Eye-tracking situated language comprehension: Immediate actor gaze versus recent action events

Dato Abashidze (dabashidze@cit-ec.uni-bielefeld.de)

Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)

Maria Nella Carminati (mcarmina@techfak.uni-bielefeld.de)
Cognitive Interaction Technology Excellence Center, Department of Linguistics
Bielefeld University, Germany

Abstract

Previous visual world eye-tracking studies have shown that when a sentential verb can refer (via tense information on the verb and on a following time adverb) to either a recent and a future action event performed by an actor, people inspected the target of the recent event more often than the (different) target of the future event. This 'recent event preference' replicated even when the frequency of future events within the experiment greatly exceeded the frequency of recent events (e.g., 75% vs 25%). The recent event preference may arise because the past action is situation-immediate and thus more relevant at the particular point in time when the sentence is processed (at that point participants have seen the past action performed and will not see the future action until after the sentence). If the situation-immediate relevance of a cue is responsible for the recent event preference, then we should be able to "overwrite" the effect of the recent action with another situation-immediate cue. Accordingly, two current eye-tracking experiments pitted the recent event preference against a situation-immediate cue, the shift in the actor's gaze to the target object. Given that interlocutors' gaze has been shown to be a powerful cue in guiding listeners' attention to objects in the visual context, we hypothesized that the actor's gaze to the future target should rapidly guide a listener's attention to it. Analyses revealed indeed that listeners' visual attention was rapidly guided to the target by the actor's gaze; crucially the gaze cue was particularly helpful in guiding looks to the *future* target. Importantly, however, we still replicated the overall preference to look at the recent target regardless of tense and gaze; and even for future gaze conditions, the preference was not immediately reversed, suggesting it is surprisingly robust in competition with a situation-specific future-biasing cue.

Keywords: Sentence comprehension, recent-event preference, actor gaze, eye tracking, visual world

Introduction

Our immediate environment contains many extralinguistic cues that we exploit for understanding language. The information present in the visual context has been shown to provide powerful cues in guiding visual attention while we process language. This rapid and to some extent predictive interplay between language and visual context is reflected in the fact that people tend to gaze at objects as they are mentioned (Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) and often even anticipate their mention (Altmann & Kamide, 1999). For instance, case marking and verb meaning can rapidly guide a listener's visual attention, permitting her to predict what will happen next (e.g. Kamide, Altmann, & Haywood, 2003; Knoeferle & Crocker, 2007). In addition, listeners tend to rapidly rely on extralinguistic cues such as a speaker's eye gaze to objects which typically precedes mention of the objects. Following a speaker's gaze is beneficial for the listener since it permits him to anticipate which object the speaker will mention next (Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012; Staudte, Crocker, Heloir, & Kipp, 2014).

How precisely do listeners rely on the many available linguistic and extralinguistic cues? Existing account of situated language comprehension predict a rapid interplay of language comprehension, (visual) attenion, and visual context effects (e.g., Altmann & Kamide, 2009; Knoeferle & Crocker, 2007). In the absence of specific evidence to the contrary, it would be tempting to predict that the various cues are all on a par in contributing toward comprehension (all else being equal). Alternatively, some cues and / or world-language relations may be preferred over others. Determining the relative contribution of different cues and world-language relations is an important step in understanding situated language comprehension (Knoeferle, Urbach, & Kutas, 2014)

The recent-event preference vs. frequency biases

Consider, for instance, a series of visual-world studies which recorded comprehenders eye gaze in a scene as they listened to spoken utterances (Abashidze, Knoeferle, Carminati, & Essig, 2011; Knoeferle, Carminati, Abashidze, & Essig, 2011). These studies all pitted two world-language relations agains one another, viz. relating a verb to recently-inspected action and its target compared with using the verb to anticipate the target of a plausible future action. In the visual world study of Knoeferle et al. (2011) participants saw an actor sitting in front of a table with two objects. First the actor performed an action on one object (e.g., sugaring strawberries) and then participants heard either *(Der Versuchsleiter zuckerte kürzlich die Erdbeeren*, 'The experimenter recently sugared the strawberries') or *(Der Versuchsleiter zuckert demnächst die Pfannkuchen*, 'The experimenter will soon sugar the pancakes'). After the sentence had ended, participants saw the actor performing the same action on the other object (i.e. sugaring pancakes). Analyses of participants' gaze record during sentence comprehension showed a predominant preference of rapidly inspecting the target of the recent event (i.e. strawberries) during and after the verb. This preference emerged even when the verb was in the present tense and the adverb in the future sentence (denoting a future event), and it lasted well into the object noun phrase.

In further experiments Abashidze, Carminati, and Knoeferle (2014) used the same critical stimuli and design but increased the frequency with which during the experiment participants saw future events (in association with a future sentence) relative to recent events (88% vs. 12% of trials). Even with this strong frequency bias in favour of the future target, the overall preference to inspect the recent event target replicated. However the frequency bias did modulate participants' visual attention in that the preferential inspection of the future event target started much earlier (by 1000 ms) than when recent and future events were equally frequent in the experiment.

Gaze as a situation-specific cue

Perhaps the short-term frequency bias, while modulating visual attention, did not override the recent-event preference because the most recent cue in the context takes precedence. Indeed, recency effects are well documented in research on memory and cognition (Glanzer & Cunitz, 1966) and on visual attention (Zelinsky, Loschky, & Dickinson, 2011).

One situation-specific cue which has been shown to rapidly guide a listener's visual attention to upcoming, *future* objects is the gaze of an interactant. For example a study by Hanna and Brennan (2007) showed that listeners followed the speaker's gaze shifts to the target before it had been mentioned (for the robustness of this finding across different settings see also Staudte et al., 2014; Macdonald & Tatler, 2013; Knoeferle & Kreysa, 2012). If comprehenders incrementally update their visual anticipation based on recent and / or situation-immediate cues, then showing the actor shift his gaze to the future action target during the verb should replace the recent action as the most recent, situation-immediate cue.

The present experiments Given that an interactant's gaze has been shown to be a powerful situation-immediate cue in guiding listener's attention to soon-to-be-mentioned objects, two eye-tracking studies pitted object-directed gaze against the recent-event preference. These studies assessed to what extent an actor's gaze towards the (recent or future) action target influences listeners' visual attention. Can the gaze towards the future object overcome the preferred inspection of the recent event target? Between the experiments, we varied when the actor shifted his gaze (either during the verb or slightly earlier at verb onset) to also examine the effects of cue onset. We used the design from Knoeferle et al. (2011, Experiment 2), described above, with the factor tense (past vs. future) and we added gaze (to the target object vs. straight ahead) as a factor. The presence of gaze should trigger earlier and more frequent looks to the appropriate target objects (recent and future). Crucially, if the gaze cue overrides the recent-event preference, it should have a stronger influence in sentences with future than past tense meaning (i.e. tense x gaze interaction).

After the eye-tracking part participants completed a memory test (Expt 1) and a gated memory test (Expt 2). Abashidze et al. (2014) reported better recall of recent target objects. If in the memory test for Experiment 1 we replicate this, and recent (vs. future) events are anchored more firmly first in working and then in short-term memory, participants should be better at recalling the target of the recent (vs. future) events. In addition in the gated test in Experiment 2 we might replicate the higher recall of past (vs. future) tense sentences. Alternatively, the gaze cue has a strong influence on visual attention and the anchoring of events in working and short-term memory. If so, then we should see better recall for the future than recent event target.

Experiments 1 and 2: Methods

Participants

Thirty-two native speakers of German (aged 19 to 32, all students of Bielefeld University) with normal or corrected-to-normal vision received 6 Euros each for their participation. All gave informed consent.

Materials and design

We used the experimental items (N=24) from Abashidze et al. (2014, see Table 1). All critical sentences had the structure NP-V-ADV-NP and two male native German speakers recorded them. The experimental sentences were always about two objects and presented in two tense conditions. In one condition, the verb was in the present tense with a time adverb (*demnächst*, 'soon') indicating the future (Table 1a). In the other condition , the verb was in the simple past, and the following time adverb (*kürzlich*, 'recently') also indicated the past (Table 1b). Only German regular verbs appeared in the critical sentences to ensure the verb was tense-ambiguous up to but excluding the word-final phoneme which disambiguated towards the simple past in the past tense condition.

As can be seen from Table 1, there were two sentences for each tense condition; this counterbalancing ensured that each object was once the target of a past and once of a future event. In turn, this ensured that visual and other characteristics of any given post-verbal target object contributed equally to each critical condition. The words in a sentence were matched for spoken syllables and lemma frequency within an item (Baayen, Piepenbrock, & Gulikers, 1995).

For every item we recorded two videos ($M_{duration}$ =5015 ms) showing a person sitting at a table in front of two objects (e.g. cucumbers and tomatoes, one on the left and one on the right). The first video showed the person performing an action on one object (e.g., flavoring cucumbers, Fig. 1-1) and the second showed the person performing the same action on the other object (e.g., flavoring tomatoes, Fig. 1-3). The position of the target objects (right vs. left) was counterbalanced across items. In addition, we created a short *gaze* video for each experimental item (for about 3-4 seconds, depending on sentence length). The video showed the actor shifting his gaze to an appropriate object (e.g., in the future condition the actor would shift gaze to the tomatoes, see Fig. 1, 2a).

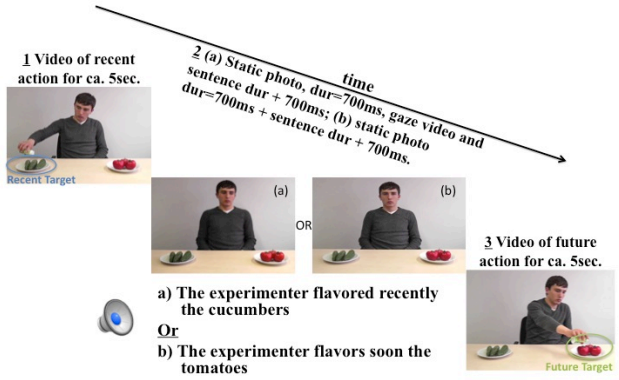
The tense factor (see Table 1, 1a vs. 1b) was crossed with actor gaze (target-directed vs. straight ahead, see Figure 1), resulting in four experimental conditions. For the *gaze condi-*

Table 1: Example experimental sentences. The indices (') indicate counterbalancing versions.

Tense Condition & counterbalancing	Sentences
1a future tense	*Der Versuchsleiter würzt demnächst die Tomaten* 'The experimenter will soon flavor the tomatoes'
1a' future tense	*Der Versuchsleiter würzt demnächst die Gurken* 'The experimenter will soon flavor the cucumbers'
1b past tense	*Der Versuchsleiter würzte kürzlich die Gurken* 'The experimenter recently flavored the cucumbers'
1b' past tense	*Der Versuchsleiter würzte kürzlich die Tomaten* 'The experimenter recently flavored the tomatoes'

tion we used the videos showing the actor shifting gaze to the target object (i.e., see Fig. 1, 2a). For the *no-gaze condition* we created a snapshot from the last frame of the first video showing the actor in a static position performing no action and looking straight ahead (i.e., see Fig. 1, 2b). Examples of the videos and the snapshot associated with the experimental sentences in Table 1 are shown in Figure 1 (1-3).

Figure 1: Sequence of events of a typical experimental trial



While we used the same 24 experimental sentences/videos in Experiments 1 and 2, the onset of the actor's gaze shift differed: In Experiment 1 this shift occurred on average 480 ms after the onset of the verb and in Experiment 2 it occurred at verb onset. Once he had shifted attention to the target, the actor continued to fixate it until the end of the sentence. In addition to the 24 experimental items we created 36 filler sentences. These ensured that participants were exposed to a range of other sentence structures and actions.

Actor gaze was one factor: in half of the trials the actor gazed at the target object and in the other half he did not. The second factor was the tense: in 50% of trials the sentence was in the past tense and in other 50% of trials in the future tense see Table 1 for counterbalancing. The resulting 8 lists used a Latin square design; each list contained every critical item in only one condition and all fillers. Each participant saw an individually pseudo-randomized version of one of the eight experimental lists. A gaze detection pretest (N=20) confirmed that participants were able to quickly detect the gaze shift.

Figure 2: An example of a display for the memory test, Expt 1 and Sequence of stages in the gated test, Expt 2

Procedure

For the eye-tracking experiment, participants were calibrated, and after a successful 9-point calibration of the eye tracker, the experiment started. Participants were told to inspect the scenes and to listen carefully to the sentences. On a given trial, a participant first saw a video of a person performing one action (e.g., flavoring the cucumbers, see Fig. 1, 1) and then a static picture (the last frame of the video) appeared. After 700 ms, the sentence started, and in the *gaze condition* the actor started to shift his eye gaze to the target object 480 ms after the verb onset in Experiment 1 and at verb onset in Experiment 2. In the *no-gaze condition* the static picture remained on the screen until 700 ms after the end of the sentence (see Fig. 1, 2b). 700 ms after the sentence had ended, participants saw a video of the actor performing the second action (e.g., flavoring the tomatoes, see Fig. 1, 3). Post-experiment, participants took part in a simple memory test in Experiment 1 and a gated memory test in Experiment 2. Finally, they were debriefed. Each experiment lasted approximately 50-55 minutes with a break after 25-30 minutes.

Memory tests

For the memory test in Experiment 1, we created two snapshots of the first and second video of each experimental item, i.e., showing the experimenter performing one of the two actions (see Fig. 2, Expt 1). The two snapshots associated with each item were combined into one display and shown to participants. Two versions were created in which the respective location of the two pictures was counterbalanced. Above the picture, one of two questions appeared:
(a) *Welche Aktion wurde VOR dem Satz durchgeführt?*
"Which action was performed before the sentence?"
(b) *Welche Aktion wurde NACH dem Satz durchgeführt?*

"Which action was performed after the sentence?"
Participants responded with a button press (e.g., if they thought that flavoring tomatoes was correct they would press the right side button for the image in Fig. 2, Expt 1).

In Experiment 2 participants took a gated memory test which provided more detailed insights into recall of sentence content on a per-constituent basis. Figure 2 (Expt 2) shows an example sentence as presented in the memory test in a 3-stage procedure. At the first stage, participants saw only the first noun phrase and the verb stem and had to verbally complete the verb tense. The second stage added the temporal adverb, and they had to recall the second noun phrase. If they were unable to do so, they received a further prompt at the third stage and had to select the correct referent out of three objects. Two of these were from that sentence trial and the third was a distractor from another filler item.

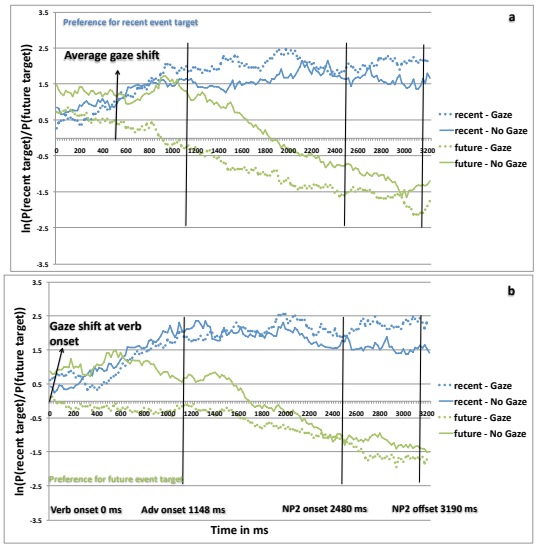
Experiment 1 and 2: Analyses and results

Eye tracking We defined a period of interest from the onset of the verb until the offset of the post-verbal object noun phrase. The measure of interest was fixations to the recent and future target objects (the cucumbers and the tomatoes) in the gaze and no-gaze conditions. We first computed gaze probabilities to the two target objects in each successive 20 ms time slots. Because looks to these two entities are not linearly independent (more looks to one object imply fewer looks to the other, and vice-versa), we computed mean log gaze probability ratios for the recent relative to the future target (ln (P (recent target)/P (future target))). A score of zero indicates that both targets are fixated equally frequently; a positive score reflects a preference for looking at the recent target over the future target, and a negative ratio indicates the opposite (see Abashidze et al., 2014; Knoeferle et al., 2011).

We used the log-gaze probability ratio to plot the time course from verb onset (Figure 3). In Figure 3 the blue lines indicate the recent condition (sentence in the past tense) and the green lines indicate the future condition (sentence in the future tense). The dotted lines indicate the gaze condition and the solid lines indicate the no-gaze condition. As can been seen in these graphs the gaze cue had an early influence on target inspection (dotted lines), but only in the future condition. In Experiment 1 participants started to preferentially inspect the future target (negative ratio) in the future condition approximately 400 ms after the gaze shift. However in Experiment 2 where the gaze shift occurred at the verb onset, ratios hover around 0 from a very early stage, suggesting the recent-event preference was eliminated. By contrast, in the past-tense condition, the gaze manipulation did not affect the distribution of attention until the end of the verb region (Expt 1) and until towards the end of the adverb region (Expt 2, see blue lines. Fig 3a-b).

For the inferential analyses, the dependent variable was the mean log gaze probability ratio averaged over the word regions (Verb, Adverb, NP2) by participants and by items, and the independent variables were gaze (gaze vs. no gaze) and tense (past vs. future). We assessed the recent-event preference in two ways (see Knoeferle et al., 2011). First, we tested the significance of the intercept overall (a positive intercept indicates a preference to inspect the recent action target). Second, we assessed significance of the intercept by condition (assessing effects of gaze and tense on the inspection preference).

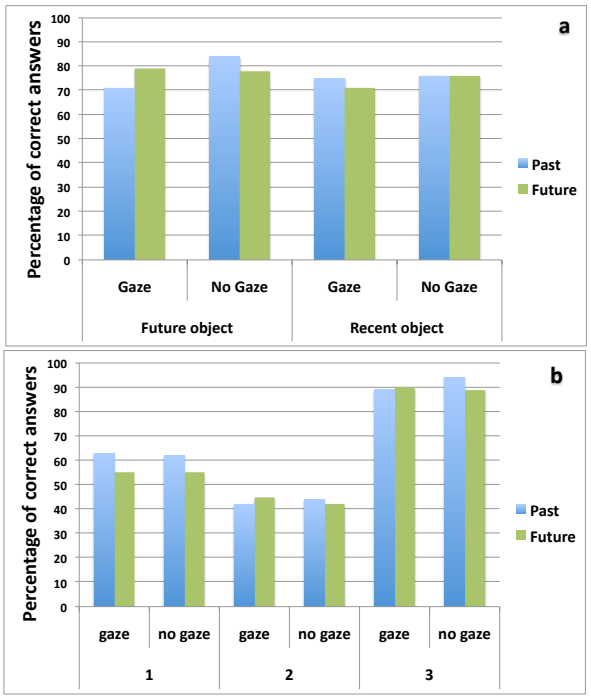
Figure 3: Mean log gaze probability ratios (ln (P(recent target/P(future target))) by condition from verb onset for Expts 1 and 2



Mixed effects ANOVAs (by participants and items) showed the grand mean (i.e., the mean of all conditions) was positive in all regions in both experiments, showing an *overall* preference for the recent target (significant intercept in all the ANOVAs by region in both experiments). Thus, the current experiments replicated the *overall* preference to look at the recent object independent of gaze cue and tense up to the very last sentential region. By contrast, pairwise comparisons revealed that gaze (vs. no gaze) enhanced looks to the future target ($p < .05$) in the future conditions in Expt 1 and 2 in the Verb and Adverb regions, suggesting a mitigation of the recent-event preference in the future tense gaze condition. With regard to the manipulated factors, a significant effect of tense (all ps $< .05$) emerged in both experiments, suggesting a reduction of the recent-event preference in the future compared with the past tense condition. In addition, the gaze effect was fully significant in the Verb region in both experiments (in the adverb region by participants in Expt 1). There was a Gaze x Tense interaction in all three regions in Experiment 1, whereas in Experiment 2 the interaction was significant by items at the verb and NP2 and marginal at the adverb ($p < .07$) regions.

Experiment 1: Memory test We calculated the percentage of correct answers by condition for participants and items separately. Figure 4a shows the average percentages (by participant) with 76% of correctly answered questions. The

graph shows that participants were more accurate in recognizing the future (78%) than the recent (74%) target objects. In logistic linear mixed effect (LME) analyses the effect of target object was significant, the effect of gaze was marginal with worse accuracy for gaze than no gaze conditions. There was a marginal interaction between object and gaze such that gaze had a stronger effect for future than past tense sentences.
Figure 4: Percentage of correct answers by object and tense (Expt 1); by tense and gaze (Expt 2)



Experiment 2: Gated memory test We calculated the percentage of correct answers by conditions for participants and items separately. The average percentages (by participant) are displayed in Figure 4b. Participants overall correctly answered 64% of the questions from all three stages. They were more accurate at stage one (59%) than two (43%), and accuracy was highest at stage three (with 90%) (see Fig. 4b, stages 1, 2 and 3). The LME analyses for stage 3 showed an effect of tense ($p < .003$, higher accuracy for past than future tense conditions) and of gaze ($p < .01$, higher accuracy without than with gaze), in the absence of an interaction.

Discussion

The current studies tested the recent-event preference (Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle et al., 2011; Abashidze et al., 2014) by pitting it against another situation-immediate cue (the actor's gaze). We assessed whether the recent-event preference replicates overall and whether participants would follow the actor's gaze to the future action target, thus effectively eliminating their preference to inspect the recent action target when the verb was ambiguous between referring to a recent action (and its target object) and a future action (and its different target object)

While participants *overall* preferred to inspect the recent event target (as indexed by reliable positive intercepts), replicating the recent event preference (Abashidze et al., 2014; Knoeferle et al., 2011), the immediate gaze cue clearly affected participants' visual inspection in the future tense condition. The eye-tracking results in Experiment 1 did show a rapid gaze cue effect especially in the *future tense conditions* during which participants started to inspect the future target already at the end of the verb, 450 ms after the gaze cue onset (see green dotted line, Fig. 3a). This effect continued until the end of the sentence. By contrast, in the no-gaze future condition participants preferentially inspected the recent event target until the middle of the adverb region. The timing of the gaze shift in relation to verb onset also affected mainly the future tense condition: in Expt 2 where the shift coincided with verb onset, the future target was looked at more and earlier than when it occurred 450 ms after verb onset (Exp 1). In fact, looks to the past and future target were balanced within 100 ms from the gaze shift in Expt 2 (see also for relevant related results, Friesen & Kingstone, 1998).

In the *past tense conditions,* the gaze effect was less pronounced and less immediate than in the future tense sentences. However, a better way to characterize this is that the inspection of the recent target during the verb was already so robust that the additional gaze cue did not lead to a further enhancement of looks to the recent target. Unlike in the future conditions, it was only towards the end of the verb region (in Expt 1) and in the middle of the adverb (in Expt 2) that gaze triggered more looks to the recent target in the gaze condition than in the no-gaze condition (see blue lines Fig 3a, b), this difference lasting until the end of the sentence.

Since gaze strongly cues the mention of upcoming objects, we might have expected a more immediate and full reversal of the recent-event preference at least in the future sentences. However, even the very early effect of gaze in the future condition in Expt 2 (Fig. 3b, green dotted line) did not lead to a sudden reversal of fixation preferences towards the future target; in fact for the first 1400 ms the log ratio hovers around zero, suggesting strong competition from the recent target.

The post-experiment memory test in Experiment 1 did not completely agree with the eye-gaze data (recall was better for future than past targets, against the recent event preference). While gaze (vs no-gaze) was beneficial in enhancing attention to future targets, it did not enhance target recall (Fig. 4 a) but was, in fact, detrimental. Perhaps gaze is only used 'on the fly' with short-lived effects on cognitive processes (see Knoeferle & Kreysa, 2012). In the gated memory-test (Experiment 2), past sentences were recalled better than future ones (in agreement with the recent event preference, Fig. 4 b). This provides some evidence for the view that past sentences anchor an event better in memory than future ones.

We can compare our findings with the results of experiments for which the design and frequency distribution was the same as for the no-gaze condition of the current experiments: In Experiment 2 by Knoeferle et al. (2011), the preference to

look at the recent object persisted until approximately 3000 ms after verb onset (when the ratio became negative). In the current experiments we instead see a considerably earlier reversal of the preference with the log ratio becoming negative at 1800 ms (Expt 1) and 2000 ms (Expt 2). Thus the fact that the actor gazed at the targets in some trials, seems to have led to an earlier shift of attention to the future target even in the no-gaze future condition of the current experiments.

Compared with a strong short-term frequency bias towards future events (Abashidze et al., 2014), the situation-immediate gaze cue had an earlier effect on the recent-event preference. When the actor shifted gaze towards the future target object, participants followed his gaze from 450 ms (in Exp 1) and from 100 ms (in Expt 2) after its onset (see Fig 3, no clear recent-event preference and log-gaze probability ratios hover around zero). By contrast, in the absence of a gaze cue when future events were more frequent than past ones (75 to 88%), participants' log-gaze probability ratio approached zero only approximately 1900ms after verb onset.

The present findings clearly shows that a situation-immediate cue modulated the recent-event preference earlier in the sentence than a short-term frequency bias towards future events. However, even gaze did not immediately reverse the preference, speaking to its robustness. The conflicting memory-test results suggest we need further experiments to assess the functional contribution of this attention preference. It could reflect an epistemic bias whereby a recent event dominates attention more than assertions about a future event. While a past event can generally be verified, a future one cannot until it has actually occurred, and until then it is uncertain if it will happen (e.g. Staub & Clifton, 2011; MacFarlane, 2003). Another possibility is that it reflects attention to whichever object representation is most highly activated in working memory.

Acknowledgments

This research was funded by the Cognitive Interaction Technology Excellence Center 277 and the SFB 673 "Alignment in Communication" (German Research Foundation, DFG).

References

Abashidze, D., Carminati, M. N., & Knoeferle, P. (2014). How robust is the recent event preference? In *Proceedings of the 36th Aannual Meating of the Cognitive Science Society.*

Abashidze, D., Knoeferle, P., Carminati, M. N., & Essig, K. (2011). The role of recent real-world versus future events in the comprehension of referentially ambiguous sentences: Evidence from eye tracking. In *European Perspectives on Cognitive Science: Proceedings of the EuroCogSci.*

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

Altmann, G., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cogn*, *111*, 55-71.

Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). The celex lexical database philadelphia: University of pennsylvania. *Linguistic Data Consortium.*

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 351–360.

Hanna, J. E., & Brennan, S. E. (2007). Speakers eye gaze disambiguates referring expressions early during face-to-face conversation. *JML*, *57*(4), 596–615.

Kamide, Y., Altmann, G. T. M., & Haywood, S. (2003). The time course of prediction in incremental sentence processing. *JML*, *49*, 133–156.

Knoeferle, P., Carminati, M. N., Abashidze, D., & Essig, K. (2011). Preferential inspection of recent real-world events over future events: evidence from eye tracking during spoken sentence comprehension. *Frontiers in Psychology*, *2*.

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: evidence from eye-movements. *JML*, *75*, 519–543.

Knoeferle, P., & Kreysa, H. (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? *Frontiers in Psy*, *3*.

Knoeferle, P., Urbach, T. P., & Kutas, M. (2014). Different mechanisms for role relations versus verb–action congruence effects: Evidence from ERPs in picture–sentence verification. *Acta Psychologica*, *152*, 133–148.

Macdonald, R. G., & Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of Vision*, *13*(4), 6.

MacFarlane, J. (2003). Future contingents and relative truth. *The Philosophical Quarterly*, *53*(212), 321–336.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*(4), 447–481.

Staub, A., & Clifton, C., Jr. (2011). Processing effects of an indeterminate future: Evidence from self-paced reading. *University of Massachusetts Occasional Papers in Linguistics*, *38*, 131–140.

Staudte, M., Crocker, M. W., Heloir, A., & Kipp, M. (2014). The influence of speaker gaze on listener comprehension: Contrasting visual versus intentional accounts. *Cognition*, *133*(1), 317–328.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 632–634.

Zelinsky, G. J., Loschky, L. C., & Dickinson, C. A. (2011). Do object refixations during scene viewing indicate rehearsal in visual working memory? *Memory & Cognition*, *39*(4), 600–613.