbs like "see". This second

manipulation followed Dahan and Tanenhaus (2004), and served to gauge the sensitivity of local

attraction to the presence of constraining context. Similar to the current study, Dahan and

Tanenhaus examined the effect of semantically constraining context, but in their study they

examined consideration of phonological cohorts. Looks to the phonological competitor was only

found with cross-spliced stimuli that favored the acoustic form of the competitor, and this effect

was not sensitive to contextual constraint in an early time window. The current study investigates a

similar issue but with adjectives that supported reference to a local competitor.

Cognitive control and sentence comprehension

While there is solid evidence for the involvement of control processes in comprehension

(e.g., Nozari et al., 2016; Sommers & Danielson, 1999; see Novick, Hussey, Teubner-Rhodes,

Harbison, & Bunting, 2014; Novick, Kan, Trueswell, & Thompson-Schill, 2009 for reviews),

relatively little work has explored the role and nature of such processes in limiting the

consideration of context-incompatible information. An exception is a recent study by Brown-

Schmidt (2009; see also Nilsen & Graham, 2009 for a similar concept in children). Brown-Schmidt

(2009) had participants play a game with the experimenter, in which they had to jointly determine

whether certain criteria were met for the arrangement of objects on a visual array. The

experimenter read sentences aloud to the participant, and participant's fixations were tracked. The

goal was to test whether between two objects of the same kind, the participant considered only the

one that the experimenter could not see (and would therefore ask about) or also the one that the

experimenter could see. The results showed that the participant's fixations on the pragmatically-

infelicitous object, in this case, the object in the common ground, could be predicted from their

incongruency scores on a linguistic Stroop task, but not on a non-linguistic no-go task in which

participants withheld a button-press response if a certain object appeared on the screen. These

results might be interpreted as compatible with a domain-specific control process: scores on a

linguistic task were predictive of the perspective taking, but scores on a non-linguistic task were

not. If so, the results can be taken to support theories that posit specialized resources for specific domains (e.g., Fedorenko, Behr, & Kanwisher, 2011) and even sub-processes within a domain (e.g., Caplan & Waters, 1999; Waters & Caplan, 1996), against those that postulate domain-general resources shared by multiple systems (e.g., Novick et al., 2009; Thompson-Schill, Bedny, & Goldberg, 2005).

However, the tasks used in Brown-Schmidt (2009) differed in more than material type: no-go tasks canonically tap into response selection and execution while Stroop also imposes strong conflict at the level of stimulus processing (Ni et al., 2000). Thus, another interpretation of Brown-Schmidt's results is that the ability that determined suppression of irrelevant information in perspective taking was suppression at the level of stimulus processing (i.e., excluding an object from possibly being a referent) and not the response (i.e., suppressing an eye movement towards a given object that might still be a cognitive candidate).

Thus, while these results suggest link between cognitive control and inhibition of competitors that were incompatible with at least one type of context (common ground), Brown-Schmidt's (2009) study leaves two questions open: (1) Are similar control processes involved in inhibiting other context-incongruent competitors? Specifically, is cognitive control involved in resolving the competition between global and local information during referent selection? (2) Are these processes domain-specific? These questions are addressed in the current study. To answer the first, we examined if suppression of semantic competitors that clash with the sentence's verb can be predicted from individuals' performance on cognitive control tasks. To answer the second, we used a variant of the Flanker task, with an embedded no-go task. The Flanker task requires suppression of irrelevant visual stimuli (the flanking objects) in order to determine the direction of the central object. We use cartoon fish as stimuli, facing left or right, and response buttons the positions of which correspond to the direction of interest (left button to indicate a central fish

facing left, right button to indicate a central fish facing right). The non-linguistic nature of stimuli and the spatial congruence of response buttons and the target direction minimizes reliance on the language system. Thus the effect size calculated as RTs in the incongruent trials (central and flanking fish facing opposite directions) – RTs in the congruent trials (all fish facing the same direction) provides an index of inhibitory control in the non-linguistic domain. We can then test if such an index predicts the magnitude of inhibitory control required to suppress the adjective competitor during sentence comprehension. A positive and reliable correlation speaks to an inhibitory control process that is shared between the linguistic and non-linguistic domains. Absence of such a correlation is consistent with specialized inhibitory control processes in each domain.

The task also involves a no-go component. Flanker trials prominently involve stimulus conflict, some response conflict, and little to no response execution difficulty, as indexed by low error rates (Ni et al., 2000), distinguishing them from no-go trials which prominently index response conflict and response inhibition. Importantly, the same non-linguistic materials were used for both trial types (Flanker and no-go). Finding an effect similar to that reported by Brown-Schmidt would support the involvement of domain-general control processes that mediate conflict resolution at the level of stimulus processing.

--- Figure 1, Figure 2 and Table 1 about here ---

Methods

Participants

Thirty-two undergraduates of the University of Pennsylvania (17 females, mean age = 21.03 ±1.87 yrs.), all right handed and native English speakers, participated in the study in exchange for payment.

Materials

The eye-tracking task

A complete list of stimuli, along with the rationale and criteria for the selection of experimental materials is presented in the Appendix. Twenty sets were created, each containing four trial types (see Table 1). A 2x2 design manipulated context (constraining vs. non-constraining verb) and local attraction (local attractor present = experimental vs. absent=control). Constraining (e.g., "eat") and non-constraining (e.g., "see") verbs were comparable in frequency (SUBTLEX; Brysbaert & New, 2009), and their inclusion was determined by norming on Amazon's Mechanical Turk (e.g., Buhrmester, Kwang, & Gosling, 2011). In the context of our paradigm, we use the term "competitor" to refer to pictures that compete for selection as a referent, based on the information available to the participant at each point in time. Local attraction was manipulated by including a picture (adjective competitor) that was compatible with the adjective, but not with the constraining verb ("heart" in Table 1). Choice of the adjective competitor was also determined by norming on Mechanical Turk, such that the adjective competitor would be at least as compatible with the adjective as the target. Seven out of the 20 adjectives were color adjectives (see the Appendix for a complete list). The control picture ("igloo" in Table 1) was selected by re-shuffling the adjective competitor pictures, such that for a given trial it was incompatible with both the adjective and the verb. Thus, adjective competitors acted as their own controls across different trials. In addition, twenty fillers were created with adjectives that, unlike in the experimental and control trials, provided no useful information in localizing the target (e.g., "good" compatible with all four pictures), and verbs that varied in how constraining they were. The use of adjectives in sentences that did not necessarily require an adjective may seem like an unnatural feature of the task, but recent work has shown that overspecification, especially with color adjectives is not unusual in speakers (Tarenskeen, Broersma, & Geurts, 2015).

Pictures were 300x300-pixel black and white line-drawings taken from either the IPNP corpus (Szekely et al., 2004), or Google images). Sentences, which had the fixed format "She will [verb] the [adjective] [noun].", were recorded by a native English speaker at 44.1 kHz. A mixed design was employed, such that each subject only encountered one trial type from each set (for a total of 20 trials, 5 of each type + 20 fillers + 4 practice trials in the beginning). There was, therefore, no repetition of auditory or visual stimuli in individual participants.

The Fish-Flanker task

The Fish-Flanker task was a variation of the classic Flanker task (Eriksen & Eriksen, 1974), with five cartoon fish, and three trial types (Figure 2). Participants indicated the direction of the central fish by pressing a button on the same side of the keyboard (left or right) as the direction that the fish was pointing. On the congruent trials, the central fish and the flanking fish all faced the same direction, while they faced opposite directions on the incongruent trials (100 trials; 50 facing left, 50 facing right for each). On the no-go trials, the flanking fish were dotted, cueing the subject not to respond (100 trials; 25 of each of the four direction combinations). There were a total of 300 trials, with 12 initial practice trials.

Apparatus

Participants were seated in a dimly-lit room, approximately 25 inches away from a 17-inch monitor with the resolution set to 1024×768 dpi. Stimuli were presented using E-Prime Professional, Version 2.0 software (Psychology Software Tools, Inc., www.pstnet.com). An Eyelink 1000 eye-tracker with chin-rest recorded participants' monocular gaze position at 500 Hz. Fish-Flanker responses were registered by E-prime via adjacent keys on a keyboard.

Procedure

Participants completed the eye-tracking task followed by the Fish-Flanker task in one session. For the eye-tracking task, they were instructed to "listen and look at the pictures" (no

response was required). They completed four practice trials, followed by 40 trials (20 critical, 20 fillers intermixed). Each trial began with a 1375 ms preview. In the first 1000 ms, the four line-drawings were presented in the four corners, and in the last 375 ms a shrinking red dot appeared at the center to draw the gaze back to the central location. After the preview, the sentence was presented through speakers at a comfortable listening volume. The position of the four pictures was randomized on every trial.

After a 5-min break, participants completed the Fish-Flanker task (12 practice trials with feedback, followed by 100 congruent, 100 incongruent and 100 no-go trials intermixed). On each trial, a fixation cross was presented at the center of the screen. The duration of presentation of the cross (ITI) was sampled from a uniform distribution ranging between 500 and 1500 ms. Next the five fish were presented at the center of the screen for 1000 ms or until a response was made. Participants responded with either the index or the middle finger of their (dominant) right hand or made no response if the trial was a no-go trial. When a response was required, the position of the response button was congruent with the direction of the central fish (e.g., left button for the fish facing left). The spatial congruency was chosen to minimize the need for verbal strategies during response selection.

Results

The eye-tracking task

Data were analyzed using Growth Curve Analysis (GCA; Mirman, Dixon, & Magnuson, 2008), a variant of multilevel modeling developed specifically to analyze time course data, in R 3.0.3 (http://www.R-project.org.). For all analyses, the pattern of fixations was analyzed using cubic orthogonal polynomial models, with random intercept and slopes for subjects. The critical effect is reflected on the interaction between condition and the model's polynomial terms. To keep the

results interpretable, only the intercept, linear and quadratic terms are included in this interaction. In discussion of the results, we focus primarily on the intercept, which reflects the average height of the curve, and can be used directly to compare the proportion of fixations in one condition vs. another. For critical (i.e., adjective competitor) effects we report the full model in tables. For all the analyses that follow, unless stated otherwise, we picked a pre-defined analysis window starting 200 ms after the onset of the adjective to allow for planning and execution of an eye movement, and ending at the average noun offset. Average duration of adjectives and nouns were 480 and 707 ms respectively, making the analysis window 987 ms.

Figures 3 and 4 show fixation proportion (±SE) to the target, verb competitor and adjective competitor when the verb was non-constraining and constraining, respectively. Local attraction was measured by comparing looks to the adjective competitor in the experimental and control conditions. When the verb was non-constraining, there were significantly more looks to the adjective competitor (heart) than control (igloo; $t = 5.05$, $p < 0.001$). Critically, when the verb was constraining, there was also reliably more looks to the adjective competitor than control ($t = 2.21$, $p = 0.034$; see Table 2 for full results). An interaction analysis revealed that the magnitude of local attraction was reliably smaller when the verb was constraining ($t = -2.096$, $p = 0.038$; See table 3 for full results). Complementary analyses of looks to the target revealed fewer looks to the target in the presence of the adjective competitor when the verb was non-constraining ($t = -2.66$, $p = 0.001$), no interaction between context and presence or absence of the adjective competitor ($t = 0.080$, $p = 0.42$), with a significant effect of the adjective competitor on target on the quadratic term ($t = 2.12$, $p = 0.04$) when the verb was constraining.

In summary, the results showed a reliable effect of local attraction in the presence of constraining context, with a timeline similar to that reported by Kukona et al. (2014, Figure 6): Local attraction started late, shortly before the noun onset, continued throughout the noun zone,

and was extinguished at the noun offset. In addition, comparison of the magnitude of local

attraction when the verb was and was not constraining revealed a reduction in the size of the effect

in the constraining context. These findings show that local attraction is reliable and its magnitude in

modulated by contextual constraints. Next we ask whether this modulation can be predicted from a

domain-general inhibitory process.

Fish-Flanker task

Mean error rate was 8.31 (SE=1.58). The majority of errors were commission errors in the

no-go condition: 6.34 (SE = 1.12). Error rates were slightly higher in the congruent (1.13, SE = 0.23)

than the incongruent (0.84, SE = 0.15) condition, but this difference was not significant (t(31) =

1.06, p = 0.30). Mean RT for correct incongruent trials (519 ms, SE = 11) was, however, significantly

longer than that of congruent trials (494 ms, SD = 10), when the distribution of log-transformed

RTs were compared (t(31) = 8.89;P < 0.001), replicating the classic congruency effect in the Flanker

task. The Flanker effect size was calculated as follows for each subject: RT(incongruent –

congruent). The average effect size was 25 ms (SE = 3).

Analysis of individual differences

Our results, in keeping with those of past studies (Dahan & Tanenhaus, 2004; Kukona et al.,

2014), showed late looks to the adjective competitor when the verb was constraining (~400 ms

after adjective onset). This analysis investigated whether looks to the adjective competitor in this

late time window were predicted by the strength of domain-general inhibitory processes. For each

individual, the magnitude of local attraction was calculated as average fixation proportions on the

adjective competitor in the experimental minus control conditions when the verb was constraining.

Participants varied considerably in their fixation proportions to the adjective competitor in the

experimental and control conditions, with an average effect size of 0.04 (SD = 0.1). This effect size

shows the magnitude of the ability to inhibit looks to the referent activated by the information in the sentence adjective.

Flanker effect size was not reliably correlated with baseline performance on the Flanker task (i.e., RT in the congruent condition; Spearman's rho = 0.19, p = 0.30). This shows that the Flanker effect size was not a reflection of the basic abilities for carrying out cognitive tasks that both Flanker and the linguistic task require (e.g., visual perception, speed of processing, etc.). This effect size can therefore be taken as a measure of the ability to inhibit irrelevant information in the visual domain. We the asked if the effect sizes in the eye-tracking and Flanker tasks were correlated. The upper and lower panels in Figure 5 show the correlations between the Flanker effect size and the number of no-go errors, respectively, with the magnitude of local attraction. The two variables were themselves not correlated (r = - 0.08, p = .66), so they were both entered as regressors in a GLM with the magnitude of local attraction as the dependent variable. Only the Flanker effect size was reliably predictive of local attraction (t = 2.17, p = 0.038; no-go effect: t = 1.23, p = 0.22; model's R^2 = 0.17).

-- Figure 5 about here --

General Discussion

This study tested the influence of earlier semantic context on inhibiting the consideration of incompatible semantic competitors of later words in the sentence, and found a reliable local attraction. This finding extended Kukona et al.'s (2014) claims of local attraction by showing that the effect could be observed in the absence of direct mapping of words to referents in the visual scene. Previously, Dahan and Tanenhaus (2004) had reported looks to phonological competitors if the cross-spliced phonetic information temporarily biased the listener towards the competitors. They, however, reported no interaction with the context, while our results showed sensitivity to

context. One possible explanation is the different nature of the biasing information (phonetic vs. semantic). However, a more likely explanation is the difference in the windows of analysis. Those authors used a narrow window of 350-500 ms from the onset of the critical word, because the main question of that study was whether context can impose an early effect on selection, while we were interested in consideration of competitors at any point. Indeed the visual inspection of their data suggests an at least numerically smaller local attraction in the presence of the constraining verb when a larger window is considered (Dahan & Tanenhaus, 2004, Figure 3), very similar to the current findings.

Regardless of the differences mentioned above, both the current study and that of Dahan and Tanenhaus (2004) found a late consideration of context-incongruent competitors. Why this late effect? Kukona et al. (2014) proposed a self-organizing model that predicts exactly such a pattern from the parallel influence of context and local attraction: the early context (e.g., "eat") activates the target and the verb competitor, whose rising activation increasingly suppresses the adjective competitor until the adjective arrives. This arrival has two consequences: (1) it drives down the activation of the verb competitor thus reducing its imposed inhibition on the adjective competitor, and (2) it directly supports the activation of the adjective competitor. Together, these two processes gradually lead to the late increased activation of the adjective competitor, unless it is suppressed by top-down control.

We then turned to the critical question of what determines the magnitude of local attraction. Kukona et al.'s (2014) simulations predict a practice effect: early in its training the model shows large local attraction, but this effect diminishes as the model receives more training. This is not surprising, given that initial processing is highly bottom-up, and it is only through feedback and learning that such bottom-up processing becomes sensitive to constraints. The model, thus, predicts that the more mature the linguistic system, the more constrained the bottom-up

processing. Comparing linguistic systems of different strengths has its challenges. One approach would be to compare local attraction in children vs. adults, where the linguistic systems are truly at different maturational stages; but so are some of non-linguistic systems such as the cognitive control system (e.g., Davidson, Amso, Anderson, & Diamond, 2006). Moreover, evidence suggests that developmental delay in cognitive control has consequences for real-time sentence comprehension in children (Choi & Trueswell, 2010; Trueswell, Sekerina, Hill, & Logrip, 1999; Weighall, 2008) although these developmental changes may be related to cognitive flexibility rather than inhibitory control (Woodard, Pozzan & Trueswell, 2016). It would thus be difficult to claim that any given difference between children and adults in local attraction is really due to the maturation of their language system alone. Another approach would be to use adult speakers, but to somehow quantify their linguistic competence. Vocabulary size (e.g., Borovsky, Elman, and Fernald, 2012) is a likely candidate, but it is unclear how much of the variability in linguistic competence of adults can be captured by that index. Moreover, above and beyond the maturation of the language system, other abilities may contribute to modulation of local attraction. This study posited that domain-general inhibitory control is one such factor.

Note that the local coherence phenomenon itself is already a challenge to one kind of mental modularity, namely, the mutual modularity of non-overlapping phrases in sentence representations. However, it is potentially compatible with the assumption that language processing is modular with respect to other cognitive processes. The difference between Kukona et al. (2014) and the current design is critical here: In the former, presence of colored objects on the screen immediately draws attention to color, as have been shown in many studies of visual search (e.g., Theeuwes, 1994), thus the induced competition is primarily visual, similar to Flanker task. Thus, it would not have been greatly surprising if the magnitude of inhibition in Kukona and colleagues' design was correlated with that in Flanker. The current design avoids the confound of strongly-guided visual capture for the following reason: we set up the materials such that (a) two of

the pictures were rated as incompatible with the adjective (i.e., banana and the fourth picture, e.g., antlers were never picked during norming to go with the adjective "red"). Of the two pictures that were compatible with the adjective, one (i.e., the target pear) had a lower probability of being associated with the adjective (see the Appendix for details). Thus it is quite unlikely that upon viewing such a scene the adjective was automatically activated. Past evidence supports this claim at least for color adjectives. Yee, Ahmed and Thompson-Schill (2012) found no evidence of color-based priming between pairs such as "cucumber" and "emerald", unless the priming task was preceded by a Stroop task which drew participants' attention specifically to color.

In summary, the probability of participants becoming aware of the upcoming competition simply by inspecting the visual scene, without hearing the sentence, is very low in the current design. It is at the point where the adjective becomes critical for referent selection (i.e., "She will eat the red…" to distinguish between banana and pear), that the adjective activates the adjective competitor. Thus, the ensuing competition is a direct result of sentence processing, as opposed to visual pop-out. We can now ask whether the ability to resolve the competition induced by sentence comprehension can be predicted from the ability to resolve the competition induced by non-linguistic visual cues as in the Flanker task.

Our results revealed that participants' ability to suppress irrelevant information in a Flanker task, but not their ability to withhold responses, predicted the magnitude of local attraction. Note, however, that the Flanker effect is indexed by RTs, which can be more sensitive in capturing variations among individuals than the error rates in the no-go task, thus making the null less reliable. However, this replicates an earlier null finding reported by Brown-Schmidt (2009), who found that no-go scores were not predictive of inhibition in perspective taking, while Stroop scores were. Similar to Flanker cost, Stroop cost taps more strongly into suppression of irrelevant information at the stimulus level than response inhibition, thus the correlation with local attraction

reflects a true suppression of linguistic information, as opposed to suppression of eye movements. Together, these findings suggest that cognitive control is involved in suppression of competitors that are in conflict with different types of contextual constraints. The results also answer a second question left open by Brown-Schmidt (2009), namely that inhibitory control processes in a non-linguistic task were still predictive of performance in sentence comprehension, suggesting the domain-generality of such processes.

In summary, these results rule out the possibility, consistent with prior work on lexical local coherence, that local attraction is simply due to low-level feature pop-out capture of attention interfering with sentence interpretation. On the other hand, prior accounts of this phenomenon have posited a purely bottom-up mechanism, while our results show that top-down cognitive control mediates the strength of local coherence effects. Collectively, these results support a sentence processing model in which activation and suppression of semantically related information are decided by the interplay of context and local attraction (as in a self-organizing model), regulated by domain-general top-down control.

Acknowledgment

References

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, *109*(1), 18–40.

Barr, D.J., & Keysar, B.(2006). Perspective taking and the coordination of meaning in language use. In: Traxler, M.J. & Gernsbacher, M.A. (eds.) *Handbook of Psycholinguistics (pp. 901-938)*. Elsevier: Amsterdam, Netherlands.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417–436.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*(5), 893–900.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(01), 77–94.

Chambers, C. G., & San Juan, V. S. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition*, *108*(1), 26–50.

Choi, Y., & Trueswell, J. C. (2010). Children's (in) ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, *106*(1), 41–61.

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*(2), 498.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433.

Kukona, A., Cho, P. W., Magnuson, J. S., & Tabor, W. (2014). Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 326.

Kukona, A., & Tabor, W. (2011). Impulse processing: A dynamical systems model of incremental eye movements in the visual world paradigm. *Cognitive Science*, *35*(6), 1009–1051.

Magnuson, J. S., Tanenhaus, M. K., & Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition*, *108*(3), 866–873.

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*(4), 475–494.

Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220–249.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., … Shankweiler, D. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience*, *12*(1), 120–133.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the

   garden-path: Improving sentence processing through cognitive control training. *Language,*
   Cognition and Neuroscience, *29*(2), 186–217.

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across

   multiple domains: Memory and language impairments following damage to ventrolateral

   prefrontal cortex. *Cognitive Neuropsychology*, *26*(6), 527–567.

Nozari, N., Arnold, J. E., & Thompson-Schill, S. L. (2014). The effects of anodal stimulation of the left

   prefrontal cortex on sentence production. *Brain stimulation*, *7*(6), 784-792.

Nozari, N., Mirman, D., & Thompson-Schill, S. L. (2016). The ventrolateral prefrontal cortex

   facilitates processing of sentential context to locate referents. *Brain and Language*, *157*, 1-13.

Nozari, N. & Thompson-Schill, S.L. (2015). Left Ventrolateral Prefrontal Cortex in Processing of

   Words and Sentences. In G. Hickok & S.L. Small (Eds.), *The Neurobiology of Language (pp. 569-*
   588). Waltham, MA: Academic Press.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in cognitive*

   sciences, *4*(4), 147-155.

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in

   young and older adults: The interaction of lexical competition and semantic context. *Psychology*

   and Aging, *14*(3), 458.

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context

   effects. *Journal of Verbal Learning and Verbal Behavior*, *18*(6), 645–659.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … Wicha, N. (2004). A

   new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–

   250.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, *50*(4), 355–370.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*(4), 427–440.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Tanenhaus, M.K. & Trueswell, J.C. (2006). Eye movements and spoken language comprehension. In M.J. Traxler and M.A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (pp. 863-900). Amsterdam, Netherlands: Elsevier.

Tarenskeen, S., Broersma, M., & Geurts, B. (2015). Overspecification of colour, pattern, and size: Salience, absoluteness, and consistency, Frontiers in Psychology, 6, doi: 10.3389/fpsyg.2015.01703

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human perception and performance*, *20*(4), 799.

Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, *15*(2), 219–224.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134.

Waters, G. S., & Caplan, D. (1996). Processing resource capacity and the comprehension of garden path sentences. *Memory & Cognition*, *24*(3), 342–355.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*(2), 75–95.

Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in

    executive function and language processing skills in child learners. *Journal of experimental child*

    psychology, *141*, 187-209.

Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless green ideas (can) prime

    furiously. *Psychological science*, *23*(4), 364-369.

Figure captions

Figure 1 – An example of the visual display on a trial of the eye-tracking task. Note that all pictures

are presented as seen here, in black and white, providing no direct bottom up cues for the adjective

(i.e., no red color in the display).

Figure 2 – The three trial types in the Fish-Flanker task. A manual response was required for the

Incongruent and Congruent trials where all fish were striped, but participants were instructed to

refrain from pressing any buttons if the flanking fish were dotted.

Figure 3 – Looks to the target (T, left panel), the adjective competitor (AC, right panel), and the verb

competitor (middle panel) in the experimental (AC = heart) and control (AC = igloo) conditions,

when the verb is non-constraining (e.g., see).

Figure 4 – Looks to the target (T, left panel), the adjective competitor (AC, right panel), and the verb

competitor (middle panel) in the experimental (AC = heart) and control (AC = igloo) conditions,

when the verb is constraining (e.g., eat).

Figure 5 – Correlations between the size of local attraction in the eye-tracking task (looks to the

adjective competitor in the experimental and control conditions when the verb was constraining)

and two measures in the Fish-Flanker task. The upper panel shows the correlation with the size of the Flanker effect (difference in RTs in the congruent and incongruent conditions for the go trials). The lower panel shows the correlation between with the number of commission errors on the no-go trials.

Table 1 – An example of a set with its four trial types. Each subject only gets one of the four.

Type	Sentence	Condition	Target	verb competitor	Adjective competitor/control	Unrelated
1	She will eat the red pear.	constraining/ experimental	pear	banana	heart	antlers
2	She will eat the red pear.	constraining/ control	pear	banana	igloo	antlers
3	She will see the red pear.	non-constraining/ experimental	pear	banana	heart	antlers
4	She will see the red pear.	non-constraining/ control	pear	banana	igloo	antlers

Table 2 – Results of comparing fixations on the adjective competitor in the experimental and

control conditions when the verb is constraining (trial types 1 and 2).

Fixed effects				
	coefficient	**SE**	**t**	**p-value**
Intercept	0.062	0.014	4.56	<0.001
linear term	-0.183	0.061	-2.98	0.004
quadratic term	-0.001	0.053	<0.01	0.994
cubic term	0.009	0.033	0.27	0.787
Condition*intercept	0.032	0.015	2.21	0.034
Condition*linear term	0.170	0.070	2.42	0.019
Condition*quadratic term	-0.153	0.061	-2.50	0.017
Random effects				
Subject intercept	**Variance**			
polynomial's intercept	0.0024			
linear term	0.1909			
quadratic term	0.1597			
cubic term	0.1136			
Subject|Condition slope	**Variance**			
polynomial's intercept	0.0036			
linear term	0.3039			
quadratic term	0.2659			
cubic term	0.2024			

Table 3 – Results of comparing the magnitude of local attraction (looks to the adjective competitor in the experimental minus control condition) when the verb is constraining and when it is not (trial type 1–2 vs. trial type 3–4).

| *Fixed effects*				
	coefficient	**SE**	**t**	**p-value**
intercept	0.100	0.018	5.45	<0.001
linear term	-0.27	0.069	-4.00	<0.001
quadratic term	0.017	0.059	0.289	0.776
cubic term	-0.067	0.025	-2.66	0.012
condition*intercept	0.087	0.023	3.74	<0.001
condition*linear term	0.210	0.080	2.64	0.009
condition*quadratic term	-0.272	0.082	-3.307	0.001
verb*intercept	-0.020	0.023	-0.876	0.383
verb*linear term	0.129	0.080	1.621	0.108
verb*quadratic term	-0.059	0.082	-0.717	0.475
condition*verb*intercept	-0.069	0.033	-2.096	0.038
condition*verb*linear term	-0.022	0.113	-0.192	0.848
condition*verb*quadratic term	0.277	0.117	2.37	0.019
Random effects				
subject intercept	**Variance**			
polynomial's intercept	0.002			
linear term	0.044			
quadratic term	0.001			
cubic term	0.005			
subject|condition*verb slope	**Variance**			
polynomial's intercept	0.009			
linear term	0.124			
quadratic term	0.112			
cubic term	0.056			
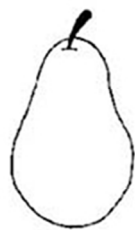
Figure 1 – An example of the visual display on a trial of the eye-tracking task. Note that all pictures are presented as seen here, in black and white, providing no direct bottom up cues for the adjective (i.e., no red color in the display).
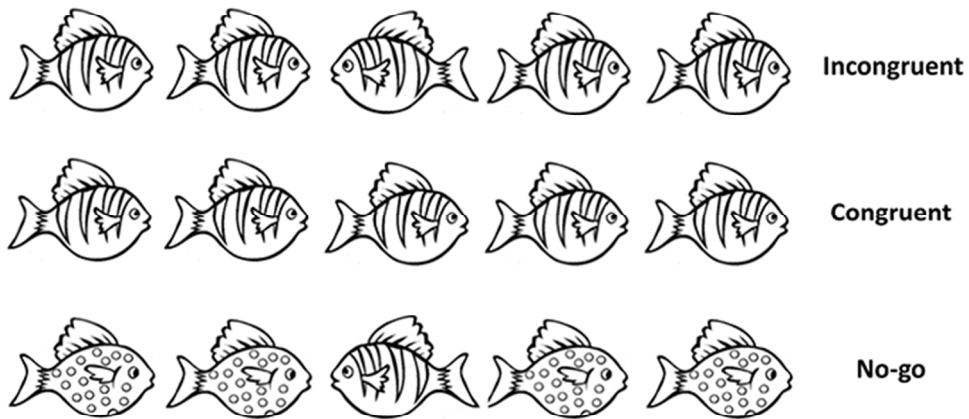63x50mm (300 x 300 DPI)

Figure 2 – The three trial types in the Fish-Flanker task. A manual response was required for the Incongruent and Congruent trials where all fish were striped, but participants were instructed to refrain from pressing any buttons if the flanking fish were dotted.
59x59mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
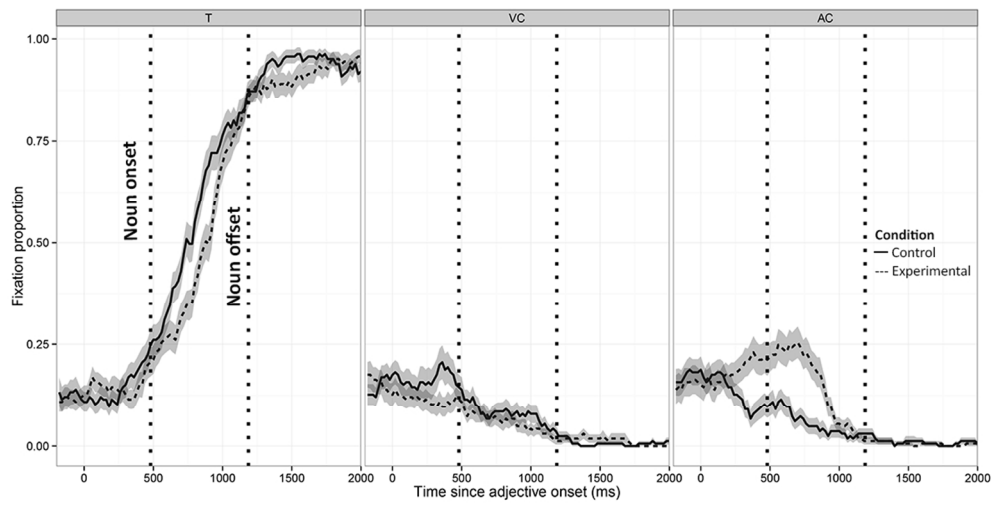45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3 – Looks to the target (T, left panel), the adjective competitor (AC, right panel), and the verb competitor (middle panel) in the experimental (AC = heart) and control (AC = igloo) conditions, when the verb is non-constraining (e.g., see).
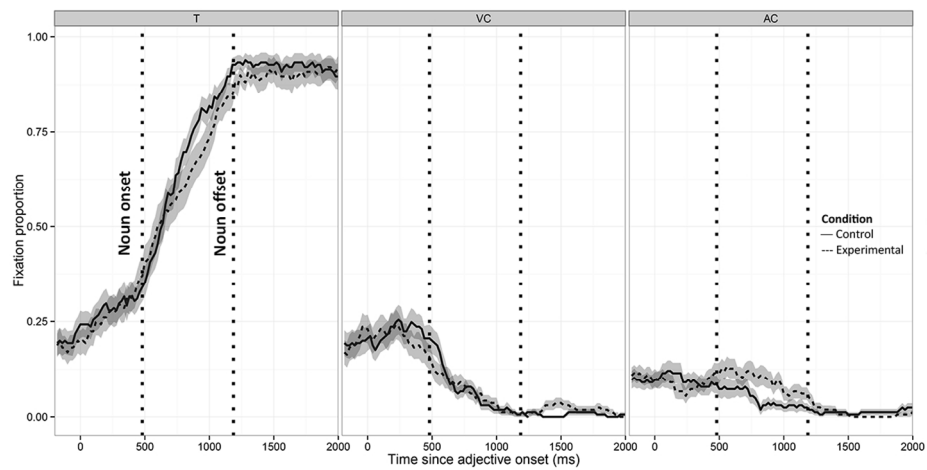101x52mm (300 x 300 DPI)

Figure 4 – Looks to the target (T, left panel), the adjective competitor (AC, right panel), and the verb competitor (middle panel) in the experimental (AC = heart) and control (AC = igloo) conditions, when the verb is constraining (e.g., eat).
118x56mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
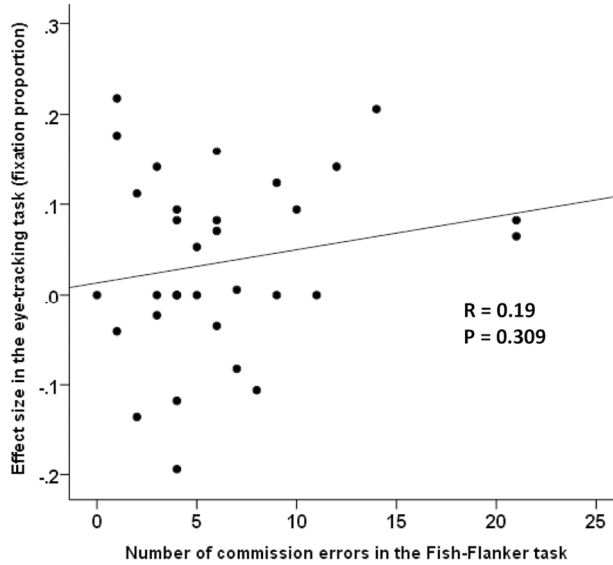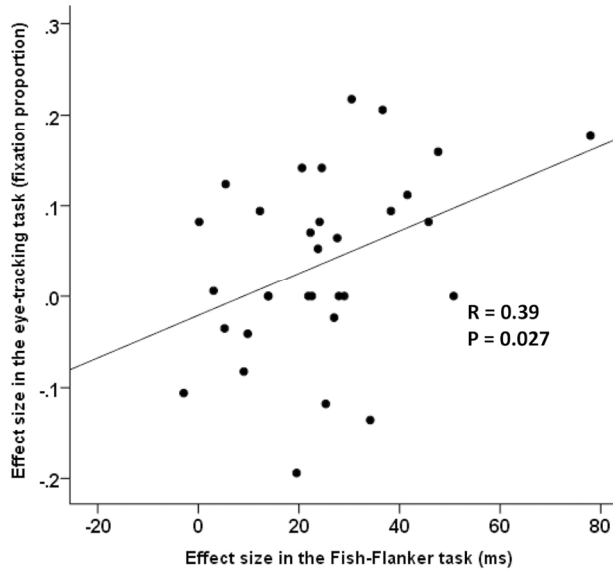51
52
53
54
55
56
57
58
59
60

Figure 5 – Correlations between the size of local attraction in the eye-tracking task (looks to the adjective competitor in the experimental and control conditions when the verb was restrictive) and two measures in the Fish-Flanker task. The upper panel shows the correlation with the size of the Flanker effect (difference in RTs in the congruent and incongruent conditions for the go trials). The lower panel shows the correlation between with the number of commission errors on the no-go trials.
396x671mm (72 x 72 DPI)

Appendix: Materials

Table A1- List of the sentences with constraining verbs in the experimental condition and three of

the four pictures: the target, the verb competitor (VC) and the adjective competitor (AC). The fourth

picture (the unrelated picture) is incompatible with neither the verb, nor the adjective and does not

play a critical role in the design. Controls for the adjective competitor are not shown in a separate

column because they are simply reshuffled ACs (see point 5 below).

	Sentence	Target	VC	AC
1	She will eat the red pear.	pear	banana	heart
2	She will play the brass saxophone.	saxophone	piano	trophy
3	She will write with the sharp pencil.	pencil	pen	needle
4	She will nibble at the red tomato.	tomato	cucumber	ladybug
5	She will feed the orange fox.	fox	pig	pumpkin
6	She will peel the white onion.	onion	orange	bone
7	She will hang the soft scarf.	scarf	picture	pillow
8	She will chop the shiny hair.	hair	asparagus	diamond
9	She will climb the wooden stairs.	stairs	mountain	birdhouse
10	She will plant the green artichoke.	artichoke	potato	frog
11	She will drink the icy juice.	juice	tea	igloo
12	She will close the wrinkled umbrella.	umbrella	door	shirt
13	She will pat the yellow bird.	bird	rabbit	lemon
14	She will wear the paper mask.	mask	pants	brownbag
15	She will lick the wedding envelope.	envelope	lollipop	veil
16	She will straddle the electric motorcycle.	motorcycle	horse	guitar
17	She will cage the red lobster.	lobster	cat	cherry
18	She will trap the prickly porcupine.	porcupine	raccoon	cactus
19	She will take off the metal bracelet.	bracelet	dress	cage
20	She will taste the hot pizza.	pizza	ice cream	fire

The materials were chosen from a larger pool of items normed on Amazon Mechanical Turk. We

collected norms from 42 people, out of whom 39 met the criteria for being a native speaker of

American English between 18 and 35 years of age. Pairs of pictures (e.g., pear/banana, pear/heart)

were presented along with the critical word (e.g., verb/adjective) and participants were asked to

choose the picture they thought matched the word. The final set was selected based on norms from

this population, using the pre-defined criteria listed below:

1. The adjective competitor and the unrelated picture were included only if none of the 39

    participants chose them in conjunction with the verb (i.e., the picture of a heart was never

    chosen to go with the verb "eat" when presented along with the picture of a pear).

2. Verb competitors were chosen such that they would have 50% - 75% chance of being

    picked given the constraining verb. For example, the verb "eat" was presented with pictures

    of pear and banana. Banana was accepted as the verb competitor if it was picked by the

    Turkers between 50% and 75% of the time. The lower bound of 50% was imposed to

    ensure that the sentence adjective was processed as an informative cue. If the verb

    competitor was so low in probability that the verb alone was sufficient to pick a unique

    target, there would be no reason for the adjective to be processed as a useful cue to the

    referent's identity. The upper bound was enforced to ensure that the targets were not ruled

    out in the first pass.

3. Both the target and the adjective competitor had to be compatible with the adjective. The

    same rule as above was applied. Given "red", heart was included if it was picked at least

    50% but no more than 75% of the time across participants, when presented among pear,

    banana, and antlers. This lower bound was critical for evoking inhibitory control.

4. Adjective competitors did not have the same onset as either the adjective or the noun. This

    constraint was imposed to ensure that local attraction was not induced by cohort

    competition.

5. Controls for the adjective competitor were assigned by reshuffling the pictures of the

    adjective competitors. The new combination was included in the norming to ensure that the

    verb and the adjective were both incompatible with the control. For example, "igloo", which

    was the adjective competitor on trial 11 in the above table, was included as the control for

the first trial, ensuring that it was incompatible with both "eat" and "red". Using the same

set of pictures as the adjective competitor allowed for the best possible control of the

experimental materials, and obviated the need for matching pictures for name-agreement,

complexity, familiarity, frequency, age of acquisition and other indexes, as such matching is

seldom perfect.

Table A2 - List of the non-constraining verbs used in the experiment.

	Non-constraining verb
1	like
2	get
3	see
4	think
5	take
6	look
7	talk (about)
8	remember
9	bring
10	watch
11	point (to)
12	observe
13	picture
14	imagine
15	notice
16	spot
17	describe
18	dislike
19	sketch
20	recognize

For the non-constraining trials, every picture was chosen at least 10% of the time across Turkers,

and none of the four pictures was chosen over 50% of the time (validating the assumption that

there was no global constraint in the case of non-constraining verbs and no singularly prominent

target).