Investigating Semantic Conflict between General Knowledge and Novel Information in Real-Time Sentence Processing

Angele Yazbec (yazbec@psy.fsu.edu)
Department of Psychology, 1107 W. Call Street
Tallahassee, FL 32306

Michael Kaschak (kaschak@psy.fsu.edu)
Department of Psychology, 1107 W. Call Street
Tallahassee, FL 32306

Arielle Borovsky (borovsky@psy.fsu.edu)
Department of Psychology, 1107 W. Call Street
Tallahassee, FL 32306

Abstract

There is extensive evidence that listeners use general knowledge to predict upcoming sentence endings; however, less is known about how novel information is integrated when there is disagreement between general knowledge and novel information. The present studies use the visual world paradigm to study the semantic competition between new information and general knowledge. Experiment 1 demonstrates that listeners learn to use limited exposure to new information and their general knowledge to anticipate sentence endings that align with the action of the sentence. Experiment 2 demonstrates participants learn to use combinatorial information from stories to elicit anticipatory eye movements to the target over the general knowledge distractor. Evidence from these experiments indicates even in the presence of semantic conflict with general knowledge, listeners rapidly increase the weight of novel information rather than general knowledge.

Keywords: visual world paradigm; sentence processing; general knowledge

Comprehending Novel Events in Real Time

Listeners actively interpret spoken language about familiar events by rapidly integrating information from multiple sources to incrementally generate expectations about upcoming input (Huettig, Rommers, & Meyer, 2011 for a review). However, not all spoken events convey highly expected information. How do listeners interpret this unexpected information in real time? Some research suggests that contextual support can prompt listeners to adapt their semantic expectations in potentially anomalous contexts. For instance, Nieuwland and Van Berkum (2006) measured N400 semantic mismatch responses to written sentences conveying normally anomalous events (e.g. a peanut in love). When these sentences were situated as plausible within a larger discourse (e.g. a peanut falling in love with an almond), participants did not show a N400 semantic mismatch effect to sentences about a peanut in love by the end of the story. Similarly, other studies find that comprehenders are faster to read pragmatic anomalies (e.g. "The mouse picked up the dynamite") when they are presented in cartoon settings that support these otherwise infelicitous events (e.g. the cartoon show *Tom and Jerry*; Filik, 2008; Filik & Leuthold, 2008). This prior research indicates that, with sufficient contextual support, it is possible to interpret otherwise unlikely events as semantically plausible. However, these findings do not answer whether participants were generating specific predictions based on the new information, or simply matching the information with the current discourse.

Recent studies suggest that listeners can use recently encountered novel events (e.g. a monkey riding in a bus) to generate predictions during spoken sentence processing (Amato & MacDonald, 2010; Borovsky, Sweeney, Elman, & Fernald, 2014). Thus, this new information can be used to support anticipatory language comprehension. Additional evidence from Kleinschmidt and Jaeger (2015) suggests that in speech perception, listeners compare statistics of a novel context to prior beliefs of how speech should sound. Subsequently, listeners rapidly adapt listening behavior by weighing newly acquired information more than prior beliefs. Thus, listeners quickly learn which source of information to rely on during comprehension.

It is still unclear, however, whether and how listeners adjudicate between cases where long established general knowledge and new information directly conflict. Nieuwland and Van Berkum (2006) provide some clues: When participants read a story containing a novel situation (e.g. a smitten peanut), by the end of the story, sentences that conveyed general knowledge about the event (e.g. the peanut was salted) elicited a strong N400 mismatch effect. This suggested that, during an extended discourse, listeners temporarily "suspended" their general knowledge about peanuts in favor of the newly relevant information. The current research seeks to disentangle these questions by using a visual world paradigm (VWP) approach to explore incremental interpretation of sentences that violate general expectations, (e.g. a pilot who flies a kite, rather than an airplane). Experiment 1 explored how listeners comprehend

isolated sentences that conflict with general knowledge without any other supporting context. Experiment 2 investigated how listeners' comprehension of these same events when proceeded by stories describing those events.

Experiment 1

This experiment explores how listeners resolve semantic competition between general knowledge and new information when isolated sentences (containing an agent, action and thematic object) are the only source of conflict. It is important to note that the thematic object is always something unexpected: participants should not anticipate this object because it does not align with their general knowledge. For example, when hearing, "The pilot flies the kite," participants should make anticipatory fixations on AIRPLANE, the expected ending based on general knowledge. Only after "kite" is spoken should participants primarily fixate on KITE. This presents direct semantic competition between novel information and general knowledge. At the beginning of the experiment, it is hypothesized that this limited information will not support anticipatory interpretation of sentence, as measured by fixations towards an image of the (to be spoken) thematic object before it is spoken. Instead, listeners should rely on general knowledge for comprehension and fixate on the object that coheres most strongly with that possibility.

After hearing a series of sentences that violate general knowledge, however, it is expected that participants should adapt their comprehension strategies such that they learn to anticipate objects related to the action of the sentence, but not the general knowledge distractor. Thus when hearing, "The pilot flies the kite," in the latter half of the experiment, participants should fixate on the flyable objects of KITE and SPACESHIP until the sentential object ("kite") is spoken.

Method

Participants

Fifty-four adults participated in the study and received course credit (M= 19.11 years, male = 14). Inclusionary criteria included: normal or corrected-to-normal vision, normal hearing, no history of diagnosis or treatment of cognitive, speech, language, or attentional issues, and monolingual English speaker.

Materials

Design

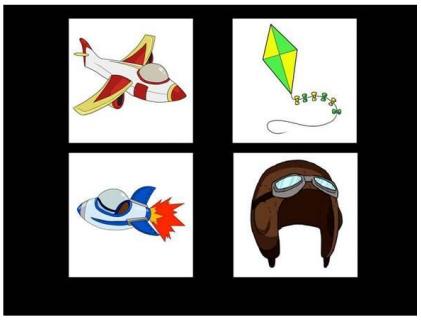
Hand-drawn, cartoon-styled pictures illustrated agents, objects, and agents acting on objects and were adjusted to a 400 X 400 pixel image on a white background. Sentences were pre-recorded by a female, native American-English speaker and sampled at a 44,150 Hz intensity level and normalized offline to 70 dB hearing level.

Agents and objects were paired such that the agent would be acting on an object that is unexpected based on general knowledge. Thus, the generally-expected object for one agent served as the target for a different agent. Each agent-object pairing was checked against Latent Semantic Analysis and the Edinburgh Associative Thesaurus to ensure the objects selected as general knowledge distractors were expected items in each pairing. Based on the combined data from these sources and previously normed relationships (Borovsky, Elman, & Fernald, 2012), 30 sentences were created. Each sentence has the following standardized construction: article 1, agent, action, article 2, object.

Sentence Comprehension Stimuli

Participants completed 15, four-alternative, forced-choice VWP tasks to assess sentence comprehension. Across all versions, image/sentence combinations were counterbalanced such that each of the 30 novel relationships was equally likely to be tested, and the locations of images of the target, verb-related distractor, agent-related distractor, and general knowledge distractor were equally likely to appear in all quadrants on the screen.

Because we were interested in the timing of fixations as the sentence unfolds, the durations of each word were controlled offline using Praat audio editing software (Boersma & Weenink, 2012) following the procedure outlined in Borovsky et al., 2014. Sentences were carefully normed such that they were all the same length and each word in each sentence started at the same time. Each trial started with 2000 ms viewing period during which the pictures appeared on the screen without auditory stimuli. After 2000 ms, the sentence was spoken and participants clicked on the object that corresponded with the sentence.



Spoken sentence:					Story Target	Agent-Related	Verb-Related	General Knowledge
The	*pilot*	*flies*	*the*	*kite.*	KITE	HELMET	SPACESHIP	AIRPLANE
The	*pilot*	*wears*	*the*	*spacesuit.*	SPACESUIT	AIRPLANE	T-SHIRT	HELMET
The	*astronaut*	*flies*	*the*	*airplane.*	AIRPLANE	SPACESUIT	KITE	SPACESHIP
The	*astronaut*	*wears*	*the*	*t-shirt.*	T-SHIRT	SPACESHIP	HELMET	SPACESUIT
The	*boy*	*flies*	*the*	*spaceship.*	SPACESHIP	T-SHIRT	AIRPLANE	KITE
The	*boy*	*wears*	*the*	*helmet.*	HELMET	KITE	SPACESUIT	T-SHIRT
Art 1	**Agent**	**Action**	**Art 2**	**Object**				

Figure 1. Illustration of the stimuli used for sentence comprehension tasks in Exp 1 and 2.

Procedure

Experimental Task

Participants sat in a stationary chair in front of a computer with a 17-inch LCD display. A five-point calibration

procedure with a black and white 20-pixel bull's-eye image ensured proper set-up and tracking of the participant's right eye. The computer running the Eyelink Experiment Builder software (SR Research, Mississauga, Ontario, Canada) presented stimuli to the participants. The instructions were to listen carefully to the sentences and then use the mouse to click on the picture that goes with the spoken sentences. Before presenting the stimuli, the 20-pixel bull's-eye image would appear on the center of the screen for drift correction. After fixating on it, the sentence comprehension trial began.

Eye Movement Recording

Eye-movements were recorded using an EyeLink 1000 remote eye-tracker with a remote arm configuration and sampled at 500 Hz. The eye-tracking camera was attached to an LCD display and adjusted so that it was 580-620 mm away from the participant's right eye. Participants wore a target sticker over the right eye to accommodate for head and eye movements relative to the camera. Eye movements, classified as saccades, fixations, and blinks by the software, were measured during the sentence comprehension task starting from the moment objects were presented on the screen until participants selected a picture. Eye movements were binned into 20 ms intervals offline for analysis.

Results

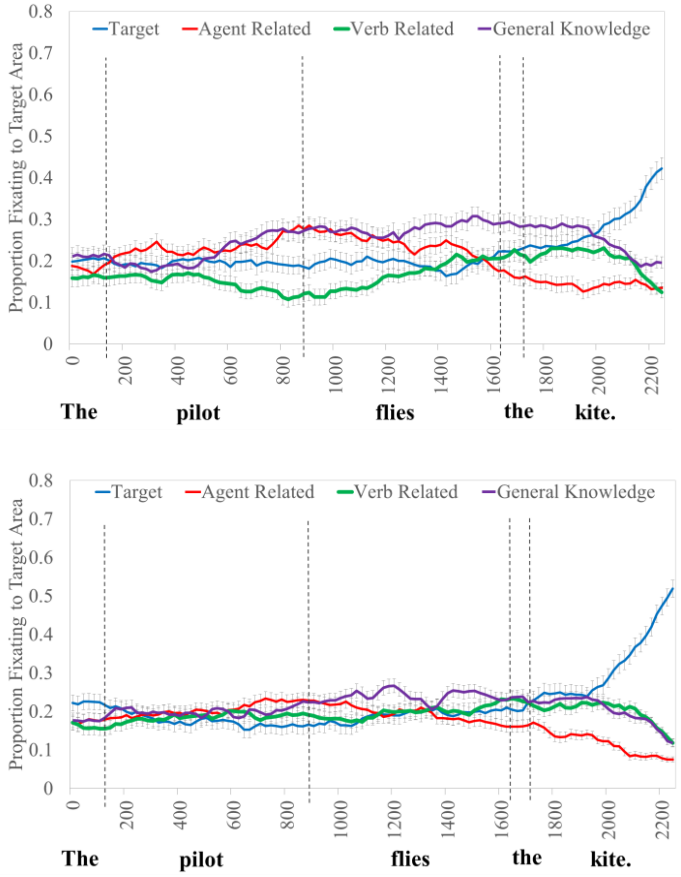
Behavioral Accuracy & Eye Movement Analysis

All analyses mentioned below are conducted only with accurate trials (accuracy= 97.9%). Many analyses were conducted to explore anticipatory fixations to the target and general knowledge distractor and whether there were any strategy changes throughout the experiment. Anticipatory fixations are defined as fixations on one object over the other objects during the action window (860-1599 ms after sentence onset) and/or second article window (1620-1720 ms after sentence onset).

Time-course Visualization

Time-course fixations across the sentence were calculated by mean proportion of time spent fixating on Target, Agent-Related, Action-Related, and General-Knowledge Distractor items in 20 ms bins, averaged across all participants.

During the first seven trials, listeners launched anticipatory fixations to the general knowledge distractor rather than the target and verb-related distractor. For example, when hearing, "The pilot flies the…" listeners anticipate AIRPLANE because according to general knowledge, airplane relates to pilots and what they fly. For the last eight trials of the experiment, there is evidence that listeners increased the weight of other possible endings; they anticipated any object that fits in with the action as a potential sentence ending, including the general knowledge distractor. For example, upon hearing the sentence, "The pilot flies the kite" they anticipate the KITE, AIRPLANE, and SPACESHIP since they relate to "flies."



Figures 2 & 3. Time-course plot of fixations to all interest areas in 20 ms bins for the first 7 and last 8 trials (*N*=54).

Analysis of Anticipatory Fixations

We measured anticipatory fixations by computing fixation proportions to the target versus the general knowledge distractor in 20 ms time bins (see Borovsky et al., 2014 for similar approach). The following main log-gaze proportions ratio were calculated: Target vs. General-Knowledge, log (P(Target/P(General)), Target vs. Verb-Related, log (P(Target/P(Verb)), Target vs. Noun-Related, log (P(Target/P(Noun))General-Knowledge vs. Target, log (P(General/P(Target)), General-Knowledge vs. Verb-Related, log (P(General/P(Verb)), and General-Knowledge vs. Noun-Related, log (P(General/P(Noun)).

Transforming the proportion of looks to each image avoids violations of linearity and homogeneity of variance and allows for an investigation in which looks to the target are relatively biased against fixations to the other objects, but not necessarily meaning fewer looks to other items. Scores of zero indicate equal number of looks to the target and distractors, positive scores show that looks to the target exceeded looks to distractors, and negative scores indicate that looks to distractors exceeded looks to the target.

A nonparametric cluster analysis was performed to determine when fixations to the target and each distractor significantly differed from each other during the sentence

(see Groppe, Urbach, & Kutas, 2011 and Maris & Oostenveld, 2007 for more detail). Results from the first seven trials reveal that listeners anticipated the general knowledge distractor over the target during the agent, action, and object windows, 720-1680 ms after sentence onset (t=68.24, p<0.01), and the verb-related distractor in the same windows, 540-1680 ms after sentence onset (t=205.34, p<0.01). Fixations to the target do not exceed those to the general knowledge distractor until the object window, 2080 ms after sentence onset (t=46.97, p<0.05). Listeners also ruled out the noun related distractor as a potential thematic object starting at the action window, 1480 ms after sentence onset (t=119.26, p<0.001). In line with our hypothesis, listeners rely heavily on general knowledge early in the experiment to anticipate sentence endings.

The last eight trials, however, reveal that fixations to the general knowledge distractor do not exceed those to either the target (t=28.76, p=0.11) or verb-related distractor (t=10.51, p=0.42) at any point during the sentence. Despite this accommodation to entertain other potential sentence endings, the noun related distractor is still not considered a possible candidate. Listeners anticipate the general knowledge distractor over the noun related distractor starting at the action window, 1420 ms after sentence onset (t=125.40, p<0.001) and the target over the noun related distractor starting at the second article window, 1700 ms after sentence onset (t=202.03, p<0.001).

While we see that objects related to the sentence action are anticipated as potential sentence endings, fixations to the target do not exceed those to the general knowledge distractor until the second article window, 2000 ms after sentence onset (t=99.14, p<0.01). This suggests that participants rapidly learned anything relating to the sentence action may be an appropriate sentence ending. Contrary to our hypothesis, however, they did not rule out the general knowledge distractor as a possible sentence ending.

Experiment 2

Findings from Experiment 1 indicated that listeners expanded their expectations of sentence endings by using their knowledge of sentence themes in addition to general knowledge to anticipate any object associated with the action of the sentence. Despite this change, the general knowledge distractor was not uniquely eliminated as a potential sentence ending by the end of the study. In Experiment 2 we ask whether the addition of stories supporting an option that conflicts with general knowledge would shift this effect more strongly with time. A possible outcome is at the beginning of the task, listeners will rely heavily on general knowledge for comprehension. About halfway through the experiment, information from the stories will override general knowledge, which would be evident in anticipatory fixations to the story target over the general-knowledge distractor. This would suggest listeners rapidly integrate new information to launch anticipatory fixations to the story target. Alternatively, there could be no anticipatory fixations to the target, but equivalent looks to the target and general knowledge distractor until listeners start to hear the object of the sentence. This would indicate that even with additional contextual support, a few exposures to novel events are not enough to override general knowledge and elicit anticipatory looks to the target.

Method

Participants

Forty-eight college aged participants participated in the study and received course credit (M= 19.07 years, male = 16). Inclusionary criteria were the same as Experiment 1.

Materials

Story Design

To familiarize participants with novel information, they listened to stories set in a cartoon world before completing a comprehension task. These stories introduced novel relationships of familiar agents, actions, and objects. In each story, three agents performed the same two actions on different objects. The general-knowledge distractor for one agent would appear as the story-target for a different agent. First, the agent would appear on the screen. Afterwards, the agent would act on an object. Subsequently, the second and third agents would be introduced, completing the same action as the first agent, but acting on different objects.

Look! There's a pilot.	*What's that? He's flying a kite.*	*What's he doing now?*	*He's wearing a spacesuit.*
		**blank screen**	
Look! There's an astronaut.	*What's that? He's flying an airplane.*	*What's he doing now?*	*He's wearing a t-shirt.*
		**blank screen**	
Look! There's a boy.	*What's that? He's flying a spaceship.*	*What's he doing now?*	*He's wearing a helmet.*
		**blank screen**	

Figure 4. Illustration of a single story block used in Exp 2. Each image was presented once with an accompanying sentence (written in italics).

Procedure

Experimental Task

The experiment was administered using the same eye tracking system and eye movement recording procedure as Experiment 1. Instructions were to listen carefully to the

stories and then use the mouse to click on the picture that goes with the spoken sentences following the stories.

Participants heard stories about novel, semantically conflicting events accompanied by depictions of the agent and the agent acting on an object. Afterwards, participants completed three sentence comprehension trials corresponding with the three agents in the story. On the screen, four of the objects from the story were presented. Before the presentation of the sentence comprehension stimuli, the 20-pixel bull's-eye image would appear on the center of the screen to serve as drift correction. After fixating it, the sentence comprehension trial began.

Results

Behavioral Accuracy & Eye Movement Analysis

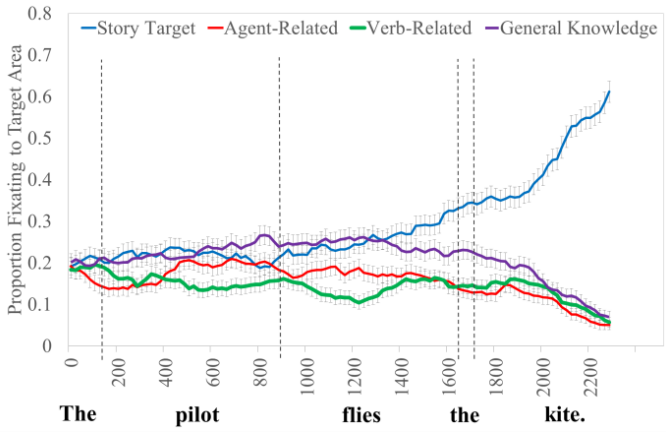
All analyses mentioned below include only accurate trials (accuracy = 98.6%). Analyses were conducted to explore anticipatory fixations to the story target over the general-knowledge distractor and whether comprehension strategy changes occurred throughout the experiment.

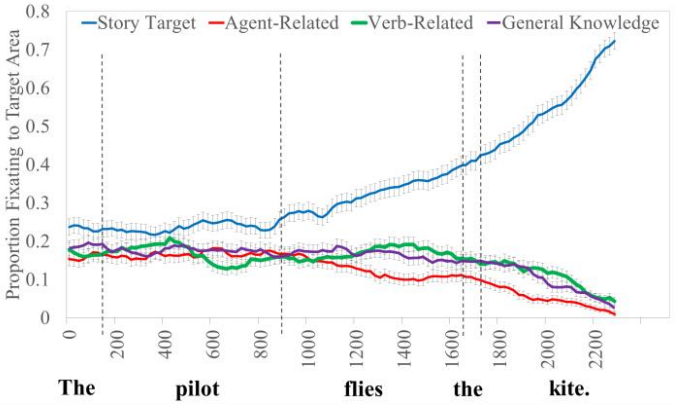
Time-course Visualization

Time-course fixations across the sentence were calculated by mean proportion of time spent fixating on the Story-Target, Agent-Related, Action-Related, and General-Knowledge Distractor items in 20 ms bins, averaged across all participants.

Visual inspection of the first seven trials shows nearly equivalent fixations to the story-target and general-knowledge distractor until after the action of the sentence is spoken. Participants are not uniquely anticipating either the story target or the general knowledge distractor as a potential sentence ending.

The remaining eight trials, however, show that listeners are able to anticipate the story target much earlier in the sentence; thus it appears that they are rapidly learning to weigh recently learned new information more than general knowledge, thus leading to more efficient comprehension.





Figures 5 & 6. Time-course plot of fixations to all interest areas in 20 ms bins for the first 7 and last 8 trials (N=48).

Analysis of Anticipatory Fixations

Due to the novel nature of the relationships in the stories, we measured whether anticipatory fixations occurred by measuring fixation proportions to the story target versus the general knowledge distractor in 20 ms time bins. This log-gaze proportion ratio was calculated: Story-Target vs. General-Knowledge, log (P(Target/P(General)).

Nonparametric cluster analyses determined when fixations to the story target exceed those of the general-knowledge distractor. Results reveal during the first seven trials, fixations to the story target do not significantly exceed those to the general-knowledge distractor until the object window, 1820 ms after sentence onset (t= 103.27, p < 0.001). Thus, there were no anticipatory fixations to the story target early in the experiment.

The remaining 8 trials, however, reveal that fixations to the story target exceed fixations to the general knowledge distractor starting at the action window, 1520 ms after sentence onset (t=213.58, p< 0.001). Thus, listeners begin to anticipate the story target over the general knowledge distractor.

For a finer grain analysis of changes in fixation patterns over time, an ANOVA was conducted to test for differences in mean fixations (averaged across all participants) to the story target and general knowledge distractor in the combined verb and second article windows (the anticipatory windows immediately before the object window) across the five experiment blocks. Significant differences in fixations to the story target and general knowledge distractor were present (F(4, 229) = 4.24, p< 0.01). Tukey's post-hoc tests revealed increased fixations to the story target over the general knowledge distractor in blocks 2 and 5 (p= 0.045) and 2 and 4 (p=0.049). There were marginally significant findings for increased fixations to the story target in blocks 1 and 5 (p= 0.072), 3 and 5 (p=0.073), 1 and 4 (p=0.078), 3 and 4 (p=0.079). Remaining comparisons were not significant. This shows that over the course of the experiment, listeners weigh recently learned information more than general knowledge. This increased weight of new information facilitated more efficient comprehension.

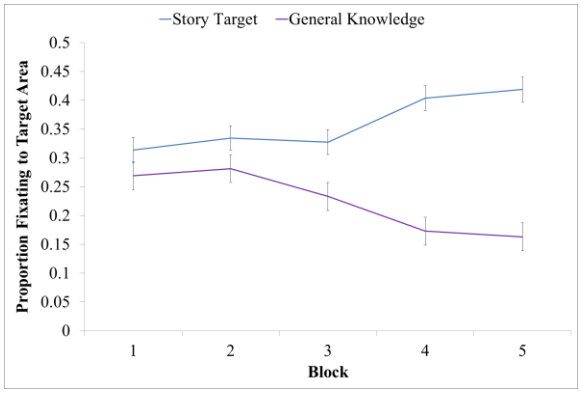


Figure 7. Mean fixations to the target and general knowledge distractor across experiment blocks (*N*=48).

General Discussion

Experiment 1 demonstrates that participants quickly considered alternative sentence endings when encountering repeated sentences that specifically conflicted with general event knowledge. Rather than generating fixations towards a general knowledge expected ending, as seen in sentence comprehension studies where sentences contain generally-expected information (e.g. Kamide, Altmann, Haywood, 2003), participants did not generate specific expectations for any particular ending. General knowledge is the primary source for comprehension (Nieuwland & Van Berkum, 2006); however, listeners quickly learned from the task that objects aligning with the action of the sentence could also be possible sentence endings. Despite increased consideration of these alternative possibilities, listeners never completely discounted their general expectations.

Experiment 2 reveals that additional discourse context via a narrated cartoon story could shift how listeners' weigh different sources of information for comprehension. During initial trials, listeners appear to weigh general knowledge and new information equally. Later trials, however, reveal that listeners weigh new information more than their general knowledge, and use that information to elicit anticipatory fixations to the story target. This suggests when general knowledge and novel information conflict, listeners rapidly learn to weigh new information more than general knowledge to understand the ongoing situation. This ability to override general knowledge is consistent with previous findings in discourse processing (e.g. Filik, 2008; Filik & Leuthold, 2008; Nieuwland & Van Berkum, 2006).

Both experiments reveal in the presence of semantic conflict between general knowledge and new information listeners rapidly increase their reliance on new information, as demonstrated by their willingness to entertain alternative sentence endings. Rapid changes in weighing new information have been demonstrated in previous studies (e.g. Kleinschmidt & Jaeger, 2015). Thus, while the utility of general knowledge never disappears, new information can outweigh general knowledge when given enough support to demonstrate the utility of the novel information.

Acknowledgements

The authors would like to thank Julie Carranza and Sam Stinchcomb for providing the vocal and artistic talents behind the stimuli used in these experiments.

References

Amato, M.S. & MacDonald, M.C. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition, 116* (1), 143-148.

Boersma, P. & Weenink, D. (2012). Praat: Doing phonetics by computer (Version 5.3.22).

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

Borovsky, A., Sweeney, K., Elman, J.L., & Fernald, A. (2014). Real-time interpretation of novel events across childhood. *Journal of Memory and Language, 73,* 1- 14.

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition, 106,* 1038-1046.

Filik, R. & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology, 45,* 554-558.

Groppe, D.M, Urbach, T.P., & Kutas, M. (2011). Mass univariate analysis of event-related potentials/fields: A critical tutorial review. *Psychophysiology, 48,* 1711-1725.

Huettig, F., Rommers, J., & Meyer, A.S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171.

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

Kiss, G.R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In: A. Aitken, R. Beiley & N. Hamilton-Smith (Eds.), *The computer and literary studies.* Edinburgh University Press.

Kleinschmidt, D.F., & Jaeger, T.F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, adapt to the novel. *Psychological Review, 122* (2), 148-203.

Landauer, T.K. & Dumais, S.T. (1997). *Psychological Review, 104,* 211-240.

Maris, E. & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG- data. *Journal of Neuroscience Methods, 164,* 177-190.

Nieuwland, M.S. & Van Berkum, J.J.A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18* (7), 1098-1111.