



# Preferential inspection of recent real-world events over future events: evidence from eye tracking during spoken sentence comprehension

Pia Knoeferle\*, Maria Nella Carminati, Dato Abashidze and Kai Essig

Cognitive Interaction Technology Excellence Cluster, Bielefeld University, Bielefeld, Germany

## Edited by:

Andriy Myachykov, University of Glasgow, UK

## Reviewed by:

Christoph Scheepers, University of Glasgow, UK

Falk Huettig, Max Planck Institute for Psycholinguistics, Netherlands

## \*Correspondence:

Pia Knoeferle, Cognitive Interaction Technology Excellence Cluster, Bielefeld University, Morgenbreede 39, Building H1, D-33615 Bielefeld, Germany.  
e-mail: knoeferl@cit-ec.uni-bielefeld.de

Eye-tracking findings suggest people prefer to ground their spoken language comprehension by focusing on recently seen events more than anticipating future events: When the verb in NP1-VERB-ADV-NP2 sentences was referentially ambiguous between a recently depicted and an equally plausible future clipart action, listeners fixated the target of the recent action more often at the verb than the object that hadn't yet been acted upon. We examined whether this inspection preference generalizes to real-world events, and whether it is (vs. isn't) modulated by how often people see recent and future events acted out. In a first eye-tracking study, the experimenter performed an action (e.g., sugaring pancakes), and then a spoken sentence either referred to that action or to an equally plausible future action (e.g., sugaring strawberries). At the verb, people more often inspected the pancakes (the recent target) than the strawberries (the future target), thus replicating the recent-event preference with these real-world actions. Adverb tense, indicating a future versus past event, had no effect on participants' visual attention. In a second study we increased the frequency of future actions such that participants saw 50/50 future and recent actions. During the verb people mostly inspected the recent action target, but subsequently they began to rely on tense, and anticipated the future target more often for future than past tense adverbs. A corpus study showed that the verbs and adverbs indicating past versus future actions were equally frequent, suggesting long-term frequency biases did not cause the recent-event preference. Thus, (a) recent real-world actions can rapidly influence comprehension (as indexed by eye gaze to objects), and (b) people prefer to first inspect a recent action target (vs. an object that will soon be acted upon), even when past and future actions occur with equal frequency. A simple frequency-of-experience account cannot accommodate these findings.

**Keywords:** visually situated sentence comprehension, eye tracking, visual context effects

## 1. INTRODUCTION

The role of prediction in language and cognition is a much-debated issue in the cognitive sciences. Prediction plays an important part in accounts of event perception (Zacks et al., 2007), in visual perception (e.g., Nijhawan, 1994; Berry et al., 1999), action anticipation (e.g., Miall et al., 1993; Wolpert et al., 1995; Aglioti et al., 2008), and in theoretical as well as modeling research on language comprehension (e.g., Elman, 1990; Hale, 2003; Federmeier, 2007; Pickering and Garrod, 2007; Levy, 2008). For language comprehension more specifically, the important role of predictive processes is evidenced by both findings from studies recording event-related brain potentials (e.g., Berkum et al., 2005; DeLong et al., 2005) and from studies tracking eye movements (e.g., Altmann and Kamide, 1999; Sedivy et al., 1999; Kamide et al., 2003a,b; see also Aborn et al., 1959; Tulving and Gold, 1963; Fischler and Bloom, 1979, for related early studies on word prediction in sentence context).

In more detail, both the current interpretation and linguistic as well as non-linguistic information from the immediate situation

can enable predictive processes during language comprehension. Visual event-related brain potential (ERP) recordings showed that when a definite article (e.g., *an*) was incongruous with the contextually most-expected noun (e.g., *kite* after *The day was breezy so the boy went outside to fly an...*), mean amplitude ERPs to the determiner were more negative going relative to when the determiner was congruous with the contextually most-expected noun (DeLong et al., 2005). Corroborating evidence for predictive processes based on the current utterance interpretation comes from "anticipatory" eye movements to target objects (i.e., eye movements to these objects before they are mentioned). Verb selectional restrictions (Altmann and Kamide, 1999), compositional noun and verb meaning, and associated world knowledge (Kamide et al., 2003a,b), prosody (Weber et al., 2006), or information structure (Kaiser and Trueswell, 2005) can each restrict the range of target objects that can be mentioned next, as evidenced by participants inspecting a target object before its mention relative to a control condition. Anticipatory gaze effects during spoken language comprehension can also be elicited by information from

the immediate non-linguistic context such as the actions that an object affords (Chambers et al., 2004), and verb-mediated depicted events (Knoeferle et al., 2005). In sum, language comprehension is characterized by a forward-looking mechanism that generates expectations about upcoming information based on the current interpretation, related linguistic, and world knowledge, as well as contextual information from the immediate situation.

In addition to information from the immediate situation, *recent* visual context information can also incrementally inform language comprehension (see Altmann, 2004; Knoeferle and Crocker, 2007; Huettig et al., 2011a), and memory task performance (Spivey and Geng, 2001). In Altmann (2004), after participants had inspected a man, a woman, a newspaper, and a cake, the screen went blank. Participants subsequently heard, for instance, *The man will eat*. . . Shortly after hearing *eat* they inspected the location where they had previously seen the cake before *cake* was mentioned. These findings corroborate the idea that semantic expectations during language comprehension are incrementally related to representations of recently inspected clipart objects (Altmann, 2004). A study by Knoeferle and Crocker (2007) extended these results to quasi-dynamically depicted clipart events and examined how visual interrogation of a scene is informed by information from events that participants had just inspected compared with events they could expect to happen in the near future. Participants saw a character (a waiter) move toward an object, interact with it (e.g., polish candelabra), and move away from it. People subsequently passively listened to an utterance that referred either to the recent action (polishing the candelabra: simple past tense: *Der Kellner polierte kürzlich die Kerzenleuchter*, “The waiter recently polished the candelabra”) or to an equally plausible action that hadn’t yet been performed (e.g., polishing crystal glasses; present tense with future meaning: *Der Kellner poliert sogleich die Kristallgläser*, “The waiter will soon polish the crystal glasses”). At the verb *poliert*. . . (“polish. . .”) the comprehension system and visual attention had a choice between anticipating the recent action target versus anticipating (and thus inspecting) the target of the as-yet-unseen future action. Participants preferentially anticipated the target of the recent (vs. the other, future) action, a gaze pattern that continued even as future tense information became available through the adverb (e.g., *sogleich*, “soon”). Verb meaning and future tense information did not elicit expectations of future events and people rather relied on the recently inspected events.

The present paper investigates in more detail how information from recent events compared with expectation of future events affects the visual inspection of (real-world) objects in visual context. Both visual anticipation and processes of accessing visual context information from working memory have been accommodated in existing accounts of situated language comprehension (see the Coordinated Interplay Account, CIA Knoeferle and Crocker, 2007). Overall, the Coordinated Interplay Account is concerned with accommodating the rapid interplay between language comprehension, (visual) attention, and subsequent feedback of non-linguistic visual information into comprehension processes. In line with existing evidence, the CIA assumes that comprehenders incrementally build an interpretation of the sentence and derive associated expectations. The (partial) interpretation built in this first stage directs attention (referentially but also

anticipatorily) to relevant aspects of visual context or representations thereof in working memory, and visual context information that is not immediately visually present experiences some decay. The representations of linguistic and non-linguistic content that are in the focus of attention are then co-indexed (e.g., grounding a verb in its action referent), and if necessary the interpretation is revised based on visual context information. As the next word is encountered, this temporally coordinated interplay continues. The three stages can overlap as the sentence is processed but they depend on each other for information.

When considering the observed preference to anticipate the recent (vs. future) event target, the CIA accommodates it via a reference-first mechanism. As people hear the sentence-initial noun “waiter,” they mostly inspect the waiter. Then they hear the verb “polish. . .” and all else being equal they first attempt to ground it in an action (representation) according to the CIA. This leads to participants inspecting the location at which the action took place. Less attention goes toward anticipating the target of future events (at least when a referential competitor – the action – has recently been seen and its target is still present). To the extent that the event representations of the recent events decay, the preference to inspect the recent-event target should decrease.

In the study by Knoeferle and Crocker, 2007, Experiment 3), however, decay was unlikely since for each critical trial only the “recent” event was depicted prior to sentence comprehension (and then referenced in the simple past in the ensuing sentence). The procedure of never depicting the future event may have created a within-experiment frequency bias toward relying more on recently depicted than on equally plausible future events. Perhaps because of this frequency bias, it has been argued that “the fact that the visual world took precedence in these studies over experiential knowledge is not surprising, of course, given that the most reliable cue to who is doing what to whom is whoever one sees doing it, not whoever one thinks is doing it. [. . .] no input is more privileged than another except insofar as one may be more predictive than the other in a given situation” (see, e.g., Altmann and Mirković, 2009, p. 596f).

These statements were made in the context of an alternative account of situated language processing by Altmann and Mirković (2009). In their model, information from the linguistic and non-linguistic visual context appears as representationally equivalent and to the extent that these two information sources are equally predictive none of them is preferred in predicting what will be mentioned next. Attention is allocated to objects through overlap between object representations (which are assumed to encode an object’s location), and linguistic representations derived from the unfolding utterance. The crux in interpreting the Altmann and Mirković account and their statements about the findings from Knoeferle and Crocker, 2007, Experiment 3 lies in understanding what these authors mean by a “reliable” cue and by input being only privileged to the extent that it is more “predictive.” The precise meaning of these terms in their paper isn’t explicitly defined, rendering their interpretation somewhat problematic. We believe that one “strong” but logically coherent interpretation of these terms within their account is that short-term and/or long-term experience of a given cue determines its predictiveness of subsequent input within their account. This is plausible since experience-based

knowledge and learning also play an important role in the Altmann and Mirković account which views language processing as governed by a mechanism that “[...] learns to anticipate, on the basis of its current and preceding input, what input may follow” (Altmann and Mirković, 2009, p. 589). Indeed, learning of statistical regularities is a hallmark of the connectionist network that Altmann and Mirković refer to in illustrating their account (Altmann and Dienes, 1999). Thus, in the absence of a clear definition of the reliability and predictiveness of a cue we instantiated predictiveness as the short-term frequency with which a participant experienced recent versus future events and long-term regularities of temporal cues in the sentences (e.g., past vs. future tense adverbs).

There are other considerations as to why a frequency-based account of Knoeferle and Crocker’s (2007, Experiment 3) findings is not implausible. In fact, in recent years it has become increasingly clear that human language comprehension and also other cognitive and motor processes are exquisitely sensitive to statistical regularities. In action execution, the recent trial-to-trial visuomotor experience can affect upcoming movement decisions (e.g., which one of two potential targets to reach for, Chapman et al., 2010). In language acquisition, statistical regularities can be exploited by children as young as 8 months for segmenting words in fluent speech (Saffran et al., 1996). Short-term linguistic experience can also modulate language production (Kaschak et al., 2006; Haskell et al., 2010) and sentence reading (Wells et al., 2009). Systematic co-variation (vs. random pairing) of novel target and distractor objects speeded up response latencies in identifying the target in a visual search task, suggesting that participants learned the associations between these two objects (Chun and Jiang, 1999).

Overall, then short-term experience of statistical regularities appears to play an important role in a number of cognitive and motor processes. To the extent that the importance of statistical regularities extends to perceptual experience of events, the frequency with which events are shown and then mentioned (“recent events”) versus the frequency with which events are performed after they were announced (“future events”), could plausibly affect how rapidly comprehenders access those events, and which ones they prefer to attend to during comprehension. An account in terms of short-term event experience could accommodate the rapid and preferred reliance on recent events when people only see recent and never future events (i.e., a bias of 100:0 toward recent events as was the case in Experiment 3 by Knoeferle and Crocker, 2007). Importantly, a short-term frequency account of the preferred reliance on recent events would further predict that as the ratio of recent versus future events that people perceive reaches a 50:50 frequency distribution (and assuming there is no linguistic frequency bias), the preferred inspection of the target of the recent event should be eliminated.

Alternatively (or in addition), the observed gaze pattern could be caused by comprehenders’ long-term linguistic experience. The recent actions in the waiter-polishing study were referred to by a verb in the simple past and an ensuing past tense adverb. The future events were indicated by a verb in the present tense with a future meaning and an ensuing future tense adverb. To the extent that the past tense verbs and adverbs may be more frequent than the present tense verbs and future adverbs, they might be processed

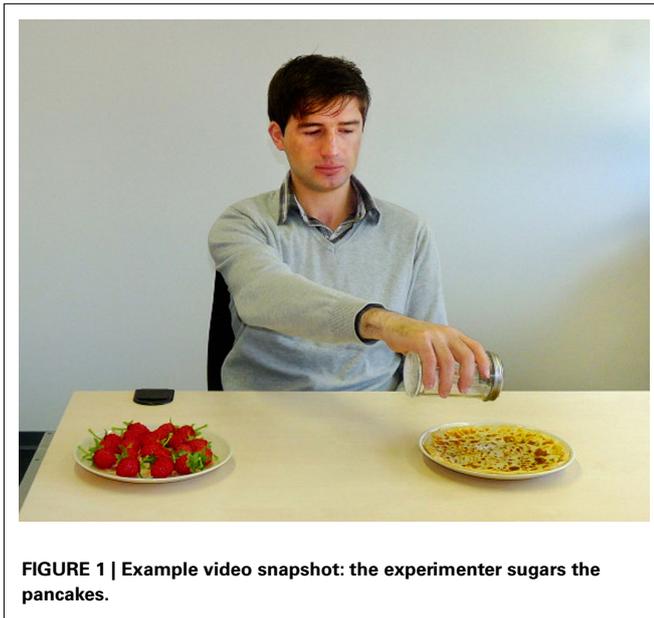
more rapidly, cueing comprehenders to preferentially inspect the target of the (recent) action that they refer to (see Dahan et al., 2001) for evidence on the effects of lexical frequency on visual attention to objects during spoken language comprehension.

In the original study (Knoeferle and Crocker, 2007, Experiment 3), people only ever saw the recent (and never the future) event on each trial. Seeing a recent event was thus a cue that was reliably followed by mention of the recent event. In contrast, hearing a sentence about a future event was never followed by actually seeing that future event, and thus an unreliable cue about future events. With a 50:50 frequency distribution, participants see an event and then it’s mentioned for half of the critical trials while on the other half of the trials, they hear an event mentioned and then they see it performed. According to the strong version of Altmann and Mirković’s (2009) account (see above), recent events should arguably be no more predictive than future events, and so a short-term frequency account would predict no particular reliance on recent actions.

Note that this is only a test of the Altmann and Mirković account to the extent that their account instantiates cue predictiveness exclusively via such statistical regularities. One could argue within their account that seeing an action increases the activation of that action representation. The action representation then overlaps with the representation of a corresponding verb, and modulates the attentional state such that the probability of an eye movement to the location associated with the activated action representation increases. In this way, the account might appear to predict more inspections to the target of the recent (vs. future) action. We think, however, that this argumentation logic doesn’t hold for a 50:50 within-experiment frequency distribution of recent versus future events. In the latter case, both remembering a recent action and anticipating a future action is equally predictive of what is mentioned/happens next. Thus, it would appear plausible that even after perception of one action activates its representations, the representations of other, relevant future events is activated just as much upon encountering the verb. Verb overlap with the future event representation could then boost the activation of those representations and modulate the attentional state such that the probability of saccades to the target of a plausible future event increases.

The Coordinated Interplay Account, in contrast, because of its mechanism of first grounding a referent would predict that even with a 50:50 frequency distribution of recent to future events, people should prefer to ground the verb in the recent action and its associated target. Thus, implementing a 50:50 frequency distribution of recent relative to future events (and controlling for linguistic frequency biases of the verbs and adverbs) would permit us to tease apart predictions of a reference-first mechanism from an account that rather emphasizes the predictive nature of an information source as instantiated by short-term frequency of event experience.

Two eye-tracking experiments and a corpus study addressed this question. To ensure that findings generalize to real-world environments, the present studies relied on real-world actions performed by the experimenter (see **Figure 1**). Experiment 1 aimed to replicate the findings from Experiment 3 in Knoeferle and Crocker (2007) with real-world action events, i.e., participants only ever



saw an event prior to sentence comprehension on each trial. The subsequent sentence either referred to that event (in the simple past) or it referred to another equally plausible event that could happen in the future. There was thus a 100:0 within-experiment frequency bias toward seeing recent (vs. future) events. By contrast for critical trials in Experiment 2, participants saw the experimenter perform one action prior to the sentence, and the other (future) action after sentence comprehension and overall in that study, the frequency distribution of recent relative to future actions was 50:50. Both the CIA and a short-term frequency instantiation of the account by Altmann and Mirković would predict a recent action preference in Experiment 1. In contrast, for Experiment 2, the CIA (but not a short-term frequency account) would appear to predict a preference to anticipate the target of the recently inspected event. An additional corpus study was conducted to gain insight into whether there was any linguistic bias such that past tense verbs and adverbs might be more frequent than present tense verbs and adverbs indicating future actions.

## 2. MATERIALS AND METHODS

### 2.1. PARTICIPANTS

Twenty-four German native speakers (aged 19 to 33,  $M = 24.83$ ; 8 males, 16 females) participated in Experiment 1, and a further twenty-four native German speakers participated in Experiment 2 (aged 19 to 33;  $M = 24.92$ , 12 males, 12 females). Participants (all students of Bielefeld University, Germany) were each paid 4 Euros to take part in the experiments. They all had normal or corrected-to-normal vision, were unaware of the purpose of the experiment and all gave informed consent in accordance with the Declaration of Helsinki.

### 2.2. MATERIALS AND DESIGN

We created twelve experimental items that each consisted of two everyday objects (e.g., strawberries and pancakes) and four sentences, recorded by a male native German speaker (see **Table 1** for

**Table 1 | Example item set for Experiments 1 and 2.**

1a	Future condition	<i>Der Versuchsleiter zuckert demnächst die Erdbeeren.</i> "The experimenter will soon sugar the strawberries."
1a'	Future condition	<i>Der Versuchsleiter zuckert demnächst die Pfannkuchen.</i> "The experimenter will soon sugar the pancakes."
1b	Recent condition	<i>Der Versuchsleiter zuckerte kürzlich die Pfannkuchen.</i> "The experimenter recently sugared the pancakes."
1b'	Recent condition	<i>Der Versuchsleiter zuckerte kürzlich die Erdbeeren.</i> "The experimenter recently sugared the strawberries."

*1a and 1b are examples of the conditions; 1a' and 1b' are the corresponding counterbalancing sentences.*

an example). Critical sentences were about the two objects and grouped into two tense conditions (future: 1a and recent: 1b). In the future condition, a present tense verb with a temporal adverb (*demnächst*, "soon") indicated the future. In the recent condition, tense was marked on the last letter of each verb (e.g., the *-e* in *zuckerte*, "sugared") and via the temporal adverb (*kürzlich*, "recently"). For the experimental sentences all words were matched for spoken syllables and lemma frequency within an item (Baayen et al., 1995). The counterbalancing versions (1a' and 1b' for 1a and 1b respectively) served to present each object once as the target of a recent, and once as the target of a future action, ensuring that visual characteristics of a post-verbal target object contributed equally to each of the two conditions.

The two objects of each experimental trial (e.g., strawberries and pancakes) could undergo the same action (e.g., sugaring). Experimental sentences about these objects began with *Der Versuchsleiter* ("The experimenter") followed by the verb (e.g., *zuckert...*, "sugar..."). Because of the counterbalancing, the two objects were equiprobable as targets of the action. Prior to the end of the verb (e.g., *zuckert...*, "sugar...") sentence tense was ambiguous and the sentence could thus either refer to a recent event (e.g., the experimenter had just sugared the pancakes) or to a future event (e.g., sugaring the strawberries). As the verb ending (*-e* in *zuckerte*, "sugared") and the adverbs were encountered, people could rely on the temporal cues to anticipate the recent versus the future event, although we know that prior research has reported weak effects of tense (Knoeferle and Crocker, 2007). However, the sentence-final noun phrase refers to the target of the recent (1b) versus future (1a) event; so, soon after people start processing this noun phrase, we should begin to see more eye gaze to the correct target (recent condition: pancakes, 1b; future condition: strawberries, 1a).

In Experiment 1, the experimenter performed only one action before the sentence for each experimental item (e.g., sugaring the pancakes), and then participants either heard a spoken sentence in the past (1b, **Table 1**) or in the future (1a, **Table 1**) condition. Participants thus saw 100:0 recent (vs. future) events and heard an equal number of sentences in the recent and future condition. In Experiment 2, the experimenter performed one action before the sentence (sugaring the pancakes), and another action after sentence presentation (sugaring the strawberries) on each critical trial such that participants not only heard equally many recent and future event sentences but also saw 50:50 recent to future events.

In addition to the twelve experimental items we created 24 filler sentences. These ensured that participants were exposed to a range of sentence and action combinations. Filler sentences were identical in the two experiments. They contained a verb in the past tense on 12 trials, and a verb in the present tense for the other 12 trials. In 8 filler trials the adverb indicated the recent past (4 trials) or the near future (4 trials). Adverbs for the other 16 filler sentences did not indicate a point in time but expressed mood, or degree of certainty of an event. The filler trials differed between the experiments in when people saw an action. In Experiment 1, the experimenter performed one action on each trial, prior to sentence presentation. In Experiment 2, for 8 of the filler trials, the experimenter conducted the action as the sentence was spoken. For another 8 filler trials, people only saw one action before sentence presentation (4 trials), or one action after sentence presentation (4 trials). For a further 8 filler trials, participants saw an action both before and after the sentence was presented. From the sentences in the two conditions and their two counterbalancing versions we created four lists using a Latin square. Each list contained every item in only one condition and all 24 filler sentences. Lists were pseudo-randomized and each participant saw an individually randomized version of one of the four experimental lists.

### 2.3. PROCEDURE

Participants were seated opposite the experimenter in front of a table. They were informed that the experiment would use an eye tracker (SMI iView X HED mobile), and they were calibrated using a 5-point calibration routine. When calibration was successful, the experiment started. Prior to the experiment, participants were instructed to look carefully at the items on the table and listen attentively to the recording played through the loudspeakers. There was no other task. For each trial, the experimenter first put the necessary objects (such as strawberries and pancakes) on the table. For the critical trials in Experiment 1, the experimenter then put sugar on the pancakes (ca. 1500 ms) and subsequently participants listened to German versions of “The experimenter sugars soon the strawberries” or “The experimenter sugared recently the pancakes” (1a and 1b, see **Table 1**). For the critical trials in Experiment 2, the experimenter performed a further action (e.g., sugaring the strawberries), after sentence presentation such that people always saw one action before, and one action after sentence presentation for the critical trials. The experiment lasted approximately 30 min. After the experiment, participants were debriefed.

### 2.4. ANALYSIS

#### 2.4.1. Eye-tracking data

For the coding of participants’ eye gaze during the experimental trials, a period of interest was defined, starting from the onset of the verb until the offset of the post-verbal NP (NP2). The onsets of the critical words in the sentence (verb, adverb, NP2) were marked in the video files using the annotation software ELAN (a tool developed at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and downloadable at <http://www.lat-mpi.eu/tools/elan>, see also Sloetjes and Wittenburg, 2008). In the videos, participants’ gaze during the trial appeared as a red circle. The duration of each frame in the videos

was 40 ms. For the period of interest, participants’ fixations were manually coded frame-by-frame as to which region of the scene was fixated in that particular frame. Three regions were defined: the recent target object, the future target object and “other” (i.e., other parts of the scene, for example, the experimenter or the background).

The measure of interest for the purpose of our study is fixations to the recent and future target objects as the sentence unfolds. Using the frame-by-frame gaze data, we first computed gaze probabilities to the two targets in each of the 40 ms time frames. 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 next computed mean log gaze probability ratios for the recent relative to the future target  $\ln(P(\text{recent target})/P(\text{future target}))$ . This measure, which expresses the bias of inspecting the recent relative to the future target, does not violate the linear independence assumption (e.g., Arai et al., 2007; Carminati et al., 2008). In this measure, 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.

For the inferential analyses we defined the following three time windows: the verb region (from verb onset until adverb onset,  $M = 1148$  ms); the adverb region (from adverb onset until the offset of the adverb,  $M = 1332$  ms) and the NP2 region (from NP2 onset until NP2 offset,  $M = 710$  ms). We aggregated mean log gaze probabilities ratios  $\ln(P(\text{recent target})/P(\text{future target}))$  over each of the three time regions of interest. A further advantage of using log-ratios (in addition to the independence assumption) is that they yield data distributions that are more suitable for parametric testing (standard probabilities often imply a violation of the homogeneity of variance assumption because they have a limited range from 0 to 1; in contrast, log-ratios can take values between minus infinite and plus infinite, which is what is required for parametric testing).

We fitted linear mixed effect (LME) models to the log probability ratios for each of the time regions, using the R-software (version 2.2.0; CRAN project; R Development Core Team, 2008)<sup>1</sup>. Separate models were fitted on log-ratios averaged over participants and items respectively (Barr, 2008). In all models, the predicted outcome was the log ratio of fixations to the recent target relative to the future target and the fixed effect predictor was condition (future vs. recent). To minimize collinearity, we used effect coding by transforming the fixed effect into a numerical value and centering it so as to have a mean of zero and a range of 1 (Baayen, 2008). Effect coding has the further advantage of allowing the coefficients of the regression to be interpreted as the main effects in a standard ANOVA (Barr, 2008). Furthermore, with this coding the intercept represents the estimate of the grand mean; therefore, applied to our particular data, a significant intercept would indicate that the mean log gaze probability ratio  $\ln(P(\text{recent target})/P(\text{future target}))$  is significantly different from zero. In turn, this would indicate that there is a significant bias toward

<sup>1</sup>Due to a sparse frequencies in the design table we could not rely on the hierarchical log-linear analyses (Scheepers, 2003; Knoeferle and Crocker, 2007).

looking at one object relative to the other, whether or not a significant effect of condition is also present (recall that a log ratio of zero would indicate that there is no such bias). For each analysis, two models were fitted, one including only the random intercept (i.e., allowing the intercept to vary across participants and items respectively) and another including both the random intercept and the random slope (i.e., allowing also the slope of the fixed predictor to vary across the random variables). These models were then evaluated using a log-likelihood ratio test (Baayen, 2008, p. 276) and the more complex model was retained only if it fitted the data significantly better than the simpler one (indicated in Table 3 with §). A coefficient was considered to be significant at  $\alpha = 0.05$  when the absolute value of  $t$  was greater than 2 (Baayen, 2008)<sup>2</sup>.

#### 2.4.2. Corpus data

For the corpus study we looked at five different corpora: the Europa Parliament Corpus (Koehn, 2005), the German Reference Corpus (COSMAS II, Kupietz et al., 2010), deWac (Baroni et al., 2009), Google, and DLex (<http://www.dlexdb.de>; Heister et al., 2011). We report two different analyses. (1) For our recent condition, we searched for the exact verb forms in the simple past and present perfect to get an estimate of how often people encounter a verb form referring to the past; for the future condition we searched for verb forms in the present tense. (2) We did a frequency count of the temporal adverbs in the two conditions (recent condition: *soeben*, “just now”; *unlängst*, “not long since”; *kürzlich*, *vorhin*, “a little while ago”; future condition: *sogleich*, “presently”; *nachher*, “subsequently”; *demnächst*, “soon”; *baldigst*, “as soon as possible”). A

third analysis in which we searched for the exact verb and adverb sequences of our items had to be abandoned due to data sparseness. We obtained frequencies of the verbs and adverbs for each item and normalized these frequencies for each corpus using the number of words in the respective corpus<sup>3</sup>. Since this resulted in small numbers, we multiplied each thus-obtained frequency by 1,000,000 to facilitate interpretation. We present descriptive frequencies of the verb forms and of the adverbs averaged across the individual items (Table 4). To ascertain whether there were reliable differences in the frequency scores for our items across the five corpora, we computed the average frequency scores across the five corpora by items (i.e., the 12 verbs used in our study and the 4 temporal adverbs for each condition). We provide the 95 percent confidence interval of the average difference scores for our verbs and adverbs in each of the two conditions (past minus present/future condition for the normalized, multiplied, and averaged scores, Table 4).

### 3. RESULTS

We first present the results of the two eye-tracking studies and subsequently the results of the corpus study. For the eye-tracking data, Figures 2 and 3 plot the mean log gaze probability ratios computed using the original 40 ms frame data, for the period from verb onset to NP2 offset, for Experiments 1 and 2 respectively. Descriptively, these two graphs reveal an overall preference for looking at the recent target relative to the future target throughout the verb and adverb, shown by the fact that during most of this period the log ratio remains well above zero (indicating that the recent target receives more looks than the future target). As participants hear the second noun, they begin to shift gaze to the future target (the referent of “strawberries”) in the future more than in the recent

<sup>2</sup>In choosing to run LME models on data aggregated up to the participant and item level separately, we follow the second approach outlined in Barr (2008) for analyzing visual-world eye-tracking data. It should be noted that this approach is essentially equivalent to running separate repeated measures mixed-design ANOVAs with participants and items as random effects.

<sup>3</sup>The only exception was the Google corpus for which we set the size to 1 since its exact size was unknown.

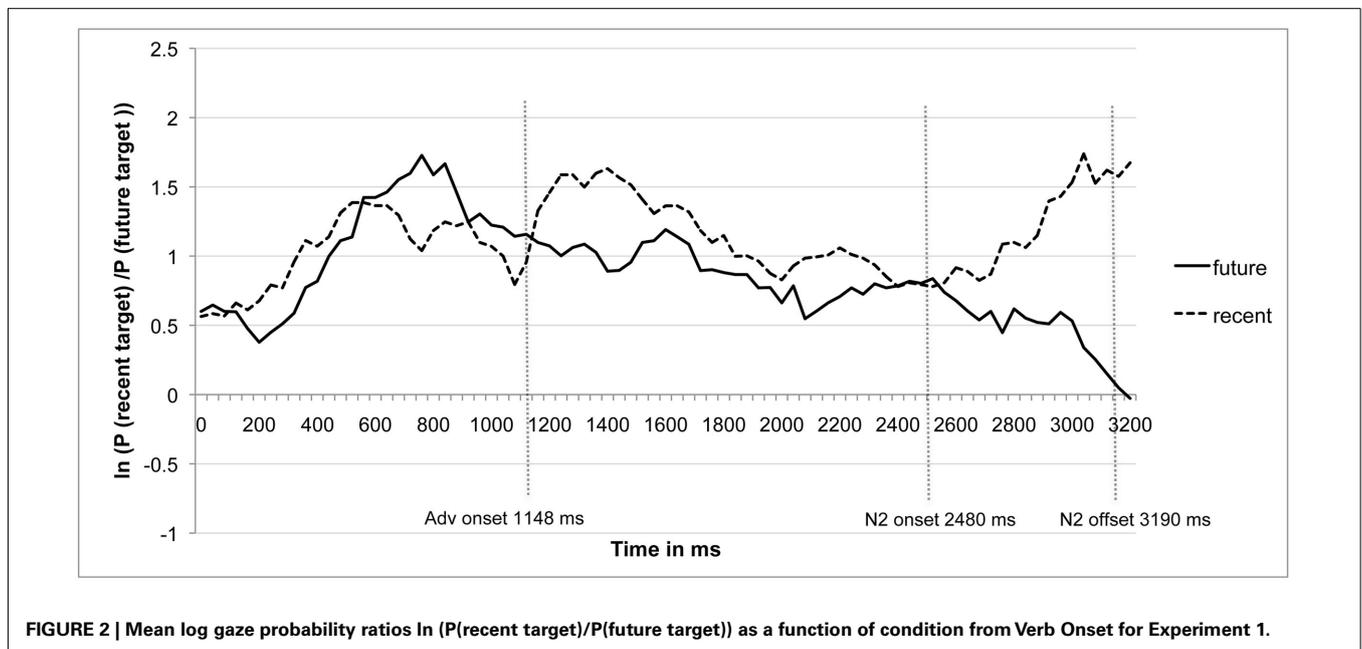


FIGURE 2 | Mean log gaze probability ratios  $\ln(P(\text{recent target})/P(\text{future target}))$  as a function of condition from Verb Onset for Experiment 1.



**FIGURE 3 |** Mean log gaze probability ratios  $\ln(P(\text{recent target})/P(\text{future target}))$  as a function of condition from Verb Onset for Experiment 2.

condition; the recent target (the referent of “pancakes”) is fixated more in the recent condition than in the future condition.

This descriptive pattern was corroborated by the per-region descriptive (Table 2) and inferential (Table 3) analyses. Table 2 shows the mean log gaze probability ratios (participants’ means) for the three time regions of interest as a function of condition, and Table 3 summarizes the results of the corresponding LME analyses. As one can see from the means in Table 2, there is a general inspection bias in favor of the recent target over the future target object, which we noted in the time course graphs (see Figures 2 and 3). Statistical analyses confirmed that this bias was reliable across all three regions in both experiments (i.e., the intercept was significantly different from zero). The positive coefficient for the intercept in Table 3 indicates that people look more at the recent than future target throughout the sentence.

In the verb region for both experiments, the visual preference for the recent target was not modulated by whether the verb was in the present (future condition) or in the past (recent condition), as evidenced by the absence of a significant main effect of condition. However, as participants incrementally processed the remainder of the sentence, the main effect of tense (future vs. recent) becomes reliable, in the final NP2 region in Experiment 1, and in the adverb and NP2 regions in Experiment 2. In the recent condition, this effect is driven by an increase in the log ratio (i.e., looks to the recent target increase and those to the future target decrease), while in the future condition there is a corresponding decrease in the log ratio (i.e., looks to the recent target decrease and those to the future target increase, see Table 2). Concurrent analyses on the same log-ratio measures using mixed-design ANOVAs with participants and items as random effects yielded results in agreement with the LME analyses (see footnote 2).

In the above LME analyses the (positive) grand mean intercept was significantly different from zero, indicating a visual bias toward the recent over the future target averaged over the two

**Table 2 |** Mean log gaze probability ratios  $\ln(P(\text{recent target})/P(\text{future target}))$  as a function of condition and time region for Experiment 1 and 2.

Time region	Future condition (present tense verb and future adverb)	Recent condition (past tense verb and adverb)
<b>EXPERIMENT 1</b>		
Verb	1.52 (0.23)	1.51 (0.23)
Adverb	1.16 (0.26)	1.64 (0.36)
NP2	0.35 (0.26)	1.78 (0.28)
<b>EXPERIMENT 2</b>		
Verb	1.43 (0.26)	1.58 (0.22)
Adverb	1.38 (0.25)	2.23 (0.28)
NP2	0.15 (0.20)	2.10 (0.35)

Standard errors are in parentheses.

conditions. To determine the extent to which this visual bias is present in the two separate conditions, particularly in the future condition, we conducted one-sample two-tailed *t*-tests on the log-ratios of participants and items respectively. These tests, adjusted for two comparisons using the Bonferroni method (new *alpha* level:  $0.05/2 = 0.025$ ), were aimed at ascertaining whether the log-ratio means for each condition are significantly different from zero. With regard to the future condition in Experiment 1, the *t*-tests were significant in both the verb and adverb region (all  $ps < 0.001$ ), but not in the NP2 region ( $p1 = 0.19, p2 = 0.16$ ). This pattern of results was replicated for the future condition of Experiment 2 (verb and adverb region all  $ps < 0.001$ ; NP2 region:  $p1 = 0.47, p2 = 0.79$ ), suggesting that the 50/50 manipulation of Experiment 2 was not able to override the visual preference for the recent object found in Experiment 1 in the verb and adverb region. As expected, the *t*-tests in the recent-event condition achieved

significance for all of the analysis regions in both Experiment 1 and 2 (all  $ps < 0.001$ ).

Table 4 shows the results from the corpus study. It displays the normalized verb and adverb frequencies for the future compared

with recent condition. The difference scores (past minus present tense) illustrate that present tense verb forms are descriptively somewhat more frequent than past tense verbs in four (European Parliament, Cosmas II, deWac, and Google) out of five of the analyzed corpora. The table also presents the normalized frequencies for the adverbs which show that the future tense adverbs are more frequent than the past tense adverbs in three (deWac, Google, and DLex) of the five analyzed corpora.

These descriptive trends, however, were not confirmed by the confidence intervals for the difference scores (past minus present tense verbs/adverbs). With the exception of the European Parliament corpus for the adverb counts, the confidence intervals for all of the corpora contained zero, suggesting that the underlying means do not differ reliably. Overall thus, past tense verbs and adverbs in our sentence stimuli do not appear to be more frequent than present tense verb forms and adverbs indicating the near future.

#### 4. DISCUSSION

Two eye-tracking studies assessed whether the frequency with which participants saw recent (vs. future) everyday events within the experiment can eliminate a previously observed preference to inspect recent-event targets more than future event targets after hearing a sentence beginning that was compatible with either event. In NP1-V-ADV-NP2 sentences the verb was referentially ambiguous between a recent action (and its associated target) and an equally plausible future action (and its different target object). When participants saw the experimenter perform only one action per trial, prior to presentation of the spoken sentence (Experiment 1), they more often inspected the target of that recent action than the target of the future event during and shortly after the verb. This confirmed that the time course and qualitative gaze pattern from a clipart eye-tracking experiment (Knoeferle and Crocker, 2007, Experiment 3) extend to real-world actions. The recent-event preference persisted even when participants saw the experimenter perform equally many actions prior to

**Table 3 | Linear mixed effect model results for Experiments 1 and 2 by time region.**

Time region	Coefficient participants	Items	t-Value participants	Items
<b>EXPERIMENT 1</b>				
Verb				
Intercept	1.51	1.14	9.20*	8.00*
Cond	-0.02	-0.11	-0.15	-0.78
Adverb				
Intercept	1.40	1.04	5.54*	11.67*
Cond	0.24	0.10	1.31	1.19
NP2				
Intercept	1.06	0.94§	5.14*	5.83§*
Cond	0.72	0.50	4.11*	3.01§*
<b>EXPERIMENT 2</b>				
Verb				
Intercept	1.50	1.30	7.69*	12.01*
Cond	0.07	0.07	0.52	0.98
Adverb				
Intercept	1.80	1.55§	8.30*	7.53§*
Cond	0.43	0.53§	2.74*	2.65§*
NP2				
Intercept	1.12§	0.98§	4.95§*	4.54§*
Cond	0.98§	0.94§	5.56§*	3.83§*

\*The effect is significant at  $\alpha = 0.05$  (using the  $|t| > 2$  criterion).

§These values refer to the model that has both random intercepts and random slopes; all other values are in respect of models with only random intercepts.

**Table 4 | Normalized frequency counts for the verb forms and adverbs in our materials averaged across the items.**

	European Parliament (25–30M)	Cosmas II (2000M)	deWac (1411M)	Google set to 1	DLex (100M)
Past tense verb forms	0.091	3.787	0.032	37175.20	8.034
Present tense verb forms	1.123	5.439	0.034	105627.80	3.498
Verb difference scores	-1.032	-1.652	-0.002	-68452.6	4.536
lower/upper 95 CI of the difference scores	-2.779/0.715	-5.843/2.541	-0.0171/0.0130	-247001.4/110096.2	-1.750/10.822
Adverbs indicating the past	58.879	27.774	0.183	417298.0	18.680
Adverbs indicating the future	11.537	17.012	0.184	421338.4	23.805
Adverb difference scores	47.343	10.762	-0.001	-4040.404	-5.125
lower/upper 95 CI of the difference scores	1.126/93.559	-20.667/42.191	-0.239/0.237	-528146.7/520065.9	-45.065/34.815

“Past tense verb forms” and “present tense verb forms” indicate the averaged and normalized frequencies for the recent and future conditions respectively. “Adverbs indicating the past” and “adverbs indicating the future” present the averaged and normalized frequency averages across the adverbs used in the recent and future conditions. “Verb difference scores” and “Adverb difference scores” present the results for subtracting the scores for verbs/adverbs in the future from those for verbs/adverbs in the recent condition. Negative difference scores indicate lower frequencies for the past than present tense verbs and the adverbs. For each corpus we show the number of tokens in millions (M) in brackets. For the verb and adverb difference scores we list first the lower and then the upper 95 percent confidence interval.

versus after sentence presentation (i.e., recent versus future actions respectively) in Experiment 2.

Overall, the data provide good evidence that people prefer to ground their expectations and visual attention during incremental language understanding more through directing their attention at the target of a recent event than at the target of another, equally plausible, future event. We examined this recent-event preference under two frequency distributions of recent relative to future events (i.e., when there was a frequency bias toward recent events in Experiment 1 and when recent and future events occurred equally often in Experiment 2). Together, these two frequency manipulations permit us to tease apart two competing accounts of how contextual information is used to inform expectations during language comprehension: while both the Coordinated Interplay Account (CIA, Knoeferle and Crocker, 2007) and a short-term frequency instantiation of cue reliability in the account by Altmann and Mirković would have predicted a reliance on recent events time-locked to the verb in Experiment 1, their predictions differ for Experiment 2. Consider their predictions for Experiment 1: The CIA incorporates a reference-first mechanism such that comprehenders upon interpreting a word and all else being equal, first look to ground it and find an appropriate referent. Upon hearing a verb, people should thus engage in a search for a suitable referent (visually by interrogating the scene, or by focusing attention on relevant representations in working memory). A short-term frequency instantiation of the account by Altmann and Mirković also predicts a rapid and preferred reliance on recent depicted events in Experiment 1 but for a different reason – because these events are more predictive of what will be mentioned next (as instantiated via a 100:0 frequency bias toward recent events).

When people saw a 50:50 distribution of recent versus future events in Experiment 2, the predictions made by these two accounts diverge. The CIA would still predict a recent-event preference based on its reference-first mechanism. In contrast, a short-term frequency instantiation of cue reliability would no longer predict a preference to inspect the recent-event target more than the future event target since neither of these two information sources is more predictive of which object will be mentioned next or of which action the verb refers to. Both events and verb/adverb forms are equally frequent within the experiment. Thus having seen one action, the ensuing sentence could 50:50 refer to that recent action vs. an equally plausible future action. The findings from Experiment 2 thus provide support against a purely frequency-based account of cue predictiveness in visually situated utterance comprehension. Apparently short-term, within-experiment perceptual and communicative experience that could immediately have informed comprehender's expectations, did not eliminate the preference to inspect the recent-event target during language comprehension.

As mentioned in the introduction, an alternative possibility is that the past tense verbs and adverbs that we used may be more frequent in long-term experience than their present tense counterparts, and that such a long-term frequency bias could guide visual attention to objects. If such a bias exists we may assume that it can rapidly guide attention, since we know that long-term word frequency has rapid effects on language processing and visual

attention in comprehension tasks during reading (e.g., Rayner and Raney, 1996), as well as during spoken language comprehension in visual contexts. For the latter situation, Dahan et al. (2001) found that people fixated objects with frequent (vs. relatively more infrequent) names faster. However, we can be relatively certain that the recent events preference indexed via visual attention that we observed in both experiments during the verb is not driven by the long-term frequency of occurrence of these words since there was no reliable frequency difference between verbs and adverbs in four out of five examined corpora.

The absence of immediate short-term frequency effects is somewhat surprising in light of existing evidence showing that short-term frequencies can affect a range of cognitive processes, among them action execution (e.g., Chapman et al., 2010), language acquisition (e.g., Saffran et al., 1996; Saffran, 2003), language production (Kaschak et al., 2006; Haskell et al., 2010), sentence reading (Wells et al., 2009), and visual perception (e.g., Chun and Jiang, 1999). And yet, participants in the present experiments were not immediately (during the verb) sensitive to the within-experiment frequency distribution of the recent compared with the future event. Had they been immediately sensitive, we should have seen no difference in target object inspection during the verb in Experiment 2. This is not to say that the 50:50 frequency manipulation in Experiment 2 (relative to Experiment 1) did not modulate visual attention. Indeed, effects of tense in Experiment 2 occurred earlier (during the adverb) than in Experiment 1 (during the post-adverb noun phrase). This confirms that our frequency manipulation was effective, in line with previously observed effects of short-term experience on cognitive processes. A short-term frequency account can accommodate the earlier tense effects in Experiment 2 compared with Experiment 1. By contrast, it cannot accommodate the visual preference for the recent-event target during the verb and adverb in the future condition in Experiment 2.

The Coordinated Interplay Account accommodates this latter gaze pattern by postulating that people prefer to first ground the verb in the recent action, and in the absence of the action they do so by inspecting the target object upon which they had previously seen the action performed. Another (speculative) possibility is that the order in which we experience events and hear them talked about affects our reliance on them during comprehension. Seeing an event and hearing it subsequently talked about as part of our experience, may anchor that event in a different way in our (working) memory compared to predicting an event that then happens. To the extent that this holds, the reported findings contribute toward delineating the role of expectation-based processes in language and cognition. They fit well with other findings that have shown older (vs. younger) adults engage less in predictive processing (Federmeier et al., 2002; Federmeier, 2007), as do high (vs. low) literates (e.g., Huettig et al., 2011b). In the present task, participants were instructed to pay attention to both the visual context and to language. When people had seen an action, they likely kept that action in their working memory. It is possible that working memory representations of the recently seen action increased visual attention to associated objects. Such a view would appear compatible with findings that suggest visual orienting can be guided by the contents of working memory in memory tasks (e.g., Spivey and Geng, 2001), and in

visual search tasks (even when they are not relevant for the ongoing search task, e.g., Olivers et al., 2006). To the extent that these findings extend to language paradigms, they underscore the role of working memory representations in language processing (see also Altmann, 2004; Knoeferle and Crocker, 2007; Huettig et al., 2011a).

This position is compatible with the Coordinated Interplay Account to the extent that the verb representation mediates the retrieval of working memory representations of an action. The result of verb-mediated referential processes is that (visual) attention goes preferentially to the location and target associated with a recent action (vs. anticipating the target of a future event). Future studies will examine role of working memory in the present findings by further increasing the frequency of future events in the

experiment and by means of a post-experiment memory test on the recent versus future actions. While it's not entirely clear yet why we observed the recent-event preference in the absence of frequency biases, it is clear that simple, short-term event experience cannot accommodate these findings.

## ACKNOWLEDGMENTS

This research was supported by the Cognitive Interaction Technology Excellence Cluster (funded by the German Research Council) and the CRC 673 "Alignment in Communication," both at Bielefeld University, Germany. We thank Eva Mende and Linda Krull for assistance with coding of the video data, Patrick Bremehr for recording the sound stimuli, and Christian Pietsch for advice on the corpus analyses.

## REFERENCES

- Aborn, M., Rubenstein, H., and Sterling, T. D. (1959). Sources of contextual constraint upon words in sentences. *J. Exp. Psychol.* 57, 171–180.
- Aglioti, S. M., Cesari, P., Romani, M., and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.* 11, 1109–1116.
- Altmann, G. T. M. (2004). Language-mediated eye-movements in the absence of a visual world: the "blank screen paradigm." *Cognition* 93, B79–B87.
- Altmann, G. T. M., and Dienes, Z. (1999). Rule learning by seven-month-old infants and neural networks. *Science* 284, 875.
- Altmann, G. T. M., and Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73, 247–264.
- Altmann, G. T. M., and Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cogn. Sci.* 33, 583–609.
- Arai, M., van Gompel, R., and Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cogn. Psychol.* 54, 218–250.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Pipenbrock, R., and Gulikers, L. (1995). *The Celex Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium.
- Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). The waCky wide web: a collection of very large linguistically processed web-crawled corpora. *Lang. Resour. Eval.* 43, 209–226.
- Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *J. Mem. Lang.* 59, 457–474.
- Berkum, J. V., Brown, C. M., Zwitserlood, P., Koojiman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 443–467.
- Berry, M. J., Brivanlou, I. H., Jordan, T. A., and Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* 398, 334–338.
- Carminati, M. N., Gompel, R. P. G. van, Scheepers, C., and Arai, M. (2008). Syntactic priming in comprehension: the role of argument order and animacy. *J. Exp. Psychol. Lang. Mem. Cogn.* 34, 1098–1110.
- Chambers, C. G., Tanenhaus, M. K., and Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 687–696.
- Chapman, C. S., Gallivan, J. P., Wood, D. K., and Milne, J. L. (2010). Reaching for the unknown: Multiple target encoding and real-time decision-making in a rapid reach task. *Cognition* 116, 168–176.
- Chun, M. M., and Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychol. Sci.* 10, 360–365.
- Dahan, D., Magnuson, J. S., and Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: evidence from eye-movements. *Cogn. Psychol.* 42, 317–367.
- DeLong, K., Urbach, T. P., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat. Neurosci.* 8, 1117–1121.
- Elman, J. L. (1990). Finding structure in time. *Cogn. Sci.* 14, 179–211.
- Federmeier, K. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505.
- Federmeier, K., McLennan, D. B., E. DeOchoa, E., and Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. *Psychophysiology* 39, 133–146.
- Fischler, I., and Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *J. Verbal Learn. Verbal Behav.* 18, 1–20.
- Hale, J. (2003). The information conveyed by words in sentences. *J. Psycholinguist. Res.* 32, 101–122.
- Haskell, T. R., Thornton, R., and MacDonald, M. C. (2010). Experience and grammatical agreement: statistical learning shapes number agreement production. *Cognition* 114, 151–164.
- Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., and Kliegl, R. (2011). dlexdb – eine lexikalische datenbank für die psychologische und linguistische forschung. *Psychol. Rundsch.* 62, 10–20.
- Huettig, F., Olivers, N., and Hartsuiker, R. J. (2011a). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychol. (Amst.)* 137, 138–150.
- Huettig, F., Singh, N., Singh, S., and Mishra, R. K. (2011b). "Language-mediated prediction is related to reading ability and formal literacy," in *Proceedings of the 17th Annual Conference on Architectures and Mechanisms for Language Processing*, Paris.
- Kaiser, E., and Trueswell, J. C. (2005). The role of discourse context in the processing of a flexible word-order language. *Cognition* 94, 113–147.
- Kamide, Y., Altmann, G. T. M., and Haywood, S. (2003a). The time course of prediction in incremental sentence processing. *J. Mem. Lang.* 49, 133–156.
- Kamide, Y., Scheepers, C., and Altmann, G. T. M. (2003b). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *J. Psycholinguist. Res.* 32, 37–55.
- Kaschak, M. P., Loney, R. A., and Borreggine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition* 99, B73–B82.
- Knoeferle, P., and Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: evidence from eye-movements. *J. Mem. Lang.* 75, 519–543.
- Knoeferle, P., Crocker, M. W., Scheepers, C., and Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition* 95, 95–127.
- Koehn, P. (2005). "Europarl: a parallel corpus for statistical machine translation," in *Proceedings of MT Summit X*, Phuket.
- Kupietz, M., Belica, C., Keibel, H., and Witt, A. (2010). "The German reference corpus DeReKo: a primordial sample for linguistic research," in *Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC 2010)*, eds N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Ojiki, S. Piperidis, M. Rosner, and D. Tapias (Valletta: European Language Resources Association), 1848–1854.
- Levy, R. (2008). Expectations-based syntactic comprehension. *Cognition* 106, 1126–1177.
- Miall, R. C., Weir, D. J., Wolpert, D. M., and Stein, J. F. (1993). Is the cerebellum a smith predictor? *J. Mot. Behav.* 25, 203–216.

- Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* 370, 256–257.
- Olivers, C. N., Humphreys, G. W., and Braithwaite, J. J. (2006). The preview search task: evidence for visual marking. *Vis. cogn.* 14, 716–735.
- Pickering, M., and Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends Cogn. Sci. (Regul. Ed.)* 11, 105–110.
- Rayner, K., and Raney, G. E. (1996). Eye-movement control in reading and visual search: effects of word frequency. *Psychon. Bull. Rev.* 3, 245–248.
- R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Saffran, J. (2003). Statistical language learning: mechanisms and constraints. *Curr. Dir. Psychol. Sci.* 12, 110–114.
- Saffran, J., Aslin, R. N., and Newport, E. (1996). Statistical learning by 8-month old infants. *Science* 274, 1926–1928.
- Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition* 89, 179–205.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., and Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition* 71, 109–148.
- Sloetjes, H., and Wittenburg, P. (2008). “Annotation by category – ELAN and ISO DCR,” in *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech.
- Spivey, M. J., and Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychol. Res.* 65, 235–241.
- Tulving, E., and Gold, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *J. Exp. Psychol.* 66, 319–327.
- Weber, A., Grice, M., and Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition* 99, B63–B72.
- Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. C., and MacDonald, M. C. (2009). Experience and sentence processing: statistical learning and relative clause comprehension. *Cognition* 58, 250–271.
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). Event perception: a mind/brain perspective. *Psychol. Bull.* 133, 273–293.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 August 2011; accepted: 28 November 2011; published online: 23 December 2011.

Citation: Knoeferle P, Carminati MN, Abashidze D and Essig K (2011) Preferential inspection of recent real-world events over future events: evidence from eye tracking during spoken sentence comprehension. *Front. Psychology* 2:376. doi: 10.3389/fpsyg.2011.00376

This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Knoeferle, Carminati, Abashidze and Essig. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

## APPENDIX

### SENTENCE STIMULI

1. Der Versuchsleiter zuckert sogleich/zuckerte soeben die Erdbeeren/Pfannkuchen.
2. Der Versuchsleiter mixt sogleich/mixte soeben den Cocktail/Milchshake.
3. Der Versuchsleiter buttert sogleich/butterte soeben die Brotscheiben/Croissants.
4. Der Versuchsleiter bewässert nachher/bewässerte unlängst die Kresse/Tulpe.
5. Der Versuchsleiter poliert nachher/polierete unlängst die Kerzenständer.
6. Der Versuchsleiter studiert nachher/studierte unlängst den Buchtitel.
7. Der Versuchsleiter öffnet demnächst/öffnete kürzlich die Saftflasche/Schuhkiste.
8. Der Versuchsleiter würzt demnächst/würzte kürzlich die Gurke/Tomate.
9. Der Versuchsleiter salzt demnächst/salzte kürzlich die Zucchini/Aubergine.
10. Der Versuchsleiter schlürft baldigst/schlürfte vorhin die Limonade/Apfelschorle.
11. Der Versuchsleiter schüttelt baldigst/schüttelte vorhin die Sojamilch/Sprühsahne.
12. Der Versuchsleiter verrührt baldigst/verrührte vorhin den Milchkaffee/Kräutertee.