How robust is the recent event preference?

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 Cluster, Department of Linguistics
Bielefeld University, Germany

Abstract

Previous eye tracking findings show that people preferentially direct their attention to the target of a recently depicted event compared with the target of a possible future event during the comprehension of a spoken sentence relating to the recent or future event (e.g., Abashidze, Knoeferle, Carminati, & Essig, 2011; Knoeferle & Crocker, 2007). This gaze pattern emerged even when the frequency of occurrence of future and recent events did not differ within the experiment, Knoeferle, Carminati, Abashidze, & Essig, 2011, Experiment 2). To further test the robustness of the recent event preference, the current studies introduced a frequency bias in favor of the future over the recent event (Experiment 1: 88% future vs. 12% past events in combination with future and past sentences; Experiment 2: 75% future vs. 25% past event). We found that increasing the frequency of the future event did result in earlier fixations to the target of the future event than previously observed (in Experiment 2 of Knoeferle, Carminati, Abashidze, & Essig, 2011). However, in the current studies we essentially replicated the same overall preference to look at the target of the recent event throughout sentence presentation. A memory test supported these results. Thus, within-experiment frequency appears to modulate the recent event preference to some extent, but cannot override it. We propose that an epistemic bias of the human mind favors assertions about past events over future ones.

Keywords: Visual context effects, Sentence comprehension, Frequency manipulation, Eye tracking

Introduction

Our everyday environment is full of visual objects and events and observers arguably focus their attention on the events that are relevant for a given situation, ignoring others. Indeed, many previous findings have shown that participants are sensitive to the visual context and that they use it for language comprehension. When participants listen to an utterance about information in a visual scene, they inspect the object in that scene in close temporal coordination to their mention (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) and they can even anticipate an object before its mention based on verb meaning. For example, in Tanenhaus’ et al. (1995) visual world study, people moved their eyes on a depicted scene while they heard a sentence involving the objects in the scene. In this and many other visual world studies, it has generally been observed that looks to objects are rapidly guided by the object name (e.g., apple narrows the domain of reference to apples) or by the linguistic input more generally (e.g., eat can restrict the range of possible direct objects to edible ones, see Altmann & Kamide, 1999).

In theory, though, such tight synchronization may not always obtain; for example, one might expect the use of linguistic information to lag behind that of visual information with visual cues being prioritized, and vice versa (Knoeferle & Crocker, 2006; 2007; see also Hartsuiker, Huettig & Olivers 2011). The challenge for researchers is then to investigate the factors (e.g., the type of visual context or the frequency of visual and linguistic cues among others) that modulate the timing and priorities in the interaction between linguistic and non-linguistic information. One phenomenon in this interaction is the recent event preference, a tendency for people to preferentially direct their attention to the target of a recently depicted event relative to the target of a possible future event, while they hear a sentence with a future tense meaning describing the future event (Knoeferle & Crocker, 2007).

Following up from the initial study by Knoeferle and Crocker (2007, Experiment 3), which used static clip-art stimuli, the recent event preference has also been investigated using dynamic real-world stimuli. In Knoeferle et al., (2011, Experiment 1), the experimenter was sitting at a table on which two objects were located (e.g. pancakes and strawberries, both of which could be sugared). The experimenter performed an event (i.e., sugaring strawberries) and subsequently participants listened to a past-tense sentence (Der Versuchsleiter zuckerte kürzlich die Erdbeeren ‘The experimenter recently sugared the strawberries’) or to a present tense sentence with a future meaning (Der Versuchsleiter zuckert demnächst die Pfannkuchen ‘The experimenter will soon sugar the pancakes’). As in the clip-art study, for a given trial, the experimenter performed only the recent event (i.e., sugaring the strawberries). The results replicated the findings from the clipart experiment: In the future tense condition, participants preferred to look at the recently-acted upon object during most of the sentence, and a clear preference to look at the future object surfaced only during the last word (NP2, when the future target was actually named). This
pattern was confirmed by the statistical results, which revealed a significant tense effect in the NP2 region, but not in the verb and adverb regions. Furthermore and importantly, in all sentence regions, including the NP2 region, an overall preference to look at the recent object emerged, which was independent of sentence tense.

**Frequency of linguistic and visual cues as a modulating factor**

One important concern regarding Knoeferle and Crocker (2007, Experiment 3) and Knoeferle et al. (2011, Experiment 1) is that the recent event preference may have arisen because participants in those two experiments only ever saw the past event performed. Because the future event was never shown in those studies, the procedure may have created a within-experiment frequency bias towards relying more on recently depicted than on equally plausible future events. This frequency imbalance was addressed in the next study (Knoeferle et al. 2011, Experiment 2) in which participants were also shown future events. In this experiment the experimenter post-sentence performed the action referring to the (previously heard) future tense sentence. Thus, during the experiment participants were shown recent and future events equally often (50:50 frequency). If the frequency with which people experience recent and future events modulates the recent event preference, we expect an earlier effect of tense (i.e., looks to the recent and future object should start to diverge earlier as a function of tense). Indeed in this experiment, the tense effect achieved significance in the adverb region (compared to the NP2 region in Knoeferle et al. 2011, Experiment 1). However, we still replicated the significant overall preference to look at the recent object independent of tense up to the last region. Thus, although the 50:50 frequency manipulations produced an earlier effect of tense, it does not appear to be strong enough to eliminate the recent event preference.

The two current experiments were designed to test the resilience of the recent event preference by introducing frequency manipulations that create a very strong bias towards the future relative to the recent event. In other words during the experiments, participants saw future events performed more often than recent ones (of which more later). The rationale of these manipulations was based on findings that within-experiment frequency biases and short-term experience could modulate sentence processing. In fact, in recent years it has become increasingly clear that language comprehension and also other cognitive and motor processes are sensitive to statistical regularities. For example, 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, Gallivan, Wood, Milne, Culham & Goodale 2010). In language, statistical regularities can be exploited by children as young as 8 months for segmenting words in fluent speech (Saffran, Aslin & Newport, 1996). Short-term language experience also modulates language production (Kaschak, Loney, & Borreggine, 2006. Haskell; Thornton & MacDonald, 2010; see also Britt, Mirman, Kornilov, & Magnuson 2014) and sentence reading (Wells, Christiansen, Race, Acheson & MacDonald, 2009).

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 visually attend to during comprehension.

In sum, the goal of these two studies was to see whether seeing more future events and hearing more future tense sentences could make the future tense and event more accessible. Increasing the frequency of future events against that of past events might produce stronger and earlier effects in the future tense condition than in the previous studies. One change that we might see is that the recent-event preference decreases or disappears completely. Alternatively, it could be that events that we have seen and then heard mentioned are more prominent in working memory than future actions even when the latter are much more frequent. It is possible that these working memory representation increase visual attention to the target of the recent event, which might lead humans to inspect a recent action target earlier and more often than the target of an equally plausible future action. In the current experiments we also gave participants a post-experiment memory test. If recent (vs. future) events are anchored more firmly first in working and then in short-term memory, participants should be better at recalling the targets of recent (vs. future) events.

**Participants**

Thirty-two German native speakers (aged 19 to 32) participated in each experiment. Participants (all students of Bielefeld University, Germany) were each paid 6 Euros for their participation. All had normal or corrected-to-normal vision, were unaware of the purpose of the experiment, and gave informed consent.

**Materials and design**

Twenty-four experimental items were used in each of the two experiments. Half of these items (12) were the same as those used in Knoeferle et al., 2011, Experiment 2, and an additional 12 were constructed using similar criteria. Each item consisted of two everyday objects (e.g., cucumbers and tomatoes, see Figure 1) and four sentences (see Table 1). All critical sentences had the structure NP-V-ADV-NP and a male native German speaker recorded them. The 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 1, 1a-a’). In the other condition, the verb was in the simple past, and the following time adverb (kürzlich, ‘recently’) also indicated the past (Table 1, 1b-b’). Only German regular verbs were used in the critical sentences. 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 both a past and future action. In turn, this ensured that visual characteristics of any given post-verbal target object contributed equally to each critical condition. Importantly, the two objects mentioned in an item could be equally plausible targets of the action expressed by the verb (e.g. both cucumbers and tomatoes can be flavored). The words in a sentence were matched for spoken syllables and lemma frequency within an item (Baayen et al., 1995).

For every item we recorded two videos, each lasting in average of 5015 ms. The scene for both videos always showed a person sitting at a table, and two objects on the table (e.g. cucumbers and tomatoes), one on the left and one on the right, at about equal distance from the person. The first video showed the person performing an action on one object (e.g., flavoring cucumbers, Fig. 1, a; Fig. 1 shows the order in which the videos were presented in a typical critical item trial) and the second showed the person performing the same action on the other object (e.g., flavoring tomatoes, Fig. 1, c). The position of the target objects (right vs. left) was counterbalanced across items. For every item we also created a snapshot (i.e., a static photo, see Fig. 1, b) showing the person in a static position performing no action and looking at the camera. Examples of the videos and the snapshot associated with the experimental sentences in Table 1 are shown in Figure 1 (a-c). The same 24 experimental sentences/videos were used in Experiment 1 and 2.

Additionally, we created a number of filler items. The frequency bias manipulation was achieved by having most filler items associated with future tense sentences (e.g., *Der Versuchsleiter zeichnet in der nächsten Zukunft ein Haus auf dem Notizblock*, literally: ‘The experimenter will draw in the near future a house on the notebook’). This manipulation is explained in detail below (see “Frequency bias manipulation”).

An important difference between Experiments 1 and 2 was that in the filler trials of Experiment 1 the first video was not shown and therefore the trial started with a static image and the future sentence (see e.g. Fig. 1, b) and ended with the video of the ‘future’ action. In this respect, filler trials differed from the experimental trials of Experiment 1, in which participants saw two videos. In Experiment 2 the filler trials were exactly as the experimental trials, i.e., with two videos being shown, one before and one after the sentence.

The experimental and filler items were combined to form 4 lists using a Latin square. Each list contained every critical item in only one condition and all fillers. Before the experiment, lists were pseudo-randomized and each participant saw an individually randomized version of one of the four experimental lists.

Figure 1: Sequence of events of a typical experimental trial

<table>
<thead>
<tr>
<th>Condition &amp; counterbalancing</th>
<th>Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a <strong>FUTURE TENSE</strong></td>
<td><em>Der Versuchsleiter würzt demnächst die Tomaten</em> ‘The experimenter will soon flavor the tomatoes.’</td>
</tr>
<tr>
<td>1a’ <strong>FUTURE TENSE</strong></td>
<td><em>Der Versuchsleiter würzt demnächst die Gurken</em> ‘The experimenter will soon flavor the cucumbers.’</td>
</tr>
<tr>
<td>1b <strong>PAST TENSE</strong></td>
<td><em>Der Versuchsleiter würzte kürzlich die Gurken</em> ‘The experimenter recently flavored the cucumbers.’</td>
</tr>
<tr>
<td>1b’ <strong>PAST TENSE</strong></td>
<td><em>Der Versuchsleiter würzte kürzlich die Tomaten</em> ‘The experimenter recently flavored the tomatoes.’</td>
</tr>
</tbody>
</table>

**Frequency bias manipulation**

As mentioned, there were 24 experimental items in each experiment. The number of filler items was 72 and 40 for Experiment 1 and 2 respectively. Thus in Experiment 1 people saw a total of 96 items (24 critical +72 fillers) and in Experiment 2 a total of 64 items (24 critical + 40 fillers). Half of experimental trials (12) showed the recent action and the other half (12) the past action. All the filler trials of Experiment 1 only showed the future action, so the number of future and recent actions seen by participants in this experiment was 84 and 12 respectively (approx 88% vs. 12% past events and sentences). In Experiment 2, 36 filler trials showed the future and recent actions and only played a future tense sentence, and the remaining 4 showed the recent and future actions and only played a past tense sentence, making a total of 48 future vs. 16 recent actions being seen over the experiment (75% vs. 25%). Thus, in both Experiment 1 and 2 there was an overwhelming bias towards future events.

**Materials for the memory test**

For the memory test 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). The two snapshots associated with each item were combined into one display. Two versions were created in which the respective location of the two pictures was counterbalanced.

Figure 2. An example of a display for the memory test

Procedure

Procedure for eye tracking study

Participants were informed that the experiment consisted of an eye-tracking experiment followed by a short memory test. During the eye-tracking experiment, participants were instructed to look at the computer display and listen carefully to the sentence. They were not given any details of what the memory test would be about. After successful calibration of the eye tracker, the experiment started. In a given experimental trial, the timing of the trial sequence is illustrated in Fig 1. Participants saw a video of a person performing one action before the sentence (e.g., flavoring the cucumbers, see Fig. 1, a) and then the static picture appeared. After 700 ms, the sentence was played out. The static picture remained on the screen until 700 ms after the end of the sentence (see Fig. 1, b). The sentence presented (see, Table 1, 1b or 1a) was either Der Versuchsleiter würzte kürzlich die Gurke ‘The experimenter recently flavored the cucumbers’ or Der Versuchsleiter würzt demnächst die Tomate ‘The experimenter will soon flavor the tomatoes’. After the static picture had disappeared, participants were shown a second video of the person performing the second action (e.g., flavoring the tomatoes, see Fig. 1, c). In the middle of the experiment participants had a short break.

Procedure for memory test

Participants did the memory test after the eye tracking session. They were assigned randomly to four counterbalancing lists and each saw a randomized order of the list. They were shown pictures such as those in Fig. 2, one for each experimental item. Above the picture, a question appeared, which could be in one of two versions: (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 were to give their response with a button press. After the memory tested ended, participants were debriefed. The experiment lasted approximately 45-50 minutes.

Analyses and results (Experiments 1 and 2)

Eye tracking

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), i.e. the end of the sentence. The measure of interest for the purpose of our study is fixations to the recent and future target objects as the sentence unfolds. We first computed gaze probabilities to the two target objects in each successive 20 ms time slots, starting from the onset of the verb until the end of the sentence. 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 (recent target)/P (future target))). 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. We used this measure to plot the time course graphs from verb onset, which are shown in Fig. 3 (a-b). In Fig.3 the dotted lines indicate the recent condition (sentence in the past tense) and the solid lines indicate the future condition (sentence in the future tense).

As can be seen (Fig. 3a-b), in both experiments looks to the future and recent object as a function of tense start to diverge towards the end of the verb region. Mixed effects ANOVAs (by participants and items) on the mean log ratios for each region (Verb, Adverb, NP2) showed a significant effect of tense (all p’s < .05) in the Adverb and NP2 regions in Experiment 1. In Experiment 2, tense was significant (only marginally by participants) also in the Verb region. Thus, the frequency manipulation in favor of the recent event did produce an earlier effect of tense in both experiments. However and importantly, as can be seen from Fig. 3a-b, in both experiments the log ratio in both tense conditions remained above zero until the late Adverb region (specifically until 2100 ms after verb onset in Experiment 1 see Fig 3, a, and 1900 ms after verb onset in Experiment 2, see Fig 3, b), indicating that, until then, the recent target received more looks than the future target even in the future tense condition. In the future tense condition, the log ratio eventually turns negative in both experiments (see Fig 3a-b), indicating a preference for the future over the recent target.

Figure 3 Mean log gaze probability ratios (ln (P(recent target)/P(future target))) as a function of condition from Verb onset for Experiment 1 and 2

a)
Furthermore, the grand mean (i.e., the mean of the two conditions) was positive in all sentence regions in both experiments, showing an overall preference for the recent target irrespective of tense. This overall preference for the recent target throughout the sentence was confirmed by a significant intercept in all the ANOVAs by region for both experiments (significant intercept=grand mean significantly different from zero). Thus, despite the strong frequency bias towards future events, the current experiments replicated the significant overall preference to look at the recent object independent of tense up to the very last sentential region.

**Memory test**

The purpose of the memory test was to assess participants’ memory of the events/actions they had seen in the experimental video sequences during the eye tracking session. In particular, we wanted to see whether actions that had been presented before the sentence (i.e., the recent actions) were remembered better than actions inspected after the sentence had been heard (i.e., the future actions).

We calculated the percentage of correct answers by condition for participants and items separately. The average percentages (by participant) are illustrated in Figure 4a and 4b. In both tests, participants correctly answered, 80% of the questions in Experiment 1 (Fig. 4, a) and 83% of the questions in Experiment 2 (Fig. 4, b). As we can see from the graphs, participants were more accurate in recognizing the recent event targets than the future event target objects in Exp 1 (83% vs. 77%) and in Exp 2 (86% vs. 80%).

In logistic linear mixed effect (LME) analyses for Experiment 1 we found marginal effects of target object ($p =0.06$) and tense ($p=0.09$), but no interaction. In Experiment 2 there were neither main effects nor interactions of target and tense in the LME analyses.

**Discussion**

In the current studies we tested the robustness of the recent event preference by introducing a within-experiment frequency manipulation, which was overwhelmingly in favor of future over recent events. In Experiment 1 the future events seen by participants constituted 88% of the total number of trials in the experiment (vs. 12% for recent events). In Experiment 2 future events made up 75% (vs. 25%) of the trials. In addition, the two experiments differed in the presentation of the filler trials. In a previous study where past and future events were seen by participants equally often in the experiment (Knoeferle et al., 2011) tense effects achieved significance from the Adverb region.

Furthermore, an overall preference emerged of looks to the recent object throughout the sentence, irrespective of tense. The frequency manipulation of the current experiments was motivated by findings showing that within-experiment frequency biases and short-term experience can modulate language processing. As such, we expected that our strong frequency bias might substantially mitigate, eliminate, or even reverse the recent event preference.

This was not the case. Although we did observe a significant tense effect earlier e.g. in the verb region in Experiment 2 than in the experiment with the 50:50 frequency manipulation, we replicated the overall preference for the recent object (up until sentence end) observed in previous experiments. The results of the memory test (in the previous studies memory tests were not administered) also suggest that recent events are remembered better than future ones.

Which cognitive mechanisms underlie the recent event preference? Before attempting to answer this question, two clarifications are in order regarding the surprising and unexpected weak effect of our strong frequency manipulation on the recent event preference. First, the learning of statistical regularities in language has been found to increase in strength over time (e.g., Kaschak & Glenberg, 2004), so it is possible that the frequency manipulation in our study did not have an immediate effect but ‘built up’ over time in the course of the experiment. To assess whether that was the case, we performed post-hoc analyses on our eye tracking data with experimental block as a factor (i.e. first vs. second half of experiment). If
learning of statistical regularities (the frequency of occurrence of a recent vs. future event, in association with a past vs. future tense sentence) takes place, and becomes stronger over time, we should find stronger tense effects in the second half than in the first half of the experiment. However, our analyses showed that the effects were as strong in the first as in the second half.

Second, it is generally assumed that not only short term but also long-term linguistic experience of statistical regularities can modulate language processing (e.g., Gries & Divjak, 2012). Thus, the recent event preference could stem from the fact that in everyday language past tense sentences (in combination with a past tense adverb) are more frequent than present tense sentences in combination with a future adverb. In this respect, Knoeferle et al., (2011) presented evidence from corpus studies showing that this is not the case. Thus, the recent event preference unlikely stems from the representation of long-term linguistic regularities.

We propose that the recent event preference reflects an epistemic bias of the human mind: Assertions about a past event command more attention than assertions about a future event because they are based on stronger evidence about event truth. In fact, while a past event can generally be verified, a future one cannot, at least not until it has actually happened, and until then it remains uncertain if it will happen (McFarlane, 2003; Staub & Clifton, 2011). An alternative explanation is that the recent event preference does not reflect an epistemic bias, but is the result of a verb-noun association that participants might have formed after seeing the first video clip (of the recent event). This would arguably be the most active association at sentence start, so it drives eye movements (this would imply that participants silently named the recent action, and the object acted upon, while they saw it performed). This explanation may contain some truth, but does not account for why participants, when encountering a future tense verb and adverb referring unambiguously to a future event, ‘unnecessarily’ linger on the recent object, and why overall they still prefer to look at the recent object until late in the sentence, even in the presence of strong frequency biases such as ours.

Acknowledgments

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

References


