Universität Bielefeld

Doctoral Thesis

Visual context effects on situated language comprehension: Evidence from eye-tracking

By
Dato Abashidze

A thesis submitted in partial fulfillment of the requirements for the degree of Doktors der Philosophie

at the Faculty of Linguistics and Literature

August 18, 2016

Examination Committee:

Prof. Dr. Pia Knoeferle	*(Humboldt University of Berlin)*
Prof. Dr. Craig Chambers	*(University of Toronto)*
Prof. Dr. Stavros Skopeteas	*(Bielefeld University)*
Dr. Annett Jorschick	*(Bielefeld University)*
Prof. Dr. Petra Wagner	*(Bielefeld University)*

Abstract

The current thesis project aimed to explore further non-linguistic context effects during language comprehension. The research in visually situated language comprehension has been conducted by applying different methods, which has revealed immediate visual-context effects, the effect of linguistic cues (e.g., verb, adverb, prepositional phrase) and non-linguistic cues (e.g., size, color, depicted action) during language processing. By contrast, less attention has been paid to the influence of events on language comprehension. Our everyday situated language refers to events and our understanding of the influence of these events onto linguistic and non-linguistic input needs to be examined.

Previous experiments that examined event influences on sentence reading time found that participants were faster in reading about ongoing than completed events while they read a past progressive or past perfect sentence (Madden & Therriault, 2009). Furthermore, eye-tracking results revealed an effect of a recently seen event on language comprehension (Knoeferle, Carminati, Abashidze, & Essig, 2011). When participants saw an action event before and after a sentence and they listened to a German (NP1-Verb-Adv-NP2) sentence referring to the recent event or to a plausible future event, their eye-movements towards a plausible future event target emerged only after the target object had been named. Most of their attention independent of sentence tense went to the target of the recent action event. This gaze pattern has been dubbed the 'recent event preference'. Thus, these results contrast with other findings which suggest that listeners are able to anticipate the target object of a plausible future event in a scene before that object is mentioned (e.g., Kamide, Scheepers, & Altmann, 2003).

The findings of the recent event influence on language processing have motivated the present thesis project and we tested the following questions: we asked, how strong the recent event influence is when pitted against various other factors such as a frequency bias towards the future event, a situation-immediate cue (the actor's gaze), and an incongruence between the verb and the recent event. In order to

examine these questions we conducted five eye-tracking experiments. Furthermore, in addition to the eye-tracking experiments we carried out memory tests.

Chapter 3 deals with the first manipulation. Experiments 1 and 2 introduced a frequency bias towards future events to test the robustness of the recent event preference. These experiments increased the number of future events and future tense sentences up to 75% in Experiment 1 and up to 88% in Experiment 2. Accordingly, the recent events and past tense sentences have been reduced in frequency within the experiments. The findings of these two experiments revealed early looks towards the future event target in the future tense condition, namely in the adverb region. However, they also replicated an overall bias in looking towards the recent event target.

Chapter 4 provides another manipulation to test this issue. Experiments 3 and 4 have pitted the gaze cue of an actor (to the future target) against the recent event preference. The actor's gaze shift occurred either during the verb in Experiment 3 or at the onset of the verb in Experiment 4. The results showed that the actor's gaze cue sustained an early inspection of the future event target at the end of the verb region. However, in the no gaze condition a similar gaze pattern towards the future event target emerged only in the later adverb region. Surprisingly, similar to Experiments 1 and 2, Experiments 3 and 4 replicated the overall recent event preference.

Since the first four experiments reported in this thesis revealed an overall recent event preference and a very strong bias in the verb region, Experiment 5 (Chapter 5) applied a strong cue against the recent event preference, whereby the recent event never matched the past tense verb. By contrast, the future tense verb and the future action event always matched referentially. Interestingly, the incongruence effects were not as strong as the other manipulations in the verb and adverb region. Although the preferential looks towards the recent event target started to decrease at the end of the verb region, inspections of the future event target occurred only at the second noun onset. As in the first four experiments, we replicated the recent event preference.

In addition, we report the results of offline memory tests. For the first two experiments, these support the analysis of the eye-movement data (the studies that had included a frequency bias towards the future target object). By contrast, for the experiments that introduced actor's gaze as a cue, the memory tests did not

reveal better memory when gaze did (vs. when it did not) bias listeners' attention to the future event target. The memory test in the last experiment that created a mismatch between the recent action and the sentential verb revealed a better performance in the future than in the past tense condition.

In the last chapter of the thesis we discuss the reported results in the context of other relevant findings from previous research. We discuss the results of the thesis experiments and we compare the main findings across the different experimental manipulations. Furthermore, we highlight the robustness of the recent event preference and at the same time we discuss its limits as revealed by our experimental manipulation. Furthermore, we argue that the results from the thesis project provide evidence in line with previous findings of visual context (recent event) influences on language processing. At the end of the chapter, for the conclusion, we summarise the key findings and point out a number of new directions for examining the issue of the recent event preference.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Pia Knoeferle for the continuous support for my thesis and related research, as well as her patience, motivation, and immense knowledge. Her guidance helped me throughout the research and writing of this thesis. In addition to the supervision itself, I am also thankful for the chance to work in a well-equipped and well organized environment at the Cognitive Interaction Technology Excellence Cluster at the University of Bielefeld. I am also very grateful to Stavros Skopeteas, who kindly agreed to review my dissertation.

I also would like to thank the members (and former members) of the Language and Cognition Research Group at the CITEC for creating a positive atmosphere for interesting discussions and feedback during our colloquiums. I wish to thank my friends and my colleagues for motivating coffee breaks. Their comments, criticisms and suggestions were very valuable for my research. My thanks go out to Maria Nella Carminati for her patience and recommendations with regard to data analyses. Special thanks to Georgina Willms with regard to English punctuation and Eva Nunnemann and Marko Dierkes for their indispensable help in the last stage of completing my thesis.

My sincere thanks also goes to Dr. Carol Madden-Lombardi and Prof. Peter Ford Dominey who provided me with the opportunity to join their team as an intern, and who gave access to the laboratory and research facilities. I further want to thank Prof. Craig Chambers and his lab members who supported me and gave me the opportunity to be at the PAL Lab and run my own study. I am grateful to the DFG for providing financial support and to all members of the CITEC Graduate School for providing a friendly and stimulating environment.

Finally, I would like to thank my family: my parents and my brothers and sister for supporting me throughout writing this thesis and my life in general. This thesis is also dedicated to many other unmentioned friends and colleagues.

Contents

List of Figures

List of Tables

Abbreviations

CIA	**C**oordinated **I**nterplay **A**ccount
LME	**L**inear **M**ixed **E**ffect
REF	**R**ecent-**E**vent **P**refence
VWP	**V**isual **W**orld **P**aradigm

Chapter 1

Introduction

In everyday life people obtain information; e.g., by seeing, listening and reading. However, in many cases, we do not know how presented visual information affects understanding of unfolding language and associated visual attention and how language processing can in turn affect ensuing visual attention.

Theoretical Background

Many experiments have been carried out to examine language comprehension using different methods. However, most studies have focused on the investigation of syntactic and semantic issues in sentence comprehension whereas less attention has been paid to non-linguistic information (e.g., Huettig, Rommers, & Meyer, 2011) and its effect on sentence comprehension. When we read a newspaper or watch a movie, we comprehend the content by combining the visual and language input. An interesting question is how comprehenders integrate these two information sources. In this line the so called "Visual World Paradigm" has been designed to examine this integration of information from a non-linguistic visual context with the unfolding interpretation in situated language comprehension. In a typical visual world paradigm study, participants see a scene including a number of objects either in the real world or on a monitor and they hear a sentence that refers to a particular target object in the scene. In some studies participants are instructed to perform a task in the experiment. For instance, they may be instructed to move an object from one place to another (e.g., Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In other studies participants task is to passively inspect the scene while listening to a sentence. They

do not have any other task (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Knoeferle & Crocker, 2007). In both cases participants' eye-movements are recorded during language comprehension.

A first study using this method was by Cooper (1974). The study tested situated language comprehension in a visual context. Participants' eye-movements were recorded while they listened to a story and viewed a display. Some objects from the display were mentioned in the story. Although participants were instructed that they could look anywhere they wanted, results showed that eye movements towards the objects were closely time locked to the spoken words. Participants even made anticipatory fixations towards a named object during the presentation of the target word or about 200 ms after its offset (Cooper, 1974). About 20 years later, Tanenhaus et al. (1995) applied this method to investigate language comprehension as participants inspected objects in different visual contexts. The study examined the effect of visual context on language comprehension as an ambiguous part of the sentence unfolded. Participants were shown a towel with an apple on it, another empty towel, a box and a pencil (one-referent context). In another context the pencil was replaced by an apple on a napkin (two-referent context). Participants were instructed to perform an action e.g., *Put the apple on the towel in the box* while their eye-movements were recorded. The prepositional phrase *on the towel* was ambiguous when the sentence referred to the one-referent context, and the first noun phrase *the apple* was ambiguous when the sentence referred to the two-referent context display. Results showed different gaze pattern during the one-referent and two-referent context. In the one-referent context, when participants heard the prepositional phrase *on the towel* they mostly inspected the empty towel indicating they had interpreted the prepositional phrase as a destination rather than a location. By contrast, in the two-referent context participants moved their eyes from one apple to the other during the first noun phrase and upon hearing the prepositional phrase they chose the correct location and looked at the apple on the towel. The authors suggested that subjects were influenced by the visual context in their understanding of the linguistic input. The eye-movement records suggested that participants had misinterpreted the ambiguous phrase as the destination in the one-referent context but not in the two-referent context. These results were taken to show that *eye-tracking data* can be used to examine rapid processing and syntactic structuring in studies of language comprehension (Tanenhaus et al., 1995, p. 1634).

Another study examined the anticipation in language comprehension as participants inspected a scene and listened to a sentence about possible activity involving objects in the scene. In a study by Altmann and Kamide (1999), participants saw a scene with a boy, a cake, a ball and other objects. While viewing the scene, participants listened to a sentence such as *The boy will eat the cake*. Results showed anticipatory eye-movements towards the cake compared with a sentence that contained a less constraining verb (*move*). When the verb *eat* was heard, participants used the verb information to anticipate the object (the cake) corresponding to a likely upcoming post-verbal noun phrase in the sentence.

Another study has investigated the effect of a scene depicting actions on language comprehension (Knoeferle et al., 2005). Participants were shown a depicted scene with a princess, a pirate, and a fencer. They listened to a German subject-verb-adverb-object (SVO) sentence in which the first noun phrase was ambiguous (it could either be the object and patient or the subject and agent of the sentence): *Die Prinzessin wäscht offensichtlich den Pirat*, 'The princess (subj/obj) washes apparently the pirate (obj)'. Despite the fact that the first noun phrase was ambiguous in the sentence, listeners used the verb and information from the scene (depicted actions) to make anticipatory eye-movements towards the patient (the pirate) already during the verb and at the beginning of the adverb.

Anticipatory eye-movements were found even early in the verb region in similar subject-verb-object sentences, when there was no ambiguity in the first noun phrase. In a study by Kamide, Scheepers, and Altmann (2003) participants saw a scene depicting a hare, a fox, and a cabbage and they listened to a German SVO sentence. When participants heard a verb *eats* in a sentence such as *Der Hase frisst gleich den Kohl*, 'The hare (subj) eats shortly the cabbage (obj)' they looked at the cabbage significantly more than at other objects in the verb region. According to the authors, this gaze pattern could be caused by case-marking in the German sentence combined with the verb information or by the subject-first preference active in German. Perhaps this information was combined sufficiently rapidly to permit listeners to predict the most plausible object in the scene (Kamide, Scheepers, & Altmann, 2003). These findings suggest that listeners are able to anticipate an upcoming post-verbal noun phrase referring to an object in the scene, before its mention.

Perhaps the anticipatory eye-movements have been observed because people tend to pay more attention to future information. Conceivably language about what

happens next in the world is particularly relevant for language users. Thus we may think precisely about the future while we use our knowledge of recent visual and language information. Our brain appears to be continuously engaged in predicting what will happen in the future based on recent experience (e.g., Bar, 2009; Hawkins & Blakeslee, 2007; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Other results suggest participants tend to show early inspections to objects representing potential arguments even if they are atypical given comprehenders' world knowledge (e.g., Boland, 2005).

However, other findings suggest that people rapidly integrate recent information and may even prefer to do so over anticipating what is coming next. For instance, in one study (Knoeferle & Crocker, 2007, experiment 3), participants listened to a sentence about plausible future events. Despite the event being plausible based on participants' world knowledge, participants showed no clear anticipation of the target of that future event. Experiment 3 by Knoeferle and Crocker (2007) examined how visual information of a depicted clipart action could have an effect on language understanding. Subjects were presented a scene depicting a waiter, candelabra, crystal glasses on a table, and other distractor objects. First, participants saw a clip-art action, where the waiter moved towards one of the objects. Next, the scene showed the waiter perform an action (e.g., polishing candelabra), and afterward participants saw another scene showing the waiter away from the object. While the scene was presented, participants heard a NP1-Verb-Adv-NP2 German sentence either in the past tense that referred to the recently seen action *Der Kellner polierte kürzlich die Kerzenleuchter*, 'The waiter polished recently the candelabra' or a sentence in the present tense with future meaning. The latter sentence referred to a plausible future action that had not been performed yet, *Der Kellner poliert sogleich die Kristallgläser*, 'The waiter polishes soon the crystal glasses'. In this experiment participants only saw a depicted clipart action before the sentence. When participants heard the verb 'polishes', they preferentially inspected the target object of the recent (vs. the other plausible future) event and their gaze pattern persisted even when the temporal adverb 'soon' became available. They inspected the crystal glasses only after they had been mentioned. Listeners ignored the information about future events provided by the sentence. It did not evoke expectations of future actions and they rather relied on the recently inspected action event for inspecting the depicted objects.

Further experiments examined the issue of the comparatively delayed inspection of

the plausible future event target using real-world action events (e.g., Abashidze, Knoeferle, Carminati, & Essig, 2011; Knoeferle, Carminati, et al., 2011). The experiment by Abashidze et al. (2011, experiment 1) used the same method as experiment 3 by Knoeferle and Crocker (2007) but the events were performed in the real-world with real objects. Participants wore a head-mounted eye tracker while they faced a person (an experimenter) sitting at a table and two objects were placed on the table (e.g., strawberries and pancakes; both objects can be sugared). First, participants saw the experimenter perform an action (e.g., sugaring strawberries) and they listened to a sentence in the past or in the future tense while the experimenter remained in a static position. A sentence in the past tense condition referred to the recently performed action *Der Versuchsleiter zuckerte gerade die Erdbeeren*, 'The experimenter sugared recently the strawberries'. In the future tense condition (with a verb in the present tense and an adverb indicating the future), the sentence referred to a plausible future action *Der Versuchsleiter zuckert demnächst die Pfannkuchen*, 'The experimenter sugars next the pancakes'. After the sentence no action was performed. This experiment replicated the preferential inspection of the recent event target. As in experiment 3 by Knoeferle and Crocker (2007) participants mostly looked at the recent event target (i.e., strawberries) irrespective of the sentence presentation. This effect has been called the recent-event preference.

The question remains, why participants failed to anticipate the plausible future event target earlier during the sentence. Participants were able to understand language about the past event even though that event was not present any more. This could similarly apply for the plausible future event. As shown in other findings, humans were able to predict upcoming information via their world knowledge and visual contexts before that information (e.g., an object) was mentioned. It is unclear to what extent a recently-perceived event could affect comprehenders' attention of a spoken sentence about a plausible future event. One could even argue that an event that we recently inspected would not be activated when an utterance refers to another future event. This is particularly the case in an experimental setup like the one by Abashidze et al. (2011); Knoeferle and Crocker (2007, experiment 2), in that half of the sentences referred to the recent event and the other half to the plausible future event.

The processing of language during an ongoing or completed event has also been examined by Madden and Therriault (2009). The authors used a word-by-word

sentence reading task, in which participants read past perfective and past imperfective sentences. The target words in the sentence were replaced by a picture of an object that was depicted either in use, i.e., depicting an ongoing action or as not in use, i.e., depicting a completed action. For example, a sentence could be *Fred was working/had worked on his "laptop" ... in the library.* Results showed an effect of a faster process when the sentence included a picture of a laptop in use (i.e., open) than of a laptop that was closed (showing a completed event) during the past imperfective sentence. By contrast, no difference has been found as a function of the laptop depiction for the processing of the past perfective sentence.

Furthermore, events that were narrated (but not depicted) can rapidly affect participants' eye-movement during later sentence comprehension. A study by Altmann and Kamide (2009) has investigated whether a mental representation derived from the narration of an event would affect participants' eye-movements towards the narrated (as opposed to actual) location of a target object. Participants saw a scene with a woman, an empty glass sitting on the floor, a table and a bottle of wine. Before the target sentences, they heard a context sentence either *a) The woman will put the glass onto the table.* Here the mental location of the glass is changed or *b) The woman is too lazy to put the glass onto the table.* Here the location is unchanged. Subsequently, participants heard a target sentence either *a) Then, she will pick up the bottle, and pour the wine carefully into the glass* or *b) Instead, she will pick up the bottle, and pour the wine carefully into the glass.* During *the glass*, participants' eye movements towards the glass did not differ in the moved and unmoved conditions; however, they fixated the table significantly more in the moved condition than in the unmoved condition. Participants' eye-movements were thus modulated by the description of the prior sentence even when that description clashed with the actual scene. These latter findings support the importance of the recent narrated context during language comprehension.

To recapitulate, when participants saw a scene with several objects and then they listened to a sentence referring to a target object of a plausible action, they were able to anticipate the target object before it was mentioned (e.g., Altmann & Kamide, 1999; Kamide, Scheepers, & Altmann, 2003; Knoeferle et al., 2005). But anticipatory eye-movements did not emerge when participants first saw a recent action event and then they listened to a sentence in the present tense with a future

meaning, referring to a plausible future event (e.g., Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011). Participants instead preferred to inspect the recent event target by a within-experiment frequency bias of future relative to recent events thoughtout the sentence.

In this thesis we investigated precisely the robustness and functionality of this gaze behaviour, viz., of the recent-event preference. We thus set out to assess by means of empirical investigation potential causes of the recent-event preference and the extent to which that preference is robust when stress-tested. In order to examine these research issues we pitted the following factors against the recent-event preference: First, we investigated a strong frequency bias towards future events. Next, we included another cue that should increase the likelihood of a future event occurring (an actor's gaze to the target of a future action). Finally, we eliminated the congruence between the recent event and the verb of the ensuing sentence to reduce reliance on the recent event.

The main questions of the thesis were: 1. Will a strong frequency bias towards future events affect and perhaps eliminate the recent-event preference? If this is the case, will the future event targets be also remembered better than the recent-event targets? 2. Will a situation-immediate cue such as an actor's gaze reduce or eliminate the recent-event preference? If so, will the future event target and future tense sentences be remembered better than the recent event target and the past tense sentences? 3. Will the incongruence of the recent event and the past tense verb increase the inspection of the future event and will participants then recall the future tense sentences better than the past tense sentences?

Outline of the Thesis

An outline of the dissertation includes the following: Chapter 2 gives an overview of findings in situated language comprehension. First, it presents the effect of depicted actions on language comprehension. Next, it reviews relevant findings on tense cues and verb meaning in visual attention and language comprehension, as well as links between attention and later memory. Furthermore, it discusses findings of frequency, of situational gaze and incongruence effects in language comprehension with or without visual contexts.

Since the goal of the thesis was to investigate the issue of the recent-event preference during language comprehension, chapters 3 to 5 present findings from five eye-tracking experiments. In all experiments we used a visual world paradigm

method. Participants were presented a visual action event and they listened to a spoken sentence that referred to the object of the event. In Chapter 3, we present findings of Experiments 1 and 2 that addressed the issue of whether the recent-event preference can be manipulated by a within-experiment frequency bias of future relative to recent events. To this end, participants saw more future events than recent events (presented before the sentence) and listened to more sentences about a future event than past tense sentences about a recent event. Other findings have shown the sensitivity in language and other cognitive processes to the statistical regularities of the input. If the introduced frequency bias, which was strongly in favour of the future events, is effective, we could expect that inspection of the future event target occurs often and early (reducing or eliminating the recent-event preference). Chapter 4 reports findings of Experiments 3 and 4 that further investigated the issue of the recent-event preference by introducing another powerful situational cue such as the gaze of the person who performed the actions. We examined whether gaze as a situational cue that is very effective in directing a listener's visual attention can override the overall recent-event preference. If the immediate gaze cue is stronger than the immediate recent event cue, then more attention should be directed towards the target objects when the gaze is present than when it is not present. In particular, more attention should go to the target of a potential future event when the sentence is about a future action and the actor shifts his gaze to the target of a possible future event.

People are sensitive to the incongruence between language and the visual context. Chapter 5 investigated whether the incongruence of the recently seen event and past tense verb could affect the preference of the recent event target in Experiment 5. If the incongruence influences the recent-event preference it would be expected that there would be an early decrease of looks towards the recent event target. These eye-tracking experiments tested the robustness of the recent-event preference. To further test its functional significance we assessed to which extent this inspection preference is mirrored in ensuing recall of sentence and event content. After each eye-tracking experiment we conducted a memory /gated memory tests. Experiments 1 to 3 examined later memory in recognising future and recent action events. The gated memory test in Experiments 4 and 5 tested recalling of the past and future tense sentence. Chapter 6 summarizes and discusses the findings. Furthermore, the chapter provides a German summary of the present thesis.

Chapter 2

Situated Language Comprehension

The present chapter reviews literature of eye-tracking and other behavioural studies that are relevant for the thesis. Among them are many studies that have used the VWP method to test visually situated language comprehension. That said, many language scenarios involve situated language. After a short introduction of some early eye-tracking experiments, we present findings on the effects of cues in the immediate visual and linguistic context on language comprehension. Furthermore, we highlight findings of depicted action effects on langue comprehension and the role of aspect/tense in sentence processing. The next section presents findings of attention in visual and language context and its effect on later memory. Towards the end of the chapter we present findings on frequency, gaze cue and incongruence manipulations (remember, these are the main manipulations in the present thesis project). At the end of the chapter we highlight once again the recent-event preference findings.

The research on situated language comprehension began almost fifty years ago. To recall, one of the first studies by Cooper (1974) showed a rapid interaction between visual information and utterance comprehension. Two decades after the research by Cooper, other studies used real-world objects and monitored eye gaze as people inspected objects and listened to related sentences (Tanenhaus et al., 1995). From that time onwards, many eye-tracking studies have examined the issue of real-time language comprehension in combination with a visual context

(e.g., Altmann & Kamide, 1999, 2009; Chambers et al., 2002; Eberhard et al., 1995; Knoeferle et al., 2005; Knoeferle, Carminati, et al., 2011).

For example a study by Eberhard et al. (1995) investigated the visual context effects during language comprehension, by introducing an unambiguous control instruction such as *Put the saltshaker that is on the envelope in the bowl*. This study was similar in its method/design to one study by Tanenhaus et al. (1995). The authors hypothesized that when participants hear an ambiguous sentence (e.g., *Put the saltshaker on the envelope in the bowl*) and they view a one-referent context display (presenting a saltshaker on an envelope and an empty envelope) they might misinterpret the location of the target object between the current location and the location of the empty envelope during *on the envelope*. By contrast, this effect should decrease or disappear during the unambiguous sentence. Results showed a distinct gaze pattern towards the target objects between the ambiguous and unambiguous sentence (see also Tanenhaus et al., 1995). Participants made more (incorrect) inspections to the empty envelope during the ambiguous sentence, which occurred more in the one-referent context than in the two-referent context. The authors explained these results as the visual context having a rapid influence on how we interpret the syntactic structure of a sentence (Eberhard et al., 1995).

A further study tested how action affordances might affect participants' language comprehension (Chambers, Tanenhaus, & Magnuson, 2004). Participants saw a scene with a glass and a bowl, both of which contained liquid eggs, and an empty bowl and flour. In another scene, the glass containing the liquid egg was replaced with a glass containing a solid egg. Participants' eye-movements were recorded while they listened to an instruction e.g., *Pour the egg in the bowl over the flour*. The authors expected that the first prepositional phrase *in the bowl* will be interpreted as a modifier when the scene showed two eggs in liquid form, in the glass and in the bowl. However, when the scene contained only one liquid egg and another egg in a solid form it was expected that participants would interpret the empty bowl as the goal during the first prepositional phrase. The gaze data analyzed from the second noun onset (*bowl*), revealed more looks at the false object (the empty bowl) when the scene contained the solid egg glass than the liquid egg glass. This gaze pattern lasted until the end of the sentence. According to the authors, the findings again showed a non-linguistic domain influence in the earliest moment of syntactic ambiguity resolution.

Another VWP study provided participants with a scene including objects such as a napkin, a box and other objects (Kako & Trueswell, 2000). Participants were instructed with a sentence such as *Now I want you to fold/pick up the napkin.* The authors were interested in the eye-movements to the target object (the napkin) during the highly constraining verb *fold* against the weakly constraining verb *pick up*. While participants could easily detect the object during the verb (*fold*) as only one object could be folded, the other verb (*picked up*) was compatible with all depicted objects. Results revealed an effect of the verb type. More looks occurred towards the target object (the napkin) during the verb *fold* than the verb *pick up*. Participants used the semantic strength of the verb for the inspection of the target object in the scene.

The above studies were presented as an introduction and as reminder of visual context effects in language understanding. In the next section we will review studies using similar manipulations, as we pay close attention to studies that revealed immediate visual and linguistic cue effects in language processing.

2.1 Immediate Visual and Linguistic Context

This section highlights findings on immediate linguistic and visual context cues, which is in line with the present project question of the recently perceived action event influence on language comprehension. On the one hand, visual cues such as the shape, the color, the size of an object and/or a depicted action can all interact immediately with language comprehension. On the other hand, immediate linguistic cues such as a preposition or the tense of a verb can modulate our language understanding.

Participants' eye movements were recorded while they moved an object on the screen (Dahan & Tanenhaus, 2005). A computer display showed four pictures e.g., a snake (target), a rope (competitor) plus two other unrelated objects. The authors wanted to examine whether or not on hearing the target word *snake*, participants would look more at the competitor object (visually similar) than other unrelated distractor objects. The results showed that participants inspected the competitor object less than the target object but approximately 300 ms after the target word onset more fixations emerged to the rope than distractor objects. These findings suggested an early visual shape similarity effect during language processing (see

also, Huettig & McQueen, 2007; Rommers, Meyer, & Huettig, 2015). A related experiment by Huettig and Altmann (2011) tested effects of object colour and associated world knowledge in language comprehension. In two experiments participants were presented a scene with colorless (black-and-white) line drawings of typical objects, such as: a frog, a cake, a glove and an arrow. Participants listened to a sentence that included a target word *frog*, and another sentence presented a color competitor word *spinach*. Results showed that in the target condition as soon as the target word *frog* became available, subjects inspected the target object significantly more than other objects. But when participants heard *spinach* their eye-movements towards the target object (the frog) occurred approximately 400 ms after the word onset compared with inspections of other unrelated objects. According to the authors, stored object color knowledge influences participants' language-mediated eye fixation but at the same time its effect seems to be weak.

Another recent study by Coco and Keller (2014, experiment 1) investigated whether visually salient objects are activated during verb processing and help to anticipate an upcoming argument of a verb. A presented scene contained objects such as an orange on a tray, a tray in a bowl, an empty bowl, another object and a girl. In one scene the colour of the bowls differed (salient scene) and in another scene they had the same colour (non-salient scene). A critical sentence was, for instance *The girl will put the orange on the tray in the bowl*. Compared to other studies e.g., by Tanenhaus et al. (1995) the study presented declarative sentences and the scene also depicted the agent of the action. The authors were interested in the gaze behaviour during the presentation of the direct object *the orange*. Results revealed a main effect of the saliency in that participants made significantly more anticipatory looks towards the goal location (the empty bowl) in the salient than in the non-salient scene.

Not only visual cues are powerful in modulating comprehenders' eye-movements during language comprehension but also linguistic cues are early modulators in situated language processing. Experiments by Sedivy, Tanenhaus, Chambers, and Carlson (1999) tested how and when subjects use contextually-defined contrast in interpreting simple pre-nominal adjectives during spoken sentence comprehension. Participants saw a display with a yellow duck, a red notebook, a pink comb and a blue pen. They were instructed with a sentence e.g., *Touch the blue pen*. For the same instruction another display had replaced the pink comb with a blue bowl. Participants used the linguistic cue of the color adjective *blue* and they inspected

the target (the pen) earlier in the first display than in the second display with two blue objects. This suggests the adjective was used at an early point for disambiguation instead of looking more at the semantically related object (i.e., the notebook). An early use of a size adjective occurred with another instruction *Pick up the tall glass* when a scene showed a small and a big glass among other big and small objects. After the linguistic information participants incrementally inspected the relevant object, the big glass more than another big object in the visual scene.

An early preposition effect has been reported in the following experiment by Chambers et al. (2002). Their study manipulated perceptual and linguistic context to examine participants' interpretation of a target object when a sentence unfolded. In experiment 1 participants saw a scene with four small objects placed in the middle of the scene and another four big objects were placed in the four corners of the scene. While inspecting the scene, they listened to an instruction e.g., *Pick up the whistle and hold it over the cross. Now put it inside/below the can.* In half of the trials a one-container condition presented a can (e.g., in which a whistle can be put). In the other half of the trials, a three-container condition presented three objects (a can, a bowl and a glass in which a whistle can be put). The authors expected an earlier anticipation of the target object in the *inside* than in the *below* sentence. This expectation relates to the fact that *inside* narrowed reference to only one object i.e. in the one-container condition while *below* did not have a comparable effect. When participants hear the preposition *below*, their attention towards the target object (the can) occurred approximately 350 ms after the target name onset. The gaze pattern towards the target object was similar in both the one- and the two-referent conditions. By contrast when participants heard the preposition *inside*, their looks towards the can in the one-container condition emerged at the end of the preposition and during the first 100 ms of the target word region. Participants looked significantly more at the can than other objects. This suggests an immediate use of the preposition to restrict the domain of reference to a single object when only one object was a plausible container. Similarly to the preposition a definite versus an indefinite article of a target can also speed up fixations toward a target object. In experiment 2 by Chambers et al. (2002) half of the critical trials used a target noun with a definite determiner (the can) and another half used an indefinite article (a can) e.g., *put the X inside the/a Y*. Results in the definite article condition showed significantly more inspections to the target object than other objects in the scene in the 200-300 ms interval

after the onset of the article. However in the indefinite condition the earlier looks towards the target emerged in the three-container instead of in the one-container conditions.

Some other studies investigated the semantic similarity effect of language-mediated eye movements. Huetting and Altmann's (2005) study examined how semantic properties of individual items can drive participants' eye-movements toward objects in a visual scene when they heard for instance a target word *piano* in a sentence such as *Eventually, the man agreed hesitantly, but then he looked at the piano and appreciated that it was beautiful.* Participants inspected a scene containing four objects including a piano (target condition). Another scene showed a trumpet (competitor condition) instead of the piano. When subjects heard *piano*, a semantically related word in the competitor condition, they looked earlier and more at the trumpet than any other individual object. However, fixations at the onset of piano towards the competitor (the trumpet) did not differ reliably from looks to the other individual objects in the scene. Interestingly, the results neither showed any significant effect in the target condition when comparing looks to the piano with looks to the other objects.

Here we discussed findings of immediate linguistic and visual context effects among others. We saw how those cues can modulate our online language comprehension. In the following section we will review studies that have examined the role of depicted actions in language comprehension.

2.1.1 Depicted Actions

The role of the depicted actions and visual context information that can impact auditory sentence comprehension several hundred milliseconds after the visual stimulus onset, is discussed in the section. Some visual world studies used contexts containing objects but no actions (e.g., Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus et al., 1995). Other studies however used depicted actions during language comprehension (e.g., Knoeferle et al., 2005; Knoeferle & Crocker, 2007).

One of the first studies by Knoeferle et al. (2005) investigated the effects of depicted actions that were described by structurally ambiguous sentences. In experiment 1, the authors examined the comprehension of initially structurally ambiguous spoken German subject-verb-object (SVO) and object-verb-subject (OVS) sentence

structures. For a particular trial participants inspected a scene with a princess, a fencer and a pirate. The scene showed two events with the princess depicted as painting and washing at the same time. One character in the scene (namely the princess) is ambiguous in her role: she is the patient of the painting event and at the same time the agent of the washing event. While participants viewed the scene they listened to a SVO sentence e.g., *Die Prinzessin wäscht offensichtlich den Piraten* 'The princess (amb) washes apparently the pirate (ACC)' or an OVS sentence e.g., *Die Prinzessin malt offensichtlich der Fechter*. 'The princess (amb) paints apparently the fencer (NOM)'. One of the expectations was that upon hearing the verb 'washes', subjects would inspect the pirate as the patient of the washing - event and they would direct more anticipatory eye movements to the agent for the SVO than for the OVS sentence. Shortly after the onset of the verb 'washes' (in the SVO sentence) and 'paints' (in the OVS sentence) did the subjects begin to look at the pirate (the target character) more often than at the fencer for SVO sentences. These results may indicate that during the adverb region, participants eye gaze shifted more towards the fencer (as an agent) in OVS than in SVO sentences. This gaze data suggests that during the adverb region the role-ambiguous character (the princess) was interpreted as the patient for the OVS sentences. The main finding is that at the beginning of the adverb people rapidly looked to the appropriate characters for SVO and OVS sentences based on the depicted actions.

A further experiment by Knoeferle and Crocker (2006), examined the verb-mediated influence of depicted actions and world knowledge. In experiment 2, participants were presented scenes that used two sources of knowledge which were in conflict. For example in one trial the scene showed a wizard, a detective, a pilot and other objects. They were displayed as the wizard looking at the pilot through a telescope while the detective was serving the pilot some food. The authors compared two different types of information in this context: the knowledge from experience (e.g., a detective spying) with the depicted action in the concurrent clipart. The pilot as the middle character was always a patient, while the other characters (e.g., a wizard and a detective) were always agents. In general, people could expect through their stereotypical knowledge what the wizard and the detective would usually do. The sentences were presented in an unambiguous OVS sentence order, a) *Den Piloten verköstigt gleich der Detektiv*, 'The pilot (obj) serves-food-to soon the detective' and b) *Den Piloten bespitzelt gleich der Zauberer*, 'The pilot (obj)

spies-on soon the wizard'. The sentences always began with an accusative case-marked noun phrase, referring to the pilot in this case. From our world knowledge, we would expect that during the verb 'spy-on' participants would look at the detective more than at the wizard. One of the questions asked in this experiment was whether at the beginning of the verb 'spy-on', subjects would direct their eyes toward the stereotypical expected agent (detective), or whether subjects would inspect the figure that is depicted as spying in the concurrent picture (wizard). It appears that participants did not immediately use their stereotypically knowledge to look to the detective. The results showed that participants inspected the wizard more than the detective at the onset of the verb 'spy-on'. The results suggest that depicted actions influence our language understanding. Participants appear to consider the depicted action relevant at the verb more than the stored stereotypical knowledge (Knoeferle & Crocker, 2006). The influence of depicted actions emerged even when the actions were no longer visible during utterance presentation. The first experiment by Knoeferle and Crocker (2007) used the material from the above presented experiment 2 (2006). But the procedure differed in a way that participants first saw the depicted action for 1500 ms and then the characters and the actions disappeared. During the sentence presentation the screen was blank and the authors tracked eye gaze to the previous location of the objects. The results replicated the effects of the depicted events, even when the scene was no longer present.

In highlighting the effect of the depicted action against stereotypical world knowledge during language comprehension, another experiment showcases the relevance of world knowledge. A study by Kamide, Altmann, and Haywood (2003) was designed to examine whether participants could use their world knowledge in anticipating an upcoming post-verbal noun phrase at the verb. The scene showed a man, a girl, a motorbike, and a carousel. During a particular trial participants listened to a sentence e.g., *The man will ride the motorbike* or *The girl will ride the carousel*. The authors were interested in the eye-movements during the verb *ride*, whether participants would use the first noun phrase (the agent, world knowledge) and verb information and anticipate the correct object during the verb. The results revealed an effect of world knowledge usage during the verb so that they inspected the correct character in the scene before the onset of the name. For example they looked at the motorbike in the *The man will ride* condition. These findings corroborates the strong influence of noun meaning, verb meaning, and world knowledge during language comprehension.

Furthermore, the effect of a recent depicted action can modulate language comprehension of an utterance about a plausible future action. In experiment 3 of their paper, Knoeferle and Crocker (2007) examined how visual information of a recent action could have an effect on language comprehension. Participants were shown a scene with the following objects: a waiter, candelabra and crystal glasses and they listened to NP1-Verb-Adv-NP2 German sentences. First, participants saw a clip-art scene where the waiter moved toward an object and then performed an action (e.g., polishing candelabra), and afterward another scene showed the character away from the object. Participants then heard either a past tense sentence such as *Der Kellner polierte kürzlich die Kerzenleuchter*, 'The waiter recently polished the candelabra' or a future tense sentence *Der Kellner poliert sogleich die Kristallgläser*, 'The waiter will soon polish the crystal glasses'. In this study, for each experimental trial, participants only saw a clip-art action scene of the recent event (the waiter polishing the candelabra). They did not see any future action after the sentence; but theoretically they did not have any reason to not think that a plausible future action would be performed after the sentence. When they heard the verb 'polish', participants preferentially looked at the target of the recent (vs. the plausible future) event and this gaze pattern persisted even when future tense became available through the adverb (e.g., 'soon'). In other words, the future tense information did not evoke expectations of future events and the preferential inspections towards the recent target lasted throughout the sentence. The inspections to the plausible future event target occurred only during the NP2 (crystal glasses).

This is an interesting finding, as it suggests a very strong influence of a recently depicted clip-art action on language about the plausible future action. One might argue that the reason for this strong influence could have been the experimental materials. A further experiment addressed this question by using real-world objects; however, the results have replicated the previous findings and showed even stronger effects of the recent action events (Abashidze et al., 2011; Knoeferle, Carminati, et al., 2011). Participants inspected the future event target only at the end of the NP2 region. The delayed inspection of the future event target (Abashidze et al., 2011, experiment 1) was investigated in further experimental manipulations and their results will be presented in more detail at the end of the present chapter.

In summary, the results discussed in the present section provide clear evidence

of depicted action effects during language comprehension. Furthermore, listeners rely more on the depicted action than their stereotypical world knowledge at least for processing unambiguous OVS sentences. Moreover, the effect of a recently seen action seems strong and long-lasting even when linguistic cues point to future events during language comprehension. The fact that the future tense information did not have a strong effect on the anticipation of the plausible future action target during the adverb region further corroborates this view.

At this point we cease discussing the effect of visual cues during sentence processing, but we continue reviewing findings on the tense cues and verb meaning in sentence comprehension in the next section.

2.1.2 Effects of Tense Cues and Verb Meaning

This section will review studies that have examined the effects of tense information in language processing, and in general verb understanding. Verbs can convey the time of an action through morphosyntactic marking and this is relevant for understanding the time of events (e.g., an action has happened in the past, is happening in the present, or will happen in the future). Moreover the verb conveys more than just past or future tense information. In our everyday language we use different types of tense information for describing an event. However, contexts in which an event has just happened and in which other events might happen in the future are interesting for examining how we integrate information from perceived events with linguistic input. On the one hand we might argue that a past tense compared with a present tense as a pointer to an event might be processed fast because of additional event input. On the one hand, however, a future tense cue might be processed fast, because humans have a tendency to anticipate what will happen next, as evidence in research on language prediction suggests (e.g., Altmann & Kamide, 1999)

In one experiment Trueswell and Tanenhaus (1991) examined the reading time of past and future sentences when participants previously received a past or future tense context. A past and future context sentence example is e.g., *Several students were/will be sitting together taking an exam yesterday/tomorrow. A proctor came up/will come up and noticed/notice one of the students cheating.* For the target sentences participants read a sentence e.g., *The student spotted by the proctor received/will receive a warning.* First, participants read a context sentence that was

displayed on a screen as an entire text and then they performed a self-paced reading task by pressing a button to uncover words. Results showed that participants read experimental sentences faster when the previous discourse was presented in the future than in the past context. The authors reported that the main interest of the article which was the manipulation between future and past context, clearly showed that subjects could use the tense context information for later sentence comprehension.

In another sentence-reading experiment participants were presented a simple or progressive past tense sentence on a monitor (Madden & Zwaan, 2003). Immediately after reading the sentence, two pictures (depicting a complete and an ongoing action) were presented above the sentence and participants were asked to choose which picture was referred to by the sentence. The results showed a faster selection of the completed action picture when the sentence was in the simple past than in the simple progressive tense. By contrast, participants were faster in processing of a sentence describing a character in the present than past tense. Carreiras, Carriedo, Alonso, and Fernández (1997) let participants read paragraphs in which the tense of the verb was manipulated. Sentences described someone?s activity either in the past or in the present and after the sentence the name of the character was presented. The results showed faster accessibility in the present than the past tense sentence.

In another experiment by Gilead, Liberman, and Maril (2013) participants were presented with sentences in the past, present, and future tense and in addition to reaction time measures the scientists measured participants' neural activity while they were reading the sentences. One of the expectations for the future tense sentences was an activation of brain regions implicated in processing intentions. For example when participants read a sentence such as *Danny will kick the ball*, one may recall a memory of planning to kick a ball or of someone else having that goal. The results of the reaction time showed faster responses in the future tense than in the past and present tense sentences. The authors further reported that the future tense did not activate all of the regions involved in processing intentions; but the activation in the 'ventromedial prefrontal cortex' (which is one of the regions implicated in processing intentions) was significantly higher during the future (vs. past / present) tense sentences (Gilead et al., 2013). Another experiment examined the processing of future and past tense sentences using a sentence-picture matching task while participants' eye-movements were recorded

(Bos, Hanne, Wartenburger, & Bastiaanse, 2014). Participants were shown a scene with black-and-white line drawing objects such as a man and two bottles (depicted in an upper and a lower location). The objects were depicted either as associated with an event in the past (an uncorked bottle signaling its content had been drunk) or as associated with an event in the future (a bottle sealed with a cork signaling that its contents could be enjoyed in the future). Participants listened to either a past or a future tense sentence and they responded to a question about whether the upper or lower picture (i.e., the "past" or "future" object) matched the sentence. Results showed longer response times for the future than past tense sentences. The eye-movement data showed overall preferential looks to the target of the past (the uncorked) compared with the future event picture (the sealed bootle).

In the following text we review studies showing verb information effects in anticipation (e.g., Altmann & Kamide, 1999, 2007; Kamide, Scheepers, & Altmann, 2003; Knoeferle & Crocker, 2006). Altmann and Kamide (1999) investigated how people react to a scene when they hear a sentence, and especially a verb that conveys something will happen in the scene. In their experiment participants inspected semi-realistic visual scenes with several clipart drawings such as a boy, a cake, and a ball. While they inspected the scene they listened to a sentence, e.g., *The boy will eat the cake*. In half of the trials the scene did not depict the target object. Despite the fact that the scene did not include the target object in half of the trial sentences, participants used the verb *eat* for anticipating the target object when the latter was depicted in the scene. Eye movements toward the cake occurred before the onset of the post-verbal phrase *the cake*. Participants thus integrated auditory and visual context information and this integration happened rapidly enough to make anticipation of the verb of an upcoming target object possible. In a similar study by Altmann and Kamide (2007) participants were shown a scene with a cat, a couple of mice, a pile of feathers and other objects. When participants listened to a future/past tense sentences such as *The cat will kill / has killed all of the mice*, during the future tense sentence, they mostly looked at the mice than other objects in the scene (after they had listened to the verb *will kill*, using the scene and the verb tense information). Similarly after the past tense sentence verb *has killed*, they inspected the pile of feathers more than the mice. In both tense sentences, the verb information was used quickly together with the world knowledge of the scene for visual anticipation of the target objects. These cues rapidly affected language comprehension.

Another experiment showed participants a scene with a fox, a hare and a cabbage and manipulated the German SVO and OVS sentence structure (Kamide, Scheepers, & Altmann, 2003). One of the research questions was whether participants would use case marking for comprehension and anticipation or whether participants would need in addition the verb as a cue for the anticipation. When participants listened to a sentence (e.g., *Der Hase frisst gleich den Kohl*, 'The hare-nom eats shortly the cabbage-acc') referring to an activity in the scene, no effect of case marking on the first noun phrase emerged. However, when participants heard the verb, they rapidly looked at a target (the cabbage). The combination of case-marking and verb information enabled participants to predict a referenced target object.

By using depicted action information and the tense of the verb, participants inspected an ambiguous character in the scene after they had heard the verb (Knoeferle & Crocker, 2006, experiment 1). Participants were shown a scene depicting a ballerina, a pirate and a fencer, in which the ballerina was an ambiguous character in the scene, depicted either as an agent or as a patient. Participants listened among others to a future tense sentence *The ballerina will splash the cellist* or a past sentence *The ballerina splashed apparently the cellist.* Even though the ballerina was an ambiguous character in the scene, participants inspected the post-verbal target character (e.g., the cellist) before it was mentioned. The combination of the depicted action and the verb information elicited the disambiguation at an early point. The above discussed findings suggest a strong tense effect in language comprehension. The tense of the verb information is quickly used to comprehend the sentence. Furthermore, the tense information of a verb is instantly used to build a temporal relation between a previous event and other events in a discourse (see also Altmann & Steedman, 1988; Trueswell & Tanenhaus, 1991).

Thus, language comprehension is informed by both scene cues and linguistic as well as world knowledge. In the next section, we consider the extent to which real-time language processing and attention impacts later processes (e.g., ensuing memory of language content). Links between attention and memory are of interest because they bring together research of language processing with learning.

2.1.3 Links between Attention and Later Memory

This section reviews studies on attention and its effects on later memory. One of the aims in reviewing studies on this topic is to clarify the state of the art on how language processing connects to later memory. This connection is relevant for the present thesis because we asked how the effects of the non-linguistic variables that we manipulated related to ensuing memory in post-experiment tests.

A study by Trueswell and Papafragou (2010) investigated the role of language interference on remembering events for a later memory test. Participants saw a series of clipart animations of a boy roller-skating towards a soccer goal. While inspecting the clipart scene participants in one condition listened to auditory input (two-digit number, linguistic interference with scene encoding) and they had to repeat these numbers until the end of the encoding phase. In another condition they only viewed the animations (no interference), such that there was no disruption for linguistic encoding of the animation. Immediately after the experiment participants were asked to perform a memory test. In addition to target animations they were shown new animations. The responses in the memory test were made verbally and the experimenter wrote down the answers. The results showed that participants were significantly better in recalling the animation in the no-interference condition than in the linguistic interference condition. The findings suggest that the encoding and ensuing recall of an event can suffer from interference through an irrelevant task.

In an eye-tracking study, Zelinsky, Loschky, and Dickinson (2011) investigated how behavioral pattern of re-inspection of an object affected participants' performance on a post-trial forced-choice target recognition task. Participants were presented a scene with nine objects, in which one object was predesignated as a target for the memory test (e.g., a butter dish). Their tasks were to inspect objects carefully and remember their location. Without participants being aware of this, the eye tracking program counted the number of fixations they made to objects. Next, participants saw an empty scene with a spatial probe at the target object location (i.e., where the butter dish had been). Afterwards, the scene showed four objects (including the target object) at randomized locations and the participants' task was to indicate which of these four objects appeared at the probe location from the previous display (e.g., the butter dish could have been presented in the lower right corner and this is where the spatial probe mask appeared). But for recognition,

the butter dish might then be presented at a different location. The results showed that participants' re-fixations to the target increased to the extent that they had fixated more intervening objects. Re-fixating the target object improved their accuracy in the recognition task by 16 percent. This suggests a clear link between how participants deploy visual attention during encoding and their ensuing success in recalling object identity (see also Zelinsky & Loschky, 2005).

Other experiments examined how (changes in) the spatial location of an object modulates ensuing memory performance. Santa (1977) found that participants were better in recognizing the visual stimuli when objects were presented at the same location as in a previous scene compared to when the location of objects changed. The role of implicit spatial information during the encoding and recognition of visual stimuli has been investigated in another study by Richardson, Spivey, Barsalou, and McRae (2003). Participants listen to a sentence e.g., *The athlete succeeds at the tournament.* During the first noun phrase they saw a picture of an athlete and during the verb the screen went blank. When they heard the last noun phrase they saw a picture of a tournament, centrally presented. After a six-trial practice block participants were presented a test block of 12 trials (including six new trials). In the test block participants saw two pictures either in a vertical or horizontal orientation. Their task was to judge whether those pictures had been presented in the previous block. The results showed faster identification of pictures in the horizontal than in the vertical orientation. What these results suggest is that spatial aspects of stimuli during encoding (i.e., the picture of the practice trials) can affect ensuing task performance.

Other experiments tested the capacity of attention and the recognition of arrays. Luck and Vogel (1997) presented participants an array of 1-12 coloured squares for a 100 ms period. After this period, a blank screen occurred for about 900 ms. Afterwards, participants saw a scene depicting arrays that were either from the prior scene or from a new scene and that differed in colour. Participants' task was to pick the arrays that they had observed in the prior scene. They performed the task perfectly for arrays of 1-3 items (see also Pashler, 1988); however, their performance decreased systematically when the number of items in the array increased. In a second experiment the authors were interested whether additional verbal information can influence the participants' task performance. The task was similar to the first experiment; additionally, in the second experiment participants saw two digits before the trials and they were asked to hold these digits in memory

and name them at the end of the trial. The results between experiments did not differ, suggesting that the capacity of recalling correct objects was not affected by additional verbal load.

Another experiment enriched a scene even with multiple objects from different categories e.g., squares, Chinese characters, polygons and shaded cubes (Awh, Barton, & Vogel, 2007). Subjects were shown a scene with either four objects or eight objects from one of the categories and these objects remained on the screen for 500 ms, followed by a 1000 ms blank period. Afterwards, a group of the objects appeared on the display. For the other half of the trials, participants saw instead other randomly chosen objects from the same category. Participants were significantly better in selecting the correct objects in the four-object than in the eight-object conditions and also better for the color objects than the cubes or shaded cubes/polygons. All of the above experiments agreed in the reduction of memory task performance when the complexity of objects increased (similar findings are reported by Alvarez & Cavanagh, 2004).

Another study tested the recall of direct-gaze or averted-gaze faces (Adams, Pauker, & Weisbuch, 2010). Participants were presented a block of 40 faces for 3.2 second each. Afterwards they performed a distractor word choice task (they had to choose particular words in the scene) for 5 minutes. As a next step they were presented with these 40 faces randomly intermixed with 40 new faces. The task was to indicate as soon as possible whether they had seen the face in the first block or not. Findings showed a better recall of the direct-gaze faces than the averted-gaze faces.

Hulme, Maughan, and Brown (1991) were interested in short-term memory of familiar and unfamiliar word pairs. The authors gave participants several words combinations to read (e.g., *switch, maths, scroll* ...) and as well as several non-word combinations (e.g., *fot, bin, zog* ...) and told to them to repeat each of the word/nonword-combinations five times. Then they measured memory span by presenting a word/nonword-combinations starting with two words increasing in number. The task was to recall words in the correct serial order. The results showed that the memory span was higher in the word than in the non-word recall and also higher for the one syllable words than two or three syllable words.

In sum, we saw that many studies showed evidence of the role of visual or language information in encoding, remembering and recalling by testing participants'

ability to recall stimuli in the later memory task. Moreover, the objects that were re-fixated in the scene, were easily recalled in the ensuing memory task. However, the results that we have reviewed provide no insight into how information from a visual context compared with linguistic input (e.g., verb tense) affects the memory of events in ensuing memory tasks. For example, when we look at one object longer than at another object, does this mean that we remember the object better in a later memory test than another object that was looked at less? Furthermore, do more fixations made in a particular tense sentence condition than in another sentence condition help to recall language information in the later memory task? For example, when we inspect an object more often while we listen to the sentence in a past tense, does the gaze pattern support the better recall of the past tense sentence than a corresponding future tense sentence? These are the questions that we examined in the present memory tests. In the following sections we turn to present findings on manipulations that are relevant for the present thesis experiments. What is the role of the frequency of language and visual input in language comprehension? We review studies that have employed different behavioural manipulations.

2.2 The Role of Frequency

We learned in the above sections that language comprehension is sensitive to different stimuli (e.g., words, or sentence structures), and to different contextual constraints (e.g., verb meaning, action depictions, tense information). Another aspect of the input that comprehenders are highly sensitive to is the frequency of stimuli (e.g., how often a specific sentence structure occurs) and additionally we can ask what effect, if any, a frequency bias has on success in a later memory task that asks participants to recall information about language and inspected events. Since the present thesis exploits frequency biases to stress-test the recent-event preference, this section reviews previous findings on frequency manipulations at the word, phrase and sentence level (e.g., Arnon & Snider, 2010; Rayner & Duffy, 1986; Turk-Browne, Scholl, Johnson, & Chun, 2010; Wells, Christiansen, Race, Acheson, & MacDonald, 2009). Furthermore, we review studies that have examined the frequency in orthographic pattern (e.g., Vanyukov, Warren, Wheeler, & Reichle, 2012) and of motor experience in a reaching task (Chapman et al., 2010).

Words that appear more frequently in language are processed more quickly than less frequent words (e.g., Morton, 1969). In a reading experiment, sentence comprehension was affected by frequent words that were included in the sentence. In an eye-tracking study by Rayner and Duffy (1986) participants read a sentence e.g., *The slow waltz/(music) captured her attention.* The high-frequency word is in the parentheses. Findings showed that participants read the sentence with *music* faster than with *waltz*. The occurrence of the infrequent word in the sentence lead to an increased inspection time on the next word. It was suggested that the infrequent words might be more difficult to access in the lexicon and to integrate with the preceding context (Rayner & Duffy, 1986). A further study investigated frequency effects in a phrasal-decision task, that was tested by conducting a reaction time experiment (Arnon & Snider, 2010). Participants read four words phrases e.g., *Don't have to worry*, for high frequency, and *Don't have any money*, for low frequency conditions. In addition most filler phrases were grammatically incorrect e.g., *I saw man the.* The task was to decide whether the word order would be possible in a sequence of an English phrase. As predicted, results showed that participants responded more slowly to the low than high high frequency phrases. Additionally, participants were faster in decision making in the second block than in the first block. These results suggest a correlation between frequency of occurrence within the experiment and processing times: the more often a phrase had been experienced, the faster it was processed.

Another experiment has investigated the effect of statistical regularities on the processing of relative clauses (Wells et al., 2009). The experiment manipulated four sessions separated each by four to eight days. For instance, participants did a self-paced reading test in session 1 and a test in session 4. The test in the final session 4 was always scheduled at least four days after the last experience test in session 3. Results showed an effect of experience in that participants read all sentences faster in the test session 4 than in the test session 1. In addition, participants' reading time seems to be sensitive to statistical regularities in the stimuli; even though the last session took place several days after the last experience-focused session (session 3), participants showed a strong effect of experience in processing relative clauses.

The effects of stimulus frequency on language processing were also examined at the word level. In an offline rating study, Reali and Christiansen (2007) compared complexity and plausibility across doubly embedded object-relative sentences. The

authors used *google* counts to measure the frequency of specific 'I plus verb' combinations. They manipulated the frequency of word co-occurrence in the deeply embedded clause e.g., *The detective who the attorney who I met distrusted sent a letter on Monday night*, (high-frequency) and *The detective who the attorney who I distrusted met sent a letter on Monday night*, (low- frequency). Participants' task was to read and rate the plausibility of sentences. Results showed that the high-frequency word combination sentences were less difficult and rated higher in plausibility than the low-frequency word combination sentences. The authors concluded that the plausibility ratings were influenced by frequency in the most deeply embedded clauses.

Other than the above reviewed reading studies the next experiment by Vanyukov et al. (2012) manipulated the frequency of orthographic patterns in a visual-search task and tested its effect on eye-movements. The authors suggested that the frequency effects in a non-reading task could show strong effects on cognition and rapidly influence eye movements. They used a paradigm where participants saw circles with an opening (a so-called 'gap') of varying orientation. Participants were given an instruction to view a horizontal row of circles from left to right and to detect any targets i.e. a cluster containing the letter *O*. After the disappearance of each cluster, participants were asked whether a target had (vs. hadn't) been present. The data did not reveal any reliable effect in the first-fixation duration in different frequency categories. However, total gaze duration decreased with more exposures and more practice in the task. The authors suggested these findings parallel the word-frequency findings in reading studies (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986).

Other than reading time changes based on frequency manipulation of words, phrases, or sentences, a study by Chapman et al. (2010, experiment 2) examined how visuomotor decision unfolds in real time. They used a paradigm where participants had to make a quick hand movement toward potential targets (circles), when only one target location had been cued before the reaching task. First they saw a fixation cross in the middle of a display and this was followed by the target display. The target display showed two circles and one of them was cued (e.g., a black filled-in circle) or none of the two circles were cued (a not filled-in circle). The two circles were presented on the left and right side on the display. During the target display participants had 325 ms to lift their finger and to prepare for the task (reaching for a target). Afterwards the reaching task display appeared

and participants had an additional 435 ms to reach one of the targets on the display. Results showed that participants moved their finger toward the middle in the target display when no cued circle had been presented in the target display. However they made a straight hand movement towards the target when a cued target had been presented before. These findings suggest that the visual-motor experience of the target (the filled-in circle) location influenced participants' rapid hand-movements within a very short time (350-500 ms).

Yet another experiment examined the effect of frequency biases in a series of visual-world experiments. In experiment 1 by Abashidze et al. (2011) participants first inspected an action event and then listened to a sentences in either a past tense condition referring to the recent event target or in a present tense condition with at future meaning referring to a plausible future action. This setup effectively instantiated a frequency bias towards the recent action and its target (participants never saw the future event performed). Accordingly, participants tended to pay more attention to the target of the recent action than to the target of the future action while listening to sentences about these events. The bias towards the recent target inspection continued until the end of the sentence. By contrast, experiment 2 presented both one action before the sentence presentation and another action after the sentence presentation. Results of the second experiment revealed earlier and more looks at the future event target (when future events were as frequent as recent ones) in the future tense than in experiment 1 (when only the recent but not the future action events were acted out).

This section has reviewed a number of experiments that on the one hand examined the effects of within-experiment manipulations of stimulus frequency at the word, phrase, and sentence level. On the other hand it has become clear that the frequency of visual cues can also affect the visuomotor or gaze behaviour to the target object. Taken together, these findings show that humans are exquisitely sensitive to stimulus frequency and that the latter can affect their language comprehension and visual attention. The present section has introduced frequency effects, thus reviewing findings that are relevant for the first two experiments of the present thesis. The next section introduces the effects of another cue (the gaze of an actor). Experiments 3 and 4 have examined precisely how the gaze of an actor might affect participants' attention allocation during language comprehension. For this reason, the next section discusses studies that have examined a situational gaze cue in visual attention and language comprehension.

2.3 The Effect of a Situational Cue - Gaze

The *language of the eyes* has a rich vocabulary, which is able to communicate mental states such as, beliefs or predictions (Frischen, Bayliss, & Tipper, 2007). The present section highlights studies that have examined the effects of a non-verbal, situated cue such as *gaze* in visual attention and language comprehension. Furthermore non-linguistic information such as gaze directed at an object can make that object more rapidly accessible than objects cued through a verb and its recent action referent (Knoeferle & Kreysa, 2012; Kreysa, Knoeferle, & Nunnemann, 2014).

In the literature, the gaze cue effect (i.e., that listeners tend to follow the gaze of a speaker) has been divided at least in two different attentional pattern: on a perceptual level, a *gaze* cue effect might be interpreted as (reflexively) increasing the visual saliency of a particular target object or scene (e.g., Driver et al., 1999; Friesen & Kingstone, 1998; Langton, O'donnell, Riby, & Ballantyne, 2006). By contrast, on a cognitive level, *gaze* may additionally be understood as a cue to the speaker's referential intentions which evokes an expectation of a target object that will be mentioned next (e.g., Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012; Kreysa & Knoeferle, 2013; Kreysa et al., 2014; Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009; Staudte, Crocker, Heloir, & Kipp, 2014).

Driver et al. (1999) investigated participants' reaction time in relation to a gaze manipulation. The study manipulated direction of gaze in a computerized face when a target letter appeared either in a gaze-congruent or in a no-gaze-incongruent location as a first factor. A second factor was the delay on each trial between the presence of gaze and the following target letter; stimulus onset asynchrony (SOA) was 100, 300, or 700 milliseconds. Results of the reaction time analyses showed a main effect of congruency, indicating faster reaction time on congruent trials. However, the effect was only significant at the 700 ms SOA. The authors concluded that the direction of gaze had a reliable effect on the letter-discrimination task, but they had expected an even earlier gaze effect. To test the delayed effect the authors performed a second experiment in which a face looking ahead was seen before the gaze cue picture. This was done because the authors thought that the reason for the delayed time-course of the cueing might have been the sudden appearance of the face cue at the centre of the screen. Presenting the neutral face for 900 ms before the gaze cue onset resulted in a significant effect already in the

300 ms SOA condition. The authors suggested that the gaze information is not robust and immediate. Whether we can or can not follow a gaze cue may depend on how long we have experienced the face/gaze beforehand (even in an uncued position).

Another reaction time study by Friesen, Moore, and Kingstone (2005) further tested the gaze cue effect. They compared a so-called 1-object condition with a 2-object condition (of which more below). Each trial began with a blank face and then a face with gaze cue appeared. Participants were instructed to keep their eyes on the nose on the computerized face in the scene and to press a button as quickly as possible when a target *square* was visible on the left or right side of the face. During half of the trials the face gaze cue was present to left or to right and during the other half of the trials no gaze cue was present. The target occurred in four SOA conditions from 105 to 1005 ms. In the 2-object condition the target (a square) occurred on the one side of the face, and a distractor (a circle) occurred on the other side of the face. The data showed that the response time was faster in the 1-object condition than in the 2-object condition and in both conditions reaction time was faster during the cue appearance than when no cue had been present. The response time decreased in line with the lengthened SOA condition. These findings demonstrate the effects of the gaze cue also in a somewhat more complex context when more than one object is presented in a scene.

Moreover, an eye-tracking study Mansfield, Farroni, and Johnson (2003) tested a person's gaze cue effect while participants first saw a face looking ahead and then shifting gaze to the left or to the right. Afterwards a target appeared either on the left or on the right side of the presented face. Even though participants were instructed to look at the middle of the display and then look at the target as soon as it would occur, their gaze shifted in the direction of the gaze cue prior to the target onset. They inspected the target significantly more when it was cued than when it was uncued.

Another experiment investigated the question of whether seeing two people (faces) engaging in eye contact could also modulate gaze following (Böckler, Knoblich, & Sebanz, 2011). First, participants saw two human faces with closed eyes. Next, they saw the faces looking at either each other (shared attention) or away from each other. The next scene showed the faces gazing at a target object (congruent) or not gazing at a target object (incongruent). Participants' task was to respond with a button press as soon as they saw the target object. Results showed an

overall significantly faster response in the shared attention condition than when the faces looked away from each other. However, the shared attention effect was not found in the incongruence trials. The results of this study suggest that humans are not only sensitive to the gaze cue when they are direct addressees but also in a shared interaction where observing others? gaze interactions modulated human reaction times (Böckler et al., 2011).

Furthermore, gaze cue effects have been tested in a social interaction task by Macdonald and Tatler (2013). Participants were sitting at a table facing an instructor and they were instructed to build blocks of mega blocks. Participants were not informed about whether the instructor would provide a gaze cue or not. Participants' eye-movements were recorded while they performed the task either in a gaze or no-gaze condition. The gaze data showed that participants were more accurate in selecting the target blocks when a gaze cue was available compared to when it was absent.

The effects of a gaze cue on a cognitive level in language comprehension has been investigated in the following studies. Hanna and Brennan (2007) examined the time and flexibility with which gaze cues in speaker/listener pairs can be used for a target-matching task. The authors tested how fast listeners would follow a speaker's gaze when an ambiguous linguistic utterance is integrated and whether the gaze affects visual attention and language processing automatically or whether it is used in a flexible manner. Participants sat in pairs facing each other and looking at their own set of objects. Each set contained six objects and other spaces that were labeled A, B, and C. Participants were randomly assigned their roles as a speaker and as a listener to move some objects on the display. Listeners' eye gaze was recorded by using a head-mounted eye tracker and the camera of the eye-tracker captured the eye gaze, to find out whether and how rapidly listeners followed the speaker's gaze. Results showed that listeners used the speaker's gaze to inspect the target object before the linguistic disambiguation. Speakers' gaze helped interlocutors to start inspecting the target more often than the competitor before the linguistic disambiguation.

Other studies by Knoeferle and Kreysa (2012) examined effects of a speaker's gaze on a listener's visual attention and language comprehension. The authors investigated whether the speaker's gaze is only robust when the listener is able to see both eyes of his interlocutor or whether gaze effects also emerge when the speaker is positioned at an angle. Participants' eye-movements were recorded when they

listened to a German spoken SVO or OVS sentence and they viewed a video either in a gaze condition or in a no-gaze condition. The results showed that the participants used the gaze of the speaker and the target object was anticipated shortly after the speaker had gazed at the target and before its mention. The data also showed that speaker gaze effects were larger for SVO than OVS sentences. Notwithstanding, the gaze effects in target anticipation generalized to non-canonical sentence structures and situations in which the speaker did not face the listeners.

Another recent study investigated gaze cue effects in human-robot interaction (Staudte et al., 2014). The authors examined whether an agent's gaze cue can be relevant during sentence processing, and whether listeners will use the gaze cue for understanding the utterance. In their study, participants saw a scene including an agent, an egg, a box, and other objects on a table in front of the agent. Participants' eye-movements were recorded while they listened to a sentence such as 'The egg is taller than the box'. During the first noun phrase 'the egg' the agent was either in a neutral position looking straight ahead or in a congruent position gazing at the egg, or gazing in a reverse order at the objects (i.e., looking at the box while 'the egg' was mentioned). The result showed a clear effect of the agent gaze cue. Participants during the NP1 inspected the egg significantly more in the congruent condition than in the reverse or neutral conditions. Similarly, the box was looked at significantly more in the reverse condition (i.e., when the robot first gazed at the box) than in the neutral and congruent conditions. In another experiment the authors decreased the number of objects to four to facilitate noticing the gaze cue and the agent was near to the objects. In that later study participants were faster to look at the box after its onset in the sentence than in the first study. To conclude, when the gaze cue was more precise it affected the listeners' gaze earlier.

The current section discussed evidence of gaze cue effects on visual attention and language comprehension. The main findings are that participants use gaze cues very rapidly in directing attention towards target objects. Furthermore the gaze effect depends of the previous experience of the gaze/face. Aside from an immediate gaze cue effect reported when the speaker/listener directly face one another, gaze effects on listener attention were also observed when the speaker did not directly face the listener. Thus, listeners followed a speaker's gaze to inspect a corresponding target referent in the scene before it was mentioned. Participants

partly relied more on the gaze cue than the language information during the incongruence condition (e.g., in the experiment by Staudte et al., 2014). Incongruence between language and the visual world has previously been examined (e.g., in sentence-picture verification, see section 2.4). The present thesis draws on this research for stress-testing the recent-event preference. Experiment 5 in the present project addresses this question. In the next section we review findings on the picture-sentence incongruence paradigm during language comprehension.

2.4 Verb-Action Incongruence

Previous studies have shown that incongruence between visual and language information affected language comprehension. The following studies discuss findings in the field and reveal comprehenders' sensitivity to picture-sentence incongruence in response times in a picture-sentence verification task (e.g., Carpenter & Just, 1975), in eye-movement recordings (e.g., Dumitru, Joergensen, Cruickshank, & Altmann, 2013; Knoeferle & Crocker, 2005; Underwood, Jebbett, & Roberts, 2004) and in event-related brain potentials (e.g., Knoeferle, Urbach, & Kutas, 2011).

In two eye-tracking experiments, Underwood et al. (2004) investigated picture-sentence verification. The authors wanted to know how comprehenders integrate visual depictions with written language. In their experiments participants were presented with visual and written sentence information at the same time (experiment 1) or sentence and scene were presented serially (experiment 2). Participants' task was to judge by pressing a button whether the sentence matched the scene or not. In experiment 1 they first saw a fixation marker for 1000 ms and then a picture and a sentence appeared either in the matching or mismatching condition. The results showed a longer response time in the mismatch than in the match conditions. The reaction time results were confirmed by further statistical analyses; more fixations occurred on the mismatching than matching display and the sentences received more inspection than the pictures. The second experiment changed the procedure in presenting sentences and pictures serially. Results revealed more accurate responses when the sentence had been presented first and this was confirmed by statistical analyses. Furthermore, the results of the total inspection time showed more fixations to the pictures than to the sentence when the picture appeared first. Overall the results in the second experiment showed an

effect of the picture/sentence order but in contrast to experiment 1 failed to find an effect of the picture-sentence incongruence.

Another experiment by Knoeferle and Crocker (2005) also used the eye-tracking method to examine sentence-picture incongruence. In addition to the incongruence this experiment manipulated the sentence word order (SVO/OVS). Participants would always first see a scene with two characters of which one character performed an action. Afterwards they read a sentence that either matched or mismatched the action scene. By contrast to Underwood et al. (2004) the authors analysed the data for each word region of the sentence. They reported an incongruence effect in the verb and adverb regions but no effects emerged in the NP1 and in the NP2 regions. Similarly to Underwood et al., the total sentence inspection data did not show any reliable incongruence effect.

In another eye-tracking picture-sentence verification study by Dumitru et al. (2013) participants' eye-movements were recorded while they were presented a picture with two characters on a display. While viewing the pictures they listened to either a conjunction (and) or a disjunction (or) sentence e.g., *Nancy examined an ant and/or a cloud*. Participants' task was to inspect the picture and to listen to the sentence and then to indicate via a button press whether the sentence matched or mismatched the picture. The matching condition showed both characters on the picture; by contrast the mismatching condition displayed one character or none of the characters. Mismatch effects emerged in the response accuracy and response time data. Participant were faster in responding to the mismatch than the match condition. Similarly, the overall gaze data revealed more fixations in the mismatch condition than in the match condition. However the fixation, during the second noun phrase were more in the match than in the mismatch condition, indicating an earlier detection of the mismatches condition (relative to the matche). Concerning the sentence type, participants were significantly faster in the conjunction than disjunction sentences.

Furthermore, an ERP study by Knoeferle, Urbach, and Kutas (2011) examined verb-action incongruence. Participants first saw a picture display with two character (a gymnast and a journalist) involved in an action. After inspecting the action, participants pressed a button to read a sentence word by word. Each word was presented for a duration of 200 ms. At the end of the sentence participants had to reply as quickly as possible whether the sentence matched or mismatched the picture. Results showed a reliable accuracy effect of understanding better the

matching than mismatching sentence-picture pairs. Similarly the reaction time data revealed a faster response in the match than in the mismatch conditions. Overall, this section has provided evidence that incongruence between a visual cotext and language can rapidly affect language comprehension.

At this stage we sum up the findings presented in this chapter. We started by presenting early findings in the visual world paradigm, followed by a discussion of effects of the visual context, as well as of linguistic and world knowledge on incremental language comprehension. Specifically, we discussed findings that provided evidence for the view that aspects of the visual context such as depicted actions, colour, size or location of an object play an important role for real-time language understanding. Furthermore we reviewed studies on the effects of tense and verb-related information during language understanding. Moreover, we gave an overview on the importance of attention for later memory. We have discussed the role of frequency, of situational gaze cues, and of incongruence manipulations for language processing. Before we begin with the experimental chapter let us recall the relevant findings that have motivated the present project questions, namely the issue of the recent-event preference.

The Recent-Event Preference

Recall the findings by Knoeferle and Crocker (2007); In their experiment when participants only saw a recent clip-art action and they listened to a sentence that either referred to the recently seen event (and its target) or to an event that likely occurs next (on a different target object), they preferred to look at the target of the recent event in both past and future tense sentence conditions. However, perhaps this strong preference emerged because of the semi-real nature of the stimuli (clipart depictions). Perhaps such semi-real depictions discourage the anticipation of future events, and exaggerate reliance on what is depicted. Ensuing experiments (Abashidze at al., 2011; Knoeferle et al., 2011) set out to replicate the recent-event preference with real-world action events. Interestingly the results replicated the findings from the clipart experiment and suggesting an even stronger recent-event preference from the real-world action, whereby participants preferentially looked at the recent action target independent of sentence tense (present tense with a future meaning or simple past tense). A clear preference to look at the plausible future object surfaced only at the offset of the last word (NP2, when the future target was named). These results replicated the recent-event preference but did not clarify why participants preferred to inspect the target of the recent action event over the target of a plausible future event. Even though participants did not see any future target acted upon in this experiment they did not have reason to direct their attention towards the recent event target, at a point in time during the sentence when they heard a present tense verb and an adverb with a future meaning. Moreover they had never seen the recent event target acted upon again in a given trial after the sentence presentation. It is thus unclear what causes the gaze pattern associated with the recent-event preference.

The authors (Abashidze et al., 2011; Knoeferle, Carminati, et al., 2011, experiment 1) suggest that one of the causes of the recent-event preference could be a frequency bias (that future events were never acted out while recent events were acted out). Acting out the recent event more frequently could have encouraged listeners to rely more on it and its target object during comprehension than on other objects related to the plausible future event. Therefore the experiment 2 by Abashidze et al. (2011) showed the recent and the future actions equally often. Participants saw a recent action before the sentence and a future action after the sentence. This frequency distribution was meant to ensure that the preferred inspection of the recent action target did not originate solely from an imbalance in

frequency of recent and future actions. However, surprisingly the results in this experiment replicated previous findings of the recent-event preference and of the delayed inspection of the future event target. The inspections of the future event target only emerged from the middle of the second noun phase.

Existing accounts of situated language comprehension predict a rapid interplay of language comprehension, (visual) attention, and visual context effects. On the one hand some studies have revealed that comprehenders anticipate an object in a visual scene through verb information, preposition meaning or through immediate depicted actions (e.g., Altmann & Kamide, 1999; Altmann, 2004; Chambers et al., 2004; Knoeferle & Crocker, 2006). Other studies found that participants process an ongoing event faster than a completed event (Madden & Therriault, 2009). Eye movements are modulated by a description of events when that description clashes with the actual depiction (Altmann & Kamide, 2009). Given that comprehenders have a preference towards inspecting the object of the recent event (Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011) maybe there exists even an epistemic reason for this behaviour (MacFarlane, 2003; Staub & Clifton Jr, 2011, for discussion). In order to resolve this question, the present thesis project stress-tests the behavioural phenomenon dubbed *the recent-event preference* and investigates to what extent the preference can hold against a frequency bias towards the future event, against a powerful situation-specific cue (the actor's gaze), and against an incongruence between visual and language context (i.e., the recent action event mismatches the past tense verb). In first two experiments reported in the following chapter, we changed the frequency distribution (the ratio of recent to future events) within the experiment.

Chapter 3

The Role of Frequency in Language Understanding

This chapter presents the first two experiments of the thesis. These experiments examined the effect of the frequency distribution of recent compared with future events during spoken sentence comprehension. We were interested in effects of increased numbers of future events and future tense sentences that biased against the recent-event preference. The changes in the frequency distribution were motivated from findings by Knoeferle, Carminati, et al. (2011) and other findings that have shown effects of frequency in language processing (see Chapter 2). As described at the end of chapter 2 a first study ensured that equally many future and recent events were acted out. Compared with previous studies in which only the recent events had been acted out, the study with a balanced frequency distribution of future and recent events had reported a somewhat earlier inspection of the future event target (Knoeferle, Carminati, et al., 2011). It is an interesting finding in this regard as the number of the events that were presented before or after a sentence affected participants' eye-movements in language processing. Participants continued to mostly inspect the recent event target until late in the sentence.

Other studies, those have manipulated the frequency, found that humans are sensitive to statistical regularities in both language processing and visuomotor activity. On a word level, findings showed that a word occurring often in language was comprehended much faster than a less frequent word (e.g., Morton, 1969). On a phrase level, a high frequency phrase has been responded to faster than a low frequency phrase (e.g., Arnon & Snider, 2010). Overall, short-term language experience can

also modulate language production (e.g., Haskell, Thornton, & MacDonald, 2010; Kaschak, Loney, & Borreggine, 2006) and sentence reading (Wells et al., 2009). Moreover, in an action execution task participants decision in hand movement has been affected by the recent trial-to-trial visuomotor experience (e.g., Chapman et al., 2010)

Recall that the main focus of the present project is to test the robustness of the recent-event preference. The question is, why participants did not anticipate the plausible future event target based on the tense cue (verb + adverb) while other findings revealed that verb tense permits comprehenders to predict post-verbal information. One possibility of this effect could be a well-known priming effect. For example, if we recall the study by Chambers et al. (2002) in which participants were presented with an instruction e.g., *Pick up the X ... Now put it inside the Y*. On hearing *Now* their eyes mostly fixated the location of the X object, from where it was picked. Furthermore, in a visual search task participants inspected a target quickly when they had seen it in a prior scene (Körner & Gilchrist, 2007). Moreover, in the experiments by Abashidze et al. (2011) participants inspected the recently acted upon object during the entire sentence presentation in the past as well as future tense condition. Recall that the sentence in the future tense condition referred to a plausible future action event.

Another reason for the lack of a future target inspection could be the greater uncertainty associated with future than recent events. While during the past tense condition participants are certain about the information presented this is not the case during the future tense condition at least until the second action event is shown (e.g., MacFarlane, 2003; Staub & Clifton Jr, 2011). However, other findings have revealed anticipatory eye-movements during the post-verbal noun phrase (e.g., Altmann & Kamide, 1999, 2007). Furthermore, It might be even an epistemic bias that exists in favour of the recently perceived visual cues. Other findings have also shown that participants relied more on information of depicted actions than on their world knowledge (e.g., Knoeferle et al., 2005).

The question we wanted to test in the following two experiments was whether by increasing the frequency of future events and future tense sentences, we can change participants' eye-movement behaviour towards the recent/future event target. Will the increased number of future events create a high certainty that the future event will happen, leading to early inspection of the future event target? Will frequently seen future events mitigate the robustness of the recent-event preference or will the

plausible epistemic bias remain despite the future event bias? Prior research has shown that frequency biases can over time create biases in cognition and action (e.g., Kaschak et al., 2006) and we were wondering whether frequency might have a similar effect in the present experiments.

If the frequency with which we experience future and recent events has an effect, much like frequency biases have had an effect in other studies (see chapter 2), then we should see an earlier inspection of the future event target than in experiment 2 by Knoeferle et al. (2011). As seen in the second experiment by Knoeferle et al. (2011), eye-movements to the future target increased only after the onset of the post-verbal noun phrase. By contrast, in the present studies we should see an earlier shift in the listener's gaze to the future event target (e.g., at the point in time when tense information is disambiguated which is at the end of the verb and during the adverb).

The present experiments have been built on the studies as by Abashidze et al. (2011). In contrast to these experiments, in the present experiments we used video materials and we examined the effects of frequency biases during spoken sentence comprehension. In addition, we wanted to know whether seeing a recent event and hearing it named modulates merely the immediate focus of attention or whether it also benefits memory and ensuing recall of the action. Experiment 1 increased the frequency of future events and future tense sentences. However, it still showed the recent action event on each trial. Experiment 2 provided further increased numbers of future tense sentences and future action events. In addition, the fillers in Experiment 2 did not show a recent action event, which further decreased its relevance within the experiment. At the end of each eye-tracking experiment, a memory test was conducted on the experimental visual stimuli.

3.1 Experiment 1

Experiment 1 investigated the effect of a strong frequency bias towards the future and thus against the recent event preference during spoken sentence comprehension. Similar to experiment 2 by Abashidze et al. (2011), sentences were either in the past or in the future tense and described events. In comparison to experiment 2 in Abashidze et al. (2011), the present experiment increased the frequency of future events and future tense sentences to 75%. This was achieved by increasing

the frequency of the filler trials in the future tense condition. We should note here that the frequency of recent and future action events of critical trials was equal. Participants listened 75% of the time to future tense sentences that all referred to the future event target. By contrast 25% of the sentences were in the past tense and referred to the recent event. In contrast to other studies that have examined the influence of frequency biases at the word, phrase, and sentence level, our experiments have created biases in the frequency of past and future tense sentences involving the visual context input. To the best of our knowledge there is no study that has created biases in the frequency of past and future tense sentences in a visual world paradigm experiment.

In comparison to findings in experiment 2 with (50/50%) (Abashidze et al., 2011; Knoeferle, Carminati, et al., 2011), eye-movements in the present experiment should show early anticipation of the upcoming event target. The expectation of the early gaze pattern to the future event target could be elicited by the instantiated frequency bias and it should emerge especially in the second part of the experiment. Prior research has shown that the learning of statistical regularities grows in strength over time (e.g., Kaschak et al., 2006). Participants' certainty that the future event will occur might thus increase and their gaze behaviour towards the future event target might reflect this interest. For example the future event target is not known while they listen to the sentence. If they are curious about the not yet seen future event they might pay more attention to potential future targets compared to targets of the recently seen event. As a result, the recent event target inspection should decrease and the overall preference might even be eliminated throughout the sentence.

Alternatively, it could happen that the recent-event preference replicates because what we already saw and listened to, is more prominent in our immediate experience. Furthermore, the recently seen event might be strongly activated in working memory and its representation might increase visual attention to the target of the recent event. Such attentional guidance would support subjects' preferential looks towards the recent target object compared with an equally plausible future event target. In addition, participants inspect the previous location of an object when a sentence describes an event in which the object could be a plausible target. This occurs even when the screen they inspect during language comprehension is blank. The gaze behavior in the blank screen (inspecting the location of an absent object) suggests that people are relying on a representation of the preceding

scene during language comprehension. Such a description can influence participants' eye-movement such that they inspect the location of previously-seen objects (Altmann, 2004).

To gain more insight into how humans use event information in the light of the frequency bias for later memory, we asked participants to perform a memory test on the visual stimuli from our experiment. We might see more looks to the future event target than to the recent event target since the future events are more frequent. If so, this might also affect ensuing memory performance. A recognition memory test revealed faster recognition for high probability stimuli (O'Brien & Raymond, 2012). Additionally, participants might remember future action events better because of a recency effect (they see the future actions performed last during each trial). In this case the most recent items would be the future event target objects. In the memory paradigm the most recent items have been recalled better than other non-recent items (e.g., Glanzer & Cunitz, 1966). Recency also has an influence on working memory (e.g., Zelinsky & Loschky, 2005).

Alternatively, however, if recent (vs. future) events are saved first in working and then in short-term memory, participants should recall the recent events better than future ones. This expectation relies on the logic that what has been encoded first would be encoded better than later information. D'Argembeau and Van der Linden (2004) examined mental time travel - participants experience associated with remembering past and imaging future events. The results showed the imagined future events were recalled with less detail than the past events. Then we might also expect better recall of the recent than future events in the memory task. This would suggest a clear relation between attention during comprehension and ensuing memory.

3.1.1 Method

Participants

Thirty-two German native speakers (Mean age=25.25; range: 19 to 32; 12 males, 20 females, all of them right-handed) took part in the experiment. All participants were from the Bielefeld University community, and they were each paid 6 Euros for their participation. None of them had learned a second language before age 4.

All had normal or corrected-to-normal vision and were unaware of the purpose of the experiment.

Materials and Design

Eye tracking: The experiment used twenty-four experimental items. Half of these items (12) were the same as those used in experiment 2 by Abashidze et al. (2011); Knoeferle, Carminati, et al. (2011), and an additional 12 new items were constructed using similar criteria. Each item consisted of two everyday objects (e.g., cucumbers and tomatoes, see Figure 3.1) and four sentences (see Table 3.1., including literal translations). All critical sentences had the structure NP-V-ADV-NP and two male native German speakers recorded them. Sentences lasted 3200 ms on average. The first noun phrase onset was at 0 ms. 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 temporal adverb (*demnächst*, 'soon') indicating the future (Table 3.1., 1a-a'). In the other condition, the verb was in the simple past, and the following temporal adverb (*kürzlich*, 'recently') also indicated the past (Table 3.1., 1b-b'). Only German regular verbs were used in the critical sentences. As can be seen in Table 3.1, there were two sentences for each tense condition; we created the counterbalancing versions (1a' and 1b') to ensure that each object was once the target of both a recent and future action event. 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, Piepenbrock, & Gulikers, 1995).

For every item we recorded two videos, the duration of videos was 5015 ms on average. As can be seen in Figure 3.1, the scene in 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. Figure 3.1 shows the order in which the videos were presented in a typical critical item trial. The first video showed the person performing an action on one object (e.g., flavoring cucumbers, Fig. 3.1., a) and the second showed the person performing the same action on the other object (e.g., flavoring tomatoes, Fig. 3.1., 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. 3.1., 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 3.1 (a-c).

TABLE 3.1: *Example experimental sentences. The indices (') indicate counterbalancing versions*

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

In addition to the critical items we created a number of filler items (N=40). The frequency bias was instantiated by having most filler items associated with future tense sentences (e.g., *Der Versuchleiter zeichnet in der nächsten Zukunft ein Haus auf dem Notizblock*, 'The experimenter will draw in the near future a house on the notebook'). A difference between experimental and filler trials was that in the filler trials participants were mostly presented future tense sentences referring to the future action. Filler sentences never used any words that appeared in the critical sentences and the objects used in the fillers were not shown in the critical videos. Thus participants saw in total 64 items (24 critical + 40 fillers). All experimental items showed both actions, a recent action before the sentence and a future action after the sentence. Half of the sentences were presented in the past tense (12) and the other half in the future tense (12). Of the 40 filler trials 36 showed the future and recent actions and only presented a future tense sentence, and the remaining 4 showed the recent and future actions and only presented a past tense sentence. In total, each participant thus saw 48 future vs. 16 recent action trials over the course of the experiment (75% vs. 25%).

We expect that the frequency bias will elicit more early eye-movements towards the plausible future event target. When participants hear the first noun phrase

'The experimenter', we expect they will mostly look at the experimenter. Both past and future tense sentences started with the same noun, for that reason we do not expect any differences in looking at the actor during the first noun phrase between the past and future tense condition. It is possible that we will find early looks to the recently-acted upon target (since it is the last thing participants were looking at before the sentence starts). On hearing the verb and the temporal adverb 'flavored recently' in the past tense condition, at the offset of the verb it becomes clear that the verb refers to the recent event target and listeners will look more at the recent event target (the cucumbers) than at the plausible future event target (the tomatoes).

However, on hearing the verb and the temporal adverb 'flavors soon' (the verb in the present tense plus the temporal adverb in the future tense) in the future tense condition, listeners might not immediately look at the future target object at the beginning of the verb region but at the end of the region they might start to direct their attention towards the future target object. An effect early during the verb is unlikely because the verb is -up to the last letter- ambiguous between the simple past and the present tense. However after a few future trials participants could notice that the verb refers to the future event and its target object on most trials. As a result, more eye-movements should be made to the future target object (the tomatoes) than to the recent target object (the cucumbers) shortly after the verb offset and immediately at the beginning of the adverb region in the future tense. The probability of this expectation should be increased by the frequency bias throughout the experiment. Another key reason for not inspecting the recent event target in the future tense during the verb and adverb is that a second action on the same object (from the recent action) has never occurred after the sentence.

When the second noun phrase 'the cucumbers' of the past tense sentence unfolds, more eye movements should go to the recent target, reflecting that participants process its name. Furthermore, significantly more eye-movements should be made to the recent target (the cucumbers) in the past tense than in the future tense condition. On the contrary, when the second noun phrase 'the tomatoes' in the future tense sentence is heard, the fixations on the tomatoes should increase and this should happen significantly more often in the future tense than in the past tense sentence. Crucially, the frequency of the future tense should decrease the recent-event preference and even override its overall bias.

The experimental and filler items were combined to form 4 lists using a Latin square design. Each list contained every critical item in only one condition and all fillers (trials in each list N=40). Before the experiment, lists were pseudo-randomized and each participant saw an individually randomized version of one of the four experimental lists.

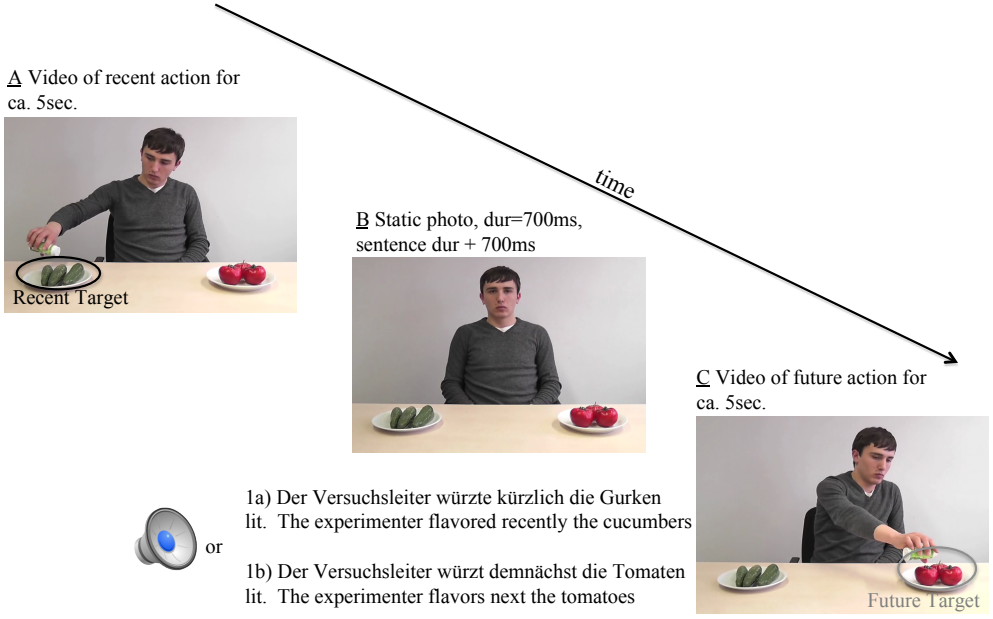
Memory test: For the memory test we created two snapshots of the first and second video of each experimental item, i.e., showing the actor performing one of the two action events (see Fig. 3.2). The two snapshots associated with each item were combined into one display. Two versions were created in which the location of the two pictures was counterbalanced. The purpose of the memory test was to assess participants' memory of the action events that they had seen in the experimental video sequences during the eye tracking session. In particular, we wanted to see whether actions that were seen before the sentence was heard (i.e., the recent actions), were remembered better than actions that were seen after the sentence was heard (i.e., the future actions). We expect a better memory performance in the future tense condition in case the frequency with which participants perceive an event plays an important role for short-term memory. However, if the recent events are anchored more firmly first in working memory and then in short-term memory, the results of the test should show a better recall of the of recent (vs. future) action events.

Procedure

Eye-tracking: Participants were seated in front of an Eyelink 1000 desktop head-stabilized eye tracker (SR Research) and the experimenter calibrated participants right eye with a 9 point dot pattern. Before the experiment, participants read the experiment instruction on the computer display. They were informed that the experiment consisted of an eye tracking experiment followed by a short memory test. During the eye tracking experiment, they should 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. For a given experimental trial, the timing of the video sequence is illustrated in Figure 3.1. Participants saw a video of a person performing one action before the sentence (e.g., flavoring the cucumbers) for ca. 5000 ms. (see Fig. 3.1., a) and then the static picture appeared. After 700 ms the sentence was played. The static picture remained on the screen until 700 ms after the end of the sentence (see Fig. 3.1., b). The sentence presented (see Table 3.1., 1a or

1b) was either *Der Versuchsleiter würzte kürzlich die Gurken*, 'The experimenter flavored recently the cucumbers', or *Der Versuchsleiter würzt demnächst die Tomaten*, 'The experimenter flavors soon 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) for ca. 5000 ms. (see Fig. 3.1., c). In the middle of the experiment participants had a short break.

FIGURE 3.1: *Sequence of events of an experimental trial, Experiment 1 and 2*



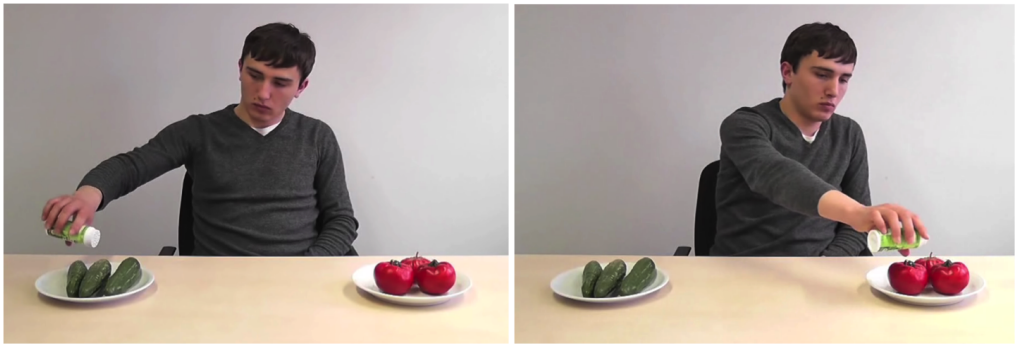
Memory test: After the eye tracking session ended, participants were asked to perform the memory test. During the memory test, participants saw pictures such as those in Figure 3.2 one for each experimental item. Above the picture, a question appeared, which was in one of two versions:

(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. Participants were assigned randomly to four counterbalancing lists and each saw a randomized order

of the list. After the memory tested ended, participants were debriefed. The experiment lasted approximately 45-50 minutes.

FIGURE 3.2: *A snapshot from the experimental video for the memory test*



Analyses

Eye-tracking: For the coding of participants' eye-movements during the experimental trials, a period of interest was defined, starting from the onset of the Verb until the offset of the post-verbal noun phrase (NP2), i.e. the end of the sentence. As for the location of the fixations on the display, we defined two areas of interest: the recent target object and the future target object. In addition there was a third interest area of the actor but it was not included in the analyses. The measure of interest for the aim of the experiment 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 slot, 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 graph shown in Figure 3.3. This graph shows participants' fixations to the two objects of interest from the onset of the Verb until the end of the second noun (NP2) in the two conditions. The dotted lines indicate the recent condition (the sentence was in past tense) and the solid lines indicate the future condition (the sentence was in future tense). In addition, we analysed the gaze data using inferential statistics. Since the aim of the ANOVA

analyses was to analyze the gaze data during individual words in the sentence, we defined the following three time windows. The Verb region (from verb onset until adverb onset, Mean duration = 1148 ms); the Adverb region (from adverb onset until NP2 onset, Mean duration = 1332 ms) and the NP2 region (from NP2 onset until NP2 offset, Mean duration = 710 ms).

Fixations were counted for the Verb region if they started in that region. For the Adverb and NP2 regions they were counted if they fell into the regions. The mean log gaze probabilities ratios (ln (P(recent target)/P(future target))) were aggregated over each of the three time regions of interest. An advantage of using log-ratios (additionally 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).

ANOVAs were performed on the log-ratios of each of the time regions. Separate ANOVAs were conducted on log ratios averaged over participants ($F1$) and items ($F2$) separately (see Table 3.3). The dependent variable was the mean log gaze probability ratio and the independent variable was tense, with two levels, past and future tense.

Furthermore, we conducted a post-hoc one-sample two-tailed t-test on the log ratios by participants and items. The goal of the test was to find out to which extent the recent-event preference persisted in the future tense condition. The test aimed at showing whether the log ratio means in the future tense condition were significantly different from zero. Recall the formula (ln (P (recent target)/P (future target), whereby a negative log-ratio that is reliably different from zero indicates a higher probability of inspections of the future than recent event target; a positive log-ratio that is reliably different from zero indicates a higher probability of inspection to the recent that future event target.

Memory test: We calculated the percentage of correct answers by condition for participants and items separately. The average percentages (by participant) are illustrated in Figure 3.4. For the further analyses of the memory test results a logistic Linear-Mixed-Effect (LME) model (binary: correct vs. incorrect) was fitted to the response data of the memory test. In this model the outcome was the response and the predictors were tense (past vs. future) and target objects

of events (recent vs. future action events). Participants and items, with their intercepts and slopes and the intercept x slope interactions, were included in the random effects of the model. The predictors were centered by transforming the fixed coding into a numerical value and centering it so as to have a mean of 0 and a range of 1 (Baayen, 2008). This effect coding has the advantage of reducing collinearity and allows the coefficients of the regression to be interpreted as the main effects in a standard ANOVA (Barr, 2008).

3.1.2 Results und Discussion

Eye-tracking: As can be seen, the graph (Fig. 3.3) shows an overall preference for looking at the recent target relative to the future target throughout the Verb and until the middle of the Adverb region. Participants inspected the recent target object for about 1900 ms from the Verb onset onward in both tense sentence conditions. During most of this period the log ratio for both the past and future conditions remains well above zero (indicating that the recent target receives more looks than the future target). Interestingly, in the beginning of the Verb preferential inspections towards the recent event target increased rapidly 400 ms after the Verb onset and the preferential inspection towards the recent target lasted throughout the Verb region. Listeners increased eye-movements towards the recent event target occurred irrespective of tense and despite the strong frequency bias towards future events. Importantly though, the two lines began to separate towards the end of the Verb region, which shows an earlier effect than in the experiment with balanced proportion of future and recent events (Knoeferle et al., 2001, Exp. 2). When participants hear the Verb in the future tense condition at the end of the Verb region they decrease their gaze toward the recent target. From the middle of the Adverb region participants inspected the future target object in the future tense condition more than in the past tense condition and correspondingly they looked more towards the recent target object in the past than future tense condition. This pattern increases steadily throughout the Adverb and the NP2 region. In the NP2 region, the future target is clearly preferred to the past target in the future tense condition (the log ratio is negative, -.926) throughout this region (see Table 3.2). The main finding, however is that despite the strong frequency bias toward the future condition and despite the earlier inspection of the future event target (by approximately 1000 ms compared with Experiment 2 by Knoeferle et al., 2011), the grand mean (see Table 3.2) in the present experiment

remained positive in all three regions, indicating an overall preference towards the recent event target.

FIGURE 3.3: *Mean log gaze probability ratios (ln (P(recent target/P(future target))) by condition from verb onset, Experiment 1*

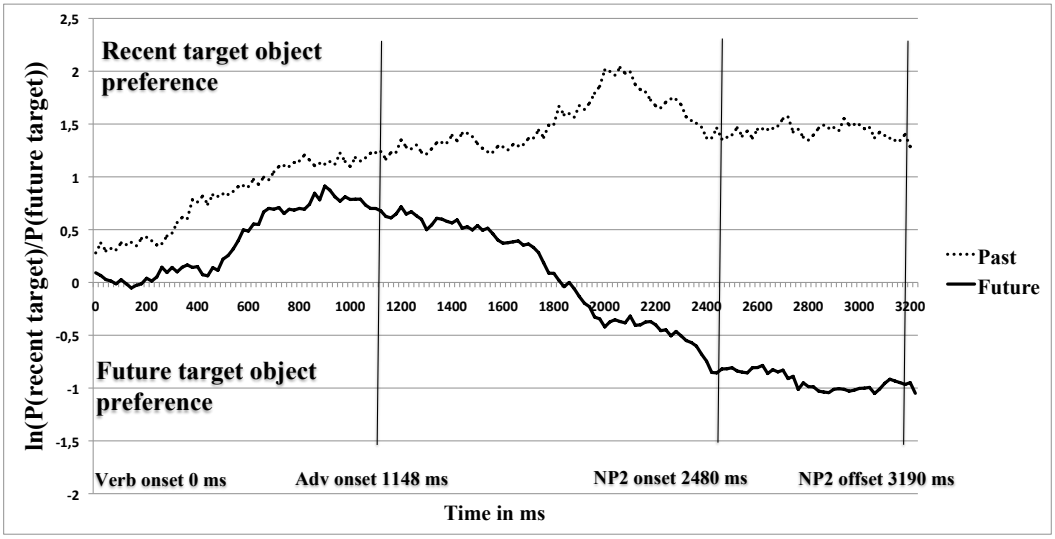


TABLE 3.2: *Grand mean and mean log gaze probability ratios (ln (P(recent target)/P (future target)) by participants as a function of condition and time region for the experiment. Standard errors (SE) in parentheses, Experiment 1*

Regions	Future tense	Past tense	Grand mean
Verb	.494 (.15)	.969 (.24)	.73 (.15)
Adv	.353 (.21)	2.13 (.23)	1.24 (.17)
NP2	-.926 (.17)	1.93 (.23)	.50 (.12)

We now turn to present the inferential analyses. Importantly, the conclusions emerging from the descriptive analyses above were mostly corroborated by the inferential analyses of the data. While the log ratio showed a positive value in all three regions in the time course presentation, the inferential analyses supported this with significant intercepts in all three regions by participants and items *ps* $< .000$ (see Table 3.3). For the Verb region, the ANOVA showed a marginal tense effect by participants $F1(df{=}1.31) = 3.18$, $p1 = .084$ and a significant effect by items $F2(df{=}1.23) = 11.48$, $p2 = .003$. This study replicated a not fully significant effect of tense for the Verb region, much like previous studies (i.e., Knoeferle, Carminati, et al., 2011). Importantly though, we can see that there is a significant tense effect in the Adverb and NP2 regions. The tense effect in the Adverb region was reliable by participants $F1(df{=}1.31) = 16.24$, $p1 = .000$

and by items $F2(df{=}1.23) = 54.66$, $p2 = .000$. The NP2 region further shows a reliable tense effect by participants $F1(df{=}1.31) = 73.26$, $p1 = .000$ and by items $F2(df{=}1.23) = 163.51$, $p2 = .000$ (see Table 3.3). Overall, the results revealed that participants made more inspections to the recent than future target object in the past tense than in the future tense conditions. This effect was reliable in all analyses (except by subjects in the Verb region).

On the one hand, the significant intercept in all three regions could be due to a possible epistemic bias towards the recent event target. We hypothesised that if the epistemic bias exists, participants might inspect the recent event target more even with a strong frequency bias in favor of future events. However, on the other hand, the frequency distribution (75/25%) in the present experiment revealed its effectiveness and resulted in an earlier inspection of the future event target (approximately 1000 ms earlier than in experiment 2 by Knoeferle et al., 2011, when frequency of recent and future events was balanced). Furthermore, the present manipulation revealed a fully significant tense effect in the Adverb and NP2 regions (but only a marginal tense effect by participants in the Verb region).

TABLE 3.3: *ANOVA analyses for the data of Experiment by region: The intercept is also given because in this case a significant intercept indicates that the grand mean is significantly different from 0, Experiment 1*

Regions	**Effect**	**F1 (df=1,31)**	**F2 (df=1,23)**	**P1**	**P2**
Verb	Intercept	23.20	44.66	.000	.000
	tense	3.18	11.48	.084	.003
Adv	Intercept	52.57	66.19	.000	.000
	tense	16.24	54.66	.000	.000
NP2	Intercept	16.51	14.88	.000	.001
	tense	73.26	163.51	.000	.000

Further, the t-tests presented in table 3.4 reveal that the means in the Verb and Adverb regions are positive (indicating a preference for the recent target, which was reliable in the Verb region only). The means are negative in the NP2 region and the effect is significant but reversed (indicating a preference now for the future target). While the significant effect in the Verb region confirms the robustness of the recent-event preference, the effect in the NP2 regions affirm the effect of the frequency bias.

TABLE 3.4: *One-sample two-tailed t-tests on the mean log ratios by word region for the future condition, Experiment 1*

Regions	t1 (df=1,31)	t2 (df=1,23)	P1	P2
Verb	3.28	3.76	.003	.001
Adv	1.69	1.03	.102	.310
NP2	-5.28	-6.14	.000	.000

Memory test: Figure 3.4 shows the percentage of correct answers, averaged by condition (by participants). Participants correctly answered 82.5% of the questions. As can be seen from the graph, participants were more accurate in recognizing the recent action events than the future action events (85.5% vs. 79.5%).

FIGURE 3.4: *Percentage of correct answers as a function of object and tense*

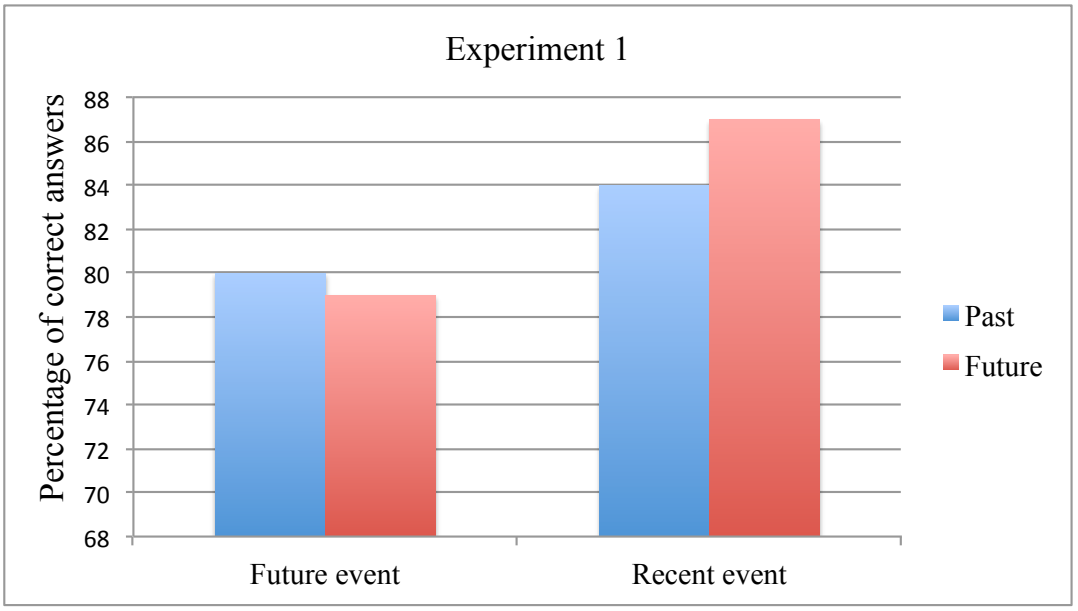


TABLE 3.5: *Linear mixed effect model results for the memory test, Experiment 1*

Effects	Coefficient	SE	z-Value	P
Intercept	1.98	0.19	11.01	.000
Object	0.21	0.16	1.28	.200
Tense	-0.12	0.17	-0.74	.460
Object x Tense	0.12	0.13	0.97	.330

Further LME analyses on the data from the memory test revealed that neither the main effects (of tense or event) nor their interaction was reliable. Even though

the data visually showed a difference, that difference was thus not confirmed by the inferential analyses. Perhaps then, the recent events are remembered better because they were fixated more than the future action events. Interestingly, this goes against the recency effect hypothesis, which should have resulted in better recognition of the future action events. In contrast, the short-term memory seems to be in line with the gaze data: the objects that are inspected more are also recalled better.

Summary

Even though with the current frequency distribution we showed evidence of an earlier gaze shift towards the future event target object in the future tense sentences than in experiment 2 with 50/50% by Knoeferle, Carminati, et al. (2011), surprisingly however, the recent-event preference has been replicated. We expected an early future event target inspection from the end of the future tense verb in the future tense condition. Contrary to our expectations participants shifted their gaze to the future target only from the middle of the Adverb region. That is around 700 ms after the verb offset. Moreover, the overall recent event target preference emerged in all three sentence regions (the verb, if only by items; the adverb, and the NP2 regions). The memory test findings with a better recognition of the recent action events supports the findings of the eye-tracking study.

It is not clear why the shift of visual attention to the future target object did not occur even earlier (given the strong 75%-25% frequency distribution). One reason could be that participants preferred to inspect the recent event target because they have always seen the recent action before the sentence. Thus, in spite of a strong tense cue, recently perceived action events seem to have a strong influence on anticipating objects in the scene. The next experiment investigates this question by further increasing the number of future events and future tense sentences and additionally by decreasing the number of the recent action events.

3.2 Experiment 2

The second experiment has further tested the issue of the recent-event preference. By adding more fillers for future events and future tense sentences, we increased the strength of the bias towards the future tense. We increased the frequency for the future tense sentences and future events up to 88%. Correspondingly the

percentage of the recent event and past tense sentences made up 12%. In addition to increasing the frequency, Experiment 2 differed from Experiment 1 further in that the filler trials did not show any recent action events. Each filler trial started with a static picture showing the experimenter in a neutral position (e.g., Fig. 1., B) with two objects on the table. Thus, the fact that most of the filler trials with a static picture (without a recent action event) might change participants' behaviour of looking at the recent event target throughout the experiment. Will not seeing a recent action in most of the trials decrease the overall inspection preference of the recent event target? Or will the recent-event preference remain robust instead of being eliminated by a strong frequency bias?

If not seeing the recent action event before the sentence in the majority of (filler) trials plays an important role in predicting an upcoming action event then we should see earlier and more looks to the future target object in Experiment 2 as compared to Experiment 1. Furthermore, we might see earlier and more looks towards the future event target due to the short-term memory experience. Increasing the number of future tense sentences and future events means more experience in the future condition that might direct listeners' attention more towards the future event target.

Additionally, as in Experiment 1 we conducted a memory test in Experiment 2. Although the first experiment failed to find any effect in the memory test, in Experiment 2 we might find a significantly better recall in the future (vs. recent) action event. This could be supported not only by increasing the number of future events and future tense sentences but also by the absence of recent action events in the fillers. If participants have overall less experience with recent actions in the visual context, then perhaps their attention and encoding of, these events will further decrease, negatively impacting later recall. But, if what we saw and heard about is bound in the working memory more strongly than what we hear and see later (i.e., we cannot confirm the truth of an event in the visual context if we have so far only heard about that event) then we might find better recall for the recent (vs. future) action event.

3.2.1 Method

Participants

Thirty-two German native speakers (Mean age=23.65; range: 18 to 30; 11 males, 1 of them left-handed, 21 females, 3 of them left-handed) participated in the experiment. All participants were students from the Bielefeld University community, and they were each paid 6 Euros for their participation. None of them had learned a second language before age 4. All had normal or corrected-to-normal vision, and were not aware of the purpose of the experiment.

Materials and Design

Eye-tracking: Methods and materials were the same as in Experiment 1. This experiment used all the critical and filler items from the Experiment 1. In addition, We created more filler items that increased (from 40 up to 72 items) the frequency of future events and future tense sentences. The frequency bias was achieved by having all filler items in the future condition. Thus, participants saw a total of 96 items (24 critical plus 72 fillers). All experimental items showed two actions, a recent action before the sentence and a future action after the sentence. Half of the sentences were presented in the past tense (12) and the other half in the future tense (12). All filler trials only showed future action events and presented future tense sentences. So, the number of future and recent actions seen by participants in this experiment was 84 and 12 respectively (88% future events and sentences vs. 12% past events and sentences). Thus, in the experiment here we instantiated an overwhelming bias towards future events.

One difference between experimental and filler trials was that each filler trials started with a picture depicting the experimenter and two objects on the table and after 700ms a future tense sentences was played out (see Fig. 3.1., b). After the sentence they saw a future event (see Fig. 3.1., c). Filler sentences never used any words from the critical sentences and the objects used in the fillers were not shown in the critical videos.

As in Experiment 1 we expected that the even stronger frequency bias towards the future condition might override the recent-event preference. When participants hear the first noun phrase 'The experimenter' we do not expect any differences in the gaze pattern towards the experimenter in the past and future tense condition. In addition to participants inspecting the experimenter during the NP1

region, we might see eye movements towards the recent action target. On hearing the Verb and the temporal Adverb 'flavors soon' we expect more attention being directed towards the future event target than the recent event target. Participants should begin to shift their gaze to the future event target earlier than in Experiment 1 at the end of the Verb region. If so, we might observe a 500-700 ms earlier inspection of the future event target than in Experiment 1. Additionally, eye-movements toward the future event target in the Adverb regions should be significantly more in the future tense than in the past tense. On hearing the Verb and the temporal Adverb 'flavored recently' in the past tense condition we expect no changes compared to Experiment 1. When the second noun phrase 'the tomatoes' in the future tense sentence is heard, the fixations on the tomatoes should continue to increase and similarly to Experiment 1 participants should inspect the tomatoes more often in the future tense than in the past tense sentence. When the second noun phrase 'the cucumbers' of the past tense sentence is encountered, participants should continue to inspect the recent target object and significantly more eye-movements should be made to recent target (the cucumbers) in the past tense than in the future tense. Crucially, the even stronger frequency bias towards future events and future tense sentences -if effective- should elicit an attention shift towards the future event target earlier than in Experiment 1 (in the Verb region) and override the overall recent-event preference. 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 (trials in each list N=72). Before the experiment, lists were pseudo-randomized and each participant saw an individually randomized version of one of the four experimental lists.

Memory test: Materials were the same as in Experiment 1. We expected a better recognition in the future tense condition for the future event target than for the recent event if the frequency distribution is effective. However, if we continue to see more visual attention to the recent event target and, if attention is related to later memory and recognition, then such a gaze pattern might coincide with better memory of the recent action event and its target.

Precedure

Procedure for the eye-tracking study and for the memory test were the same as in Experiment 1. Due to increasing the number of the filler items, the length of the eye-tracking study increased and we made sure to give a break to participants in the middle of each experimental list. Afterwards participants performed the memory

test and at the end they were debriefed. The experiment lasted approximately 55 minutes.

Analyses

Analyses for the *eye-tracking* study and for the *memory test* were identical to Experiment 1.

3.2.2 Results and Discussion

Eye-tracking: The results of the log gaze probability ratios from Experiment 2 are shown in Figure 3.5. Similar to Experiment 1, the gaze pattern indicates that we replicate the preferred inspection to the recent event target. Surprisingly, at the beginning of the Verb region, and for 800 ms onward, participants' attention towards the recent target did not differ between the future and past tense condition. This gaze pattern suggests that the tense information of the Verb did not modulate the gaze pattern towards the recent target neither in the future nor in the past tense sentence conditions. Only at the end of the Verb region the lines start to diverge, indicating a decrease in eye-movements to the recent event target in the future tense condition. However, the participants' bias on looking at the recent target object continues until the middle of the Adverb region irrespective of tense; the log ratio in both tense conditions remained above zero until 2100 ms after the Verb onset. This gaze pattern indicates that until then, the recent event 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, which means a preferential inspection of the future event target over the recent event target. In the second noun phrase region, inspection of the future event target continued in the future tense condition. In the past tense condition the preferential gaze pattern to the recent event target persisted until the end of the sentence.

Event though the present experiment increased the frequency of future events and future tense sentences, the instantiated frequency bias was not able to eliminate the recent-event preference. Moreover the tense effect did not reach significance in the Verb region. Mixed effects ANOVA (by participants and items) on the mean log ratios for the Adverb region showed a significant tense effect by participants $F1(df{=}1.31) = 14.30$, $p1 = .001$ and by items $F2(df{=}1.23) = 32.36$, $p2 = .000$.

FIGURE 3.5: *Mean log gaze probability ratios (ln (P(recent target/P(future target))) by condition from verb onset, Experiment 2*

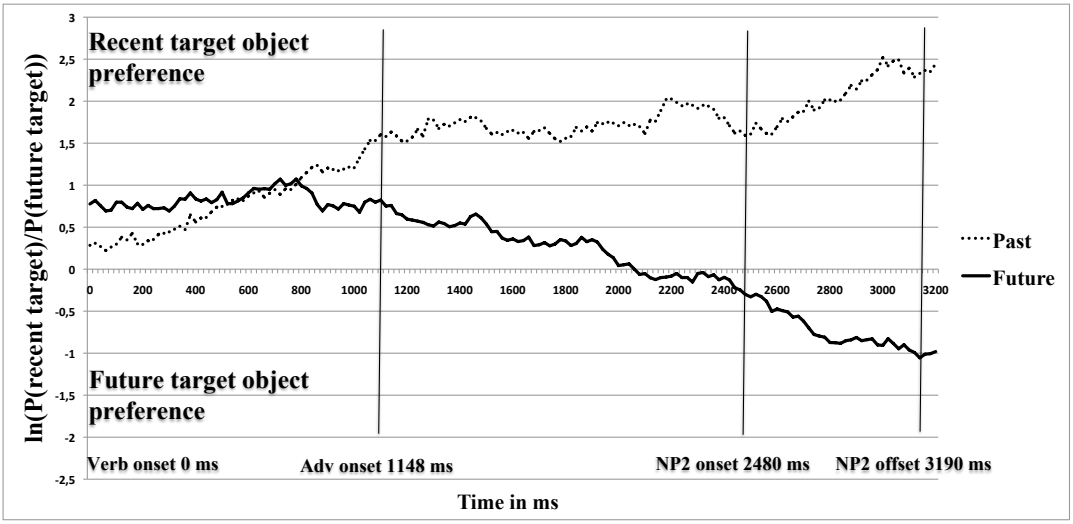


TABLE 3.6: *Grand mean and mean log gaze probability ratios (ln (P(recent target)/P (future target)) by participants as a function of condition and time region. Standard errors (SE) in parentheses, Experiment 2*

Regions	Future tense	Past tense	Grand mean
Verb	1.11 (.27)	1.11 (.23)	1.11 (.15)
Adv	.278 (.22)	2.11 (.39)	1.19 (.20)
NP2	-.502 (.33)	2.46 (.40)	.98 (.23)

The main effect of tense became fully significant in the NP2 region by participants $F1(df{=}1.31) = 36.03$, $p1 = .000$ and by items $F2(df{=}1.23) = 97.79$, $p2 = .000$.

Despite the more pronounced frequency bias in the present experiment than in Experiment 1, the present experiment replicated the overall recent-event preference. Participants overall looked significantly more at the recent event target than at the future target independent of the tense of the sentence $ps < .001$ (see the intercept, Table 4.6). However, we should note that the increased frequency in the present experiment seems to have reduced the recent-event preference approximately 150 ms earlier in the future tense condition (of Experiment 2) compared to Experiment 1.

However, surprisingly, in contrast to Experiment 1, the tense effect in the Verb region disappeared. It seems that the recent-event preference persisted in the Verb region for Experiment 2. One explanation could be that for the majority of the trials (including all filler trials), the recent action event was not shown. The

absence of the recent action event might have biased participants to focus their attention on the target of the recent action for the few trials in which recent action events did occur.

However, the recent-event preference decreased rapidly and the looks to the future and recent target objects as a function of tense began to diverge from the end of the Verb region. Importantly, the preferential looks to the future event target started approximately 800 ms after the Adverb onset, which is about 1000 ms earlier than in the experiment with a balanced (50/50%) frequency distribution of recent and future action events (Knoeferle et al., 2011). This comparatively earlier emergence of tense effects suggest that the frequency bias did modulate the recent event inspection preference.

TABLE 3.7: *ANOVA analyses for the data of Experiment 2 by region: The intercept is also given because in this case a significant intercept indicates that the grand mean is significantly different from 0, Experiment 2*

Regions	**Effect**	**F1 (df=1,31)**	**F2 (df=1,23)**	**P1**	**P2**
Verb	Intercept	52.22	106.66	.000	.000
	tense	.000	2.15	.990	.160
Adv	Intercept	35.28	42.41	.000	.000
	tense	14.30	32.36	.000	.000
NP2	Intercept	17.48	35.90	.001	.000
	tense	36.03	97.79	.000	.000

Similar to Experiment 1 the post-hoc *t*-tests in the present experiment revealed positive means in the Verb and Adverb regions (Table 3.8). The effect was reliable by participants and items in the Verb region and only by items in the Adverb region (indicating a preference for the recent target). The means are negative in the NP2 region and the effect is significant by items and only marginal by participants (indicating a preference now for the future target). Interestingly, increasing the number of future event and future tense sentences in the present experiment even further (relative to Experiment 1) did not reveal an earlier future event target inspection in the future tense compared with the results from Experiment 1. As in Experiment 1, the recent-event preference was confirmed in the Verb region and even in the Adverb region by participants; more inspections of the future event target in the future tense condition emerged only in the NP2 region.

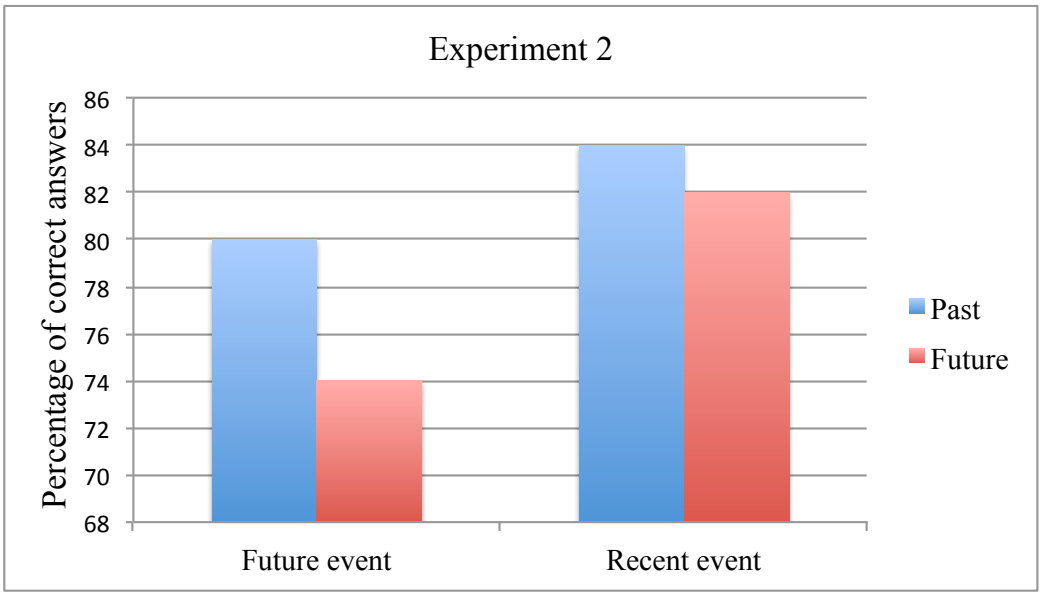
Memory test: Figure 3.6 shows the percentage of correct answers, averaged by condition (by participants). Participants correctly answered 80% of the questions.

TABLE 3.8: *One-sample two-tailed t-tests on the mean log ratios by word region for the future condition, Experiment 2*

Regions	t1 (df=1,31)	t2 (df=1,23)	P1	P2
Verb	4.17	7.40	.000	.000
Adv	1.25	2.98	.219	.007
NP2	-1.90	-5.61	.067	.000

As can be seen from the graph, participants were more accurate in recognizing the recent action events than the future event action events (83% vs. 77%). The target actions were recognised better in the past tense than in the future tense conditions, even with a frequency bias that was more pronounced in Experiment 2 than in Experiment 1.

FIGURE 3.6: *Percentage of correct answers as a function of object and tense*



The LME analyses of the memory test showed marginal effects of target event ($p < .06$) and tense ($p < .09$), but there was no interaction between tense and event. Interestingly, the memory test results have replicated the findings in Experiment 1. In Experiment 1, the enhanced recognition of recent action events over the future action events was not reliable. By contrast, in Experiment 2 the enhanced recognition of the recent actions approached significance. These results are in line with the view that more attention to an object during comprehension can benefit later memory. Alternatively, they support the view that actions that occurred first during a trial are better encoded in memory than future actions that occurred last during a trial.

TABLE 3.9: *Linear mixed effect model results for the memory test, Experiment 2*

Effects	Coefficient	SE	z-Value	P
Intercept	1.76	0.24	7.43	.000
Object	0.27	0.14	1.84	.060
Tense	-0.18	0.11	-1.70	.090
Object x Tense	-0.09	0.13	-0.67	.500

Summary

In summary, the results revealed more inspection of the recent event target than future event target during the Verb region irrespective of tense in Experiment 2, and modulated somewhat by tense in Experiment 1 (recall the tense effects during the Verb region in Experiment 1). However, as a result of our frequency distribution, at the end of the Verb, eye-movements towards the recent target started to decrease and from the middle of the Adverb region more attention was directed to the future event target in the future tense than in the past tense condition.

That said, our frequency distribution (75% of all trials showed future events and presented future tense sentences in Experiment 1; 88% of all trials showed future events and presented future tense sentences in Experiment 2) elicited earlier eye-movements (by approximately 1000 ms) towards the future event target than in the experiment by Knoeferle, Carminati, et al. (2011, experiment 2). The increased gaze to the future target in the future tense condition emerged descriptively early on; however, the *t*-tests did not show a reliable preference for participants' inspecting the future target in early regions (i.e., in the adverb region).

Our findings provide evidence of a frequency effect and suggest that the recent-event preference is sensitive to changes in the frequency distribution of recent relative to future events and past tense relative to future tense sentences. However, surprisingly, at the same time the analyses demonstrate the robustness of the overall recent-event preference in both experiments. This has been partly supported by the results from the memory test, suggesting that participants were more accurate in recognising the recent (vs. future) action and its target object. In

conclusion, the recent-event preference seems to be sensitive to the frequency distribution although the present manipulations were not able to override the overall recent-event preference preference.

Perhaps the recent-event preference results from participants' attention to situation-specific cues. They pay particular attention during sentence comprehension to cues (such as a recent action) that they have just experienced in the immediate context. If this is the case, and if the situational presence of a cue is key in guiding attention during spoken comprehension, then pitting the recent action against another situation-immediate cue should eliminate the preference. To assess this possibility, Experiments 3 and 4 pitted the gaze of the actor in the videos against the recent actions.

Chapter 4

Gaze Cue Effects during Language Understanding

This chapter presents the third and fourth experiments of the thesis. The experiments investigated the issue of the recent-event preference by manipulating a situation-specific cue (the gaze of the actor in the videos).

Recall that Experiments 1 and 2 have examined whether the recent-event preference can be eliminated by means of a strong frequency bias towards future events. Despite the strong frequency bias towards the future event, we replicated an overall recent-event preference (as reflected in a reliable intercept with a positive valence in all three word-regions of both sentence tense conditions).

In the present experiments we wanted to test the preference of the recent event by pitting a powerful situational cue such as an actor's gaze against the recent-event preference. Object based gaze is very effective in directing a listener's visual attention during language comprehension (e.g., Hanna & Brennan, 2007).

Previous research has applied a human gaze cue in different paradigms (see chapter 2); for instance, on a perceptual level in that the gaze cues reflexively human attention (e.g., Friesen & Kingstone, 1998; Mansfield et al., 2003) and on a cognitive level during sentence comprehension (e.g., Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012). Furthermore, not only the gaze of a human interlocutor but even the object-based gaze of a robot can rapidly affect language processing in a human listener (e.g., Staudte et al., 2014). Existing research of a situational gaze

cue manipulation has revealed that listeners follow an instructor's gaze in anticipating a target object before the object was named (e.g., Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012).

To recall the experiments reviewed in chapter 2, a simple target-matching task has investigated the gaze cue effect in a speaker/listener pair (Hanna & Brennan, 2007). The authors found that the listeners used the gaze cue of the instructor to identify the correct target object before the linguistic information disambiguated reference. In their setting, the speaker faced the listener frontally such that it was relatively easy to perceive which object the speaker fixated at a given point in time. Furthermore Knoeferle and Kreysa (2012) found effects of a speaker's gaze on listener's visual attention and language comprehension even when the speaker did not fully face the listener and when perceiving her gaze was arguably more difficult. In another experiment while participants were instructed by an experimenter, they performed their task better when the instructor's gaze was present than it was absent (Macdonald & Tatler, 2013).

Recall the question: What is the cause of the recent-event preference? A possible hypothesis is that comprehenders visually ground the verb of a sentence in the recent event; that is, when they hear the verb, they prefer to associate it with a recently-inspected event and its associated target (Abashidze et al., 2011; Knoeferle & Crocker, 2007). Perhaps the gaze behavior is elicited by an invariant mechanism, or it might only emerge when recent events are predictive of what comes next. In the present experiments situated cues such as gaze shift were manipulated. Can we find an earlier effect of the inspection of future event targets when the gaze cue is present? If this is the case, more attention to the future event target might lead to an early reduction of the recent-event preference and we might even override the overall preference. Prior research has shown that gaze can have an immediate effect; however, we can not predict whether it will be strong enough against another immediate cue such as a recent action event.

If humans are more sensitive to a later situated cue (in this case; the gaze cue) than we should see an earlier increase in looks to the future event target. This effect could occur immediately after the gaze cue onset. But if the preceding recent event cue exerts a stronger influence on the listeners' visual attention than a later gaze shift of the actor during the sentence comprehension, we should replicate the recent-event preference.

In sum, a speaker's gaze is a powerful cue in directing a listener's visual attention to objects during language comprehension. The present experiments used the video materials that have been used in Experiments 1 and 2. Additionally, we recorded gaze cue videos. Experiment 3 examined the effects of the actor's gaze shift to the future target object when the onset of the gaze shift occurred 380 ms after the verb onset. In Experiment 4 we strengthened the gaze cue by showing it earlier (at the onset of the verb). After the eye-tracking part of Experiment 3, we conducted a memory test on the visual experimental stimuli. Likewise, after the eye-tracking part of Experiment 4, we conducted a (gated) memory test on the experimental linguistic stimuli.

4.1 Experiment 3

Experiment 3 examined the effect of a powerful situated cue such as an actor's gaze shift, which we pitted against the recent-event preference during spoken sentence comprehension. Identically to Experiments 1 and 2, participants saw short videos showing an event before and after the sentence. As before, one half of the sentences were presented in the past tense, the other half in the present tense with a future meaning. Events presented before the sentences and after the sentences occurred equally often (see Fig. 4.1). In contrast to Experiments 1 and 2, the present experiment introduced a gaze cue as a second factor. In half of the trials participants saw a video of an actor (a second experimenter) gazing at a target object; in the other half of the trials, they saw a static picture of the actor looking ahead (see Fig. 4.1., B, (a) and (b)). In contrast to other studies, our experiment has pitted the effects of an actor's gaze shift on the listener's language comprehension and visual attention against the effect of a recently-seen event and participants' preference to inspect its target.

For instance, the gaze shift in the study by Knoeferle and Kreysa (2012) started on average 950 ms after the verb onset, whereas in our experiment the gaze cue started on average 380 ms after the verb onset. The goal was to examine at what point in time the actor's gaze cue would direct comprehenders' visual attention to a plausible future event target. Furthermore, we examined whether guidance of the listener's visual attention by the actor's gaze would be strong enough to override the overall recent-event preference. Alternatively, if there is an invariant

recent-event preference then we may find that the actor's gaze cannot modulate the preferential inspection of the recent event target.

Our expectations are: first, if gaze has an effect on directing visual attention, then we should see earlier and more looks to the appropriate object (recent and future) when gaze is present compared with when it is not present. Next, gaze might be particularly effective in enhancing looks to the future event target i.e., with gaze the recent-event preference might disappear and we thus would not necessarily expect that the actor's gaze can further enhance this preference. By contrast, for the future tense condition, the actor's gaze to the future event target might guide the listener's attention away from the recent target and towards the future target. In the study by Knoeferle and Kreysa (2012) gaze cue effects were more pronounced in the SVO than OVS sentence condition. In addition, participants fixated the target character 358 ms earlier in the gaze than in the no-gaze condition. Given the findings by Knoeferle and Kreysa (2012), the manipulation in the present study might help us elicit early gaze cue effects, thus diminishing the recent-event preference (to the extent that the latter preference can be modulated by another situation-immediate cue such as the actor's object based gaze)

When listeners see the actor's gaze shift to an object, this could allow participants to anticipate the plausible future action event a few hundred milliseconds after the gaze shift. If this is true, it may eliminate the overall recent event inspection preference during the verb. Thus, more looks towards the future event target in the second half of the experiment should occur if there is a learning effect.

Alternatively, participants may not look at the future target earlier because people may pay more attention to recent events than to possible future events. It is possible that a recent real world event has a stronger immediate effect on language comprehension than a situated cue, such as an actor's gaze. Especially, when the event has appeared prior to the gaze cue - there is a temporal precedence. If this is the case, then we should continue to see more looks to the target of the recent (vs. future) event even when gaze shift cues (the target of) a plausible future event.

Identically to Experiments 1 and 2, participants completed a memory test at the end of the eye-tracking session. The aim of the test was to see whether the gaze cue has a clear effect on recalling the future vs. recent action event. Other experiments have found that participants were better to complete a task when it was described with both speech and gestures, than when it was only described with language

(e.g., Beattie & Shovelton, 1999). Thus, we might find a better recall for the action events from the gaze than the no-gaze condition. However, findings by Knoeferle and Kreysa (2012) showed that the gaze cue had a short lived effect and was not beneficial for the later task (but note that Knoeferle and Kreysa reported some evidence for an effect of participants' gaze following on their response times in a post-sentence verification task).

If the gaze cue effect is more pronounced in the future tense condition, then we might find a better recall of the future event than the recent event in the future condition. However, if the recent even preference replicates, then we should see better memory accuracy for the recent than the future event (on the assumption that we replicate the results from Experiments 1 and 2 and if memory performance mirrors attention during comprehension).

4.1.1 Method

Participants

Thirty-two native German speakers participated in Experiment 3 (mean age=23.31; range: 18 to 32; 11 males, 2 of them left-handed, 21 females, 1 of them left-handed). Participants from the Bielefeld University community were each paid 6 Euros for their participation. None of them had learned a second language before age 4. All had normal or corrected-to-normal vision and participants were not aware of the purpose of the experiment.

Materials and Design

Eye-tracking: The present experiment used the same experimental materials as Experiments 1 and 2: 24 experimental items and 40 filler items. Each item consists of two videos with two everyday objects (e.g., tomatoes and cucumbers, see Fig. 4.1). One video was shown before the sentence, and a second video was shown after the sentence (see Table 3.1). For the purpose of this study we additionally recorded short videos (lasting about 3-4 seconds, depending on the length of sentences) for every item, showing an actor gazing at a target e.g., at tomatoes for the future action target or at cucumbers for the recent action target (see Fig. 4.1., B). Gaze has been added as a second factor in the present experiment. In half of the trials the actor gazed at the target object and in the other half he did not. These gaze cue videos have been shown in half of the experimental trials - 6 in the future

tense condition and 6 in the past tense condition (see Table 4.1). For the other half of the experimental trials participants saw snapshots i.e., a static photo (see Table 4.1) showing the actor in a static position looking at the camera. Examples for the videos and the snapshot associated with the sentences in Table 3.1., 1a and 1b are shown in Figure 4.1. The tense factor was identical to Experiments 1 and 2. In half of trials the sentences were in the past tense and in the other half of the trials, they were in the present tense with a future meaning (see Table 3.1 for the counterbalancing).

TABLE 4.1: *Counterbalancing the gaze (G) and no-gaze (NoG) conditions, Experiment 3*

Tense	*Exp items*	*Fillers*	*Total*
Future	12: 6 G / 6 NoG	18: 9 G / 9 NoG	30: 15 G / 15 NoG
Past	12: 6 G / 6 Nog	18: 9 G / 9 NoG	30: 15 G / 15 NoG
Total	24	36	60

We expect comprehension processes to manifest themselves in the eye-movement pattern when subjects process the tense of the sentence as well as the gaze shift towards the target. During the first noun phrase 'The experimenter' listeners will expectedly inspect the actor more than any other object and this should be similar in both future and past tense conditions.

When hearing the verb in the simple past tense and the adverb with the past meaning 'flavored recently' we expect participants to identify the target object of the recent event early and to look mostly at the recent event target viz., the cucumbers (e.g., Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011). When the gaze cue is present more eye-movements might occur toward the recent event target in the gaze than in the no-gaze condition (e.g., Knoeferle & Kreysa, 2012).

However, since there is a preference to inspect the target of the recent event, it is not clear whether the gaze cue will make much of a difference in directing attention to the recent event target. By contrast, on hearing the verb and the adverb in the future condition 'flavors next', participants should predominantly look at the future target object (the tomatoes) when the actor inspects the future target ahead of its mention compared with the condition in which he looks straight ahead. Thus, in the gaze condition participants should look more at the future than recent event target at the end of the verb region if they can rapidly exploit the

gaze cue. Similarly, gaze cue effects have been shown approximately 350 ms after the onset of a gaze cue shift (Knoeferle & Kreysa, 2012). If exploiting the gaze cue has a learning effect (such that anticipation of the future target is learned and transferred to other situations), we might even find an earlier gaze shift towards the future event target in the no-gaze condition in the present experiment than in experiment 2 by Knoeferle, Carminati, et al. (2011). Taken together, earlier and more looks to the future event target object should decrease preferential looks towards the recent event target. This should happen in the early sentence regions and this could even override the overall recent-event preference.

When participants hear the second noun phrase 'the cucumbers', their visual attention to the cucumbers should continue since the target object is confirmed by the NP2. Whether or not gaze effects will modulate participants' attention during the NP2 region, is unclear. A decreased gaze cue effect in the later sentence regions has been shown in Knoeferle and Kreysa (2012). When hearing the second noun phrase 'the tomatoes' referring to the future event target, participants should keep their attention on the tomatoes. If this pattern does not occur before, then they should inspect the future target (the tomatoes) during the NP2 more than the other object (the cucumbers).

Overall, eye-movements to the future event target might not occur as early and as often as the inspection of the recent target object. However, the situated gaze cue should be helpful in the future tense condition for predicting the upcoming future event. In sum, the presence of gaze should trigger earlier and more frequent looks to the appropriate target objects (recent and future). Crucially, if the gaze cue overrides the recent-event preference, it should have a stronger influence in sentences with future than past tense meaning (i.e., tense x gaze interaction).

The experimental and filler items were combined to form 8 lists using a Latin square design; each list contained every critical item in only one condition and all 40 fillers. Each participant saw an individually pseudo randomised version of one of the eight experimental lists.

Memory test: The materials for the memory test were identical to the materials in Experiments 1 and 2. The aim of the memory test was to examine participants' later memory of the recent and future action events. If gaze cues help to direct visual attention and, if they benefit encoding of information for later recall, participants will benefit from the prior visual attention in the memory test. The

action events from the gaze condition should be recalled better than the action events from the no-gaze condition. Moreover, if gaze has a stronger effect in the future than in the past condition, recall of the future event should be better than recall of the recent event. But, if there are more looks at the recent (vs. future) event target this might increase the probability of better recall of the recent event. However, if the gaze cue has no beneficial effect on short term memory, because it is short lived as found by Knoeferle and Kreysa (2012), then we should not see any better recall in the gaze condition.

Gaze detection test

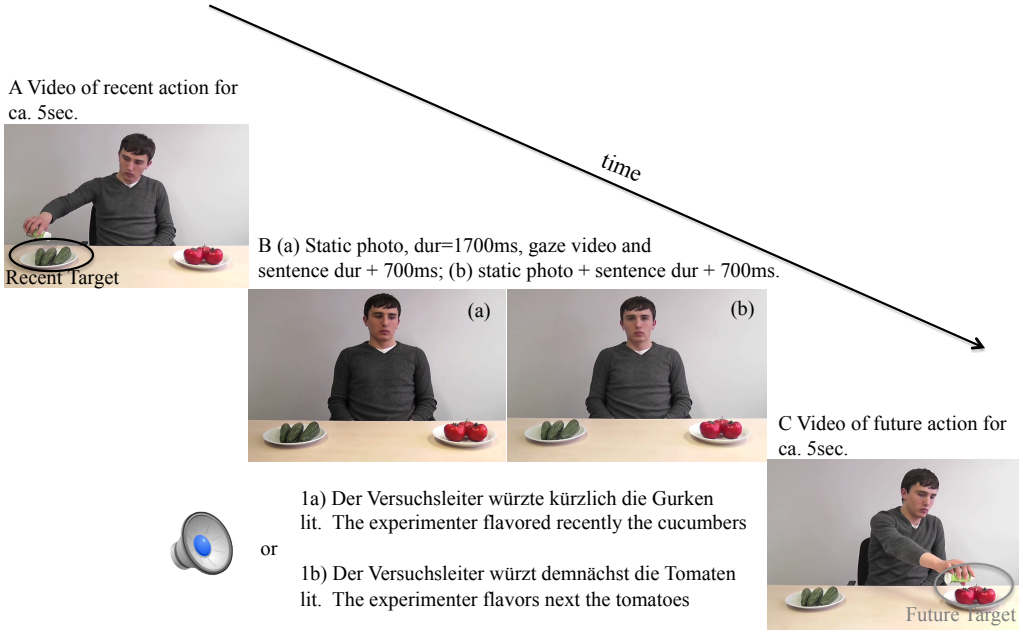
Previous studies have found that humans are very accurate to quickly notice the gaze direction of an interlocutor in a normal conversation (e.g., Pusch & Loomis, 2001). To find out how long it takes the onset of the actor's gaze shift to the target we asked participants (N=20) to take part in a pre-test. They were all students aged 20-30 and received 1 Euro for their participation. The participants were instructed as follows: "you will see a short video where an actor is sitting at the table and two objects are located on the table. In some videos the actor will gaze at one of the objects. Your task is to press the button as soon as possible when you notice the actor is going to gaze at an object". The results showed that participants took 380 ms on average to detect the gaze shift.

Procedure

Eye-tracking: The procedure was similar to the one used in Experiments 1 and 2. Except, in the present experiment participants inspected a gaze video during the sentence processing for half of the trials. First, participants saw a video showing an actor performing an action - a recent action (e.g., flavoring cucumbers, Fig. 4.1., A). Afterwards a past or future tense sentence was presented and in half of the trials a video showed an actor gazing at the target object. The gaze shift started 380 ms after the verb onset (e.g., a gaze shift to the future target - the tomatoes, see Fig. 4.1., B, (a)) and the actor kept his gaze on the tomatoes until the end of the sentence. 700 ms after the end of the sentence participants saw a second video in which the actor performed a second action - the future action (e.g., flavoring tomatoes, see Fig. 4.1., C). At the end of the eye-tracking section participants were asked to perform a memory test.

Memory test: the procedure of the memory test was identical to Experiments 1 and 2. After the memory test participants were debriefed. The experiment lasted approximately 40-45 minutes.

FIGURE 4.1: *Sequence of events of an experimental trial, Experiment 3 and 4*



Analyses

Eye-tracking: The data was analyzed in a similar way as in Experiments 1 and 2. Figure 4.2 shows participants' fixations to the two objects of interest: we present the time course of visual attention from the onset of the Verb until the end of the second noun (NP2) in the two conditions. The present experiment manipulated 2 factors: sentence tense (past vs. future) and gaze to target object (gaze vs. no-gaze). The darker lines indicate the recent condition (sentence was in the past tense) and the light grey lines indicate the future condition (sentence was in the present tense with future meaning). The dotted lines present the data from the gaze condition (the target object was gazed-at by the actor during the verb) and the solid lines present the data from the no-gaze condition (the target object was not inspected by the actor).

For the inferential analyses the time regions were the same as in Experiments 1 and 2. We aggregated the mean log gaze probability ratios ln (P (recent target)/P (future target)) over each of the three time regions of interest. The mean log gaze probability ratios for each region as well as the tense and the gaze/no-gaze

condition and the grand mean are provided in Table 4.2. Next we performed ANOVAs on the log ratios for each of the time regions with the independent factors tense and the factor gaze (gaze vs. no-gaze). Additionally, we performed one-sample two-tailed t-tests on the log ratios only for the future tense condition. The ANOVAs and the post-hoc one-sample t-tests results are summarized in Tables 4.3 and 4.4 respectively.

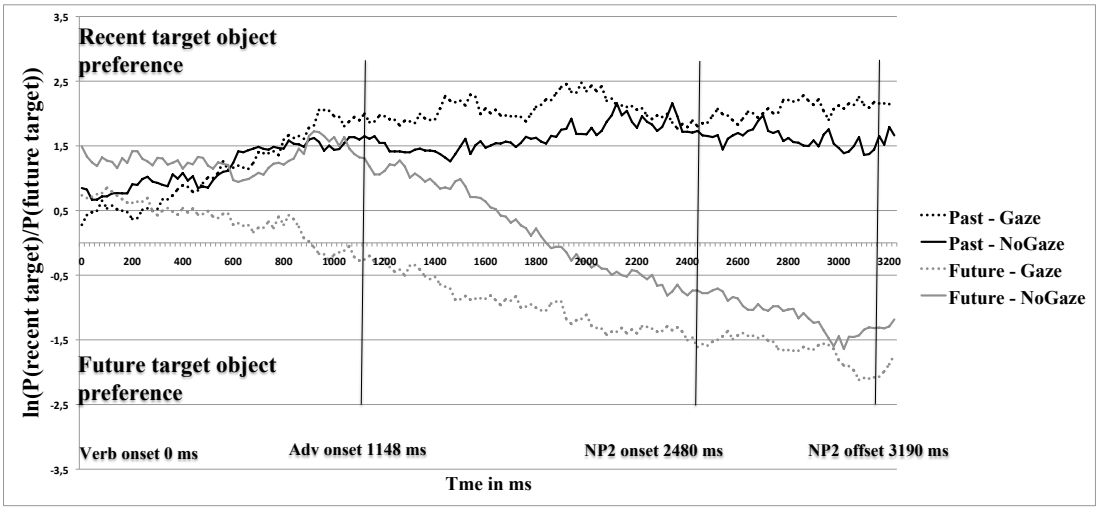
Memory test: We calculated the percentage of correct answers by condition for participants and items separately. The average percentages (by participants) are illustrated in Figure 4.3. The accuracy data was analyzed using LME analyses (see Experiments 1 and 2). In contrast with Experiments 1 and 2, the current analyses included another predictor, gaze (gaze vs no-gaze) in addition to the tense factor. The predictors were centered by transforming the fixed effect coding into a numerical value and centering it so as to have a mean of 0 and a range of 1. The results are presented in Table 4.4.

4.1.2 Results and Discussion

Eye tracking: Descriptively, the graph in Figure 4.2 shows that the gaze cue (dotted lines) has an effect in both future and past tense conditions. While in the future tense the effect of gaze was stronger and appeared earlier during the sentence, the past tense shows a slow and weak effect. The Verb region shows an overall preference for looking at the recent target relative to the future target from the Verb onset for about 900 ms in both gaze and no-gaze condition. However, when the actor gazed at the target, an earlier effect can be observed in the future tense condition (the value became negative 900 ms after the Verb onset - the light gray dotted line, see Fig. 4.2). Thus, in the future tense condition, gaze cue effects emerged from about 450 ms after the onset of the gaze shift. Similarly, gaze effects emerged around 350 ms after the onset of the speaker's gaze shift in the experiment by Knoeferle and Kreysa (2012).

By contrast, in the no-gaze future condition, the log-gaze probability values became negative (which means that more inspection went to the future event target) about 1000 ms later (see the light grey solid line, which was above zero until 1900 ms after the onset of the verb). However, the means of the log ratio for both the past and future conditions remain above zero for 1900 ms (indicating that the recent target receives more looks than the future target).

FIGURE 4.2: *Mean log gaze probability ratios (ln (P(recent target/P(future target))) by condition from verb onset, Experiment 3*



These no-gaze condition findings replicated the results from Experiments 1 and 2. The recent event target was preferentially inspected until the middle of the Adverb region in the future condition (the light gray solid line, Fig. 4.2). As expected, the gaze cue effect in the past tense condition did not emerge in the Verb region (see darker lines in Fig. 4.2). In the future tense condition the gaze cue effect started from the Verb region and continuously increased until the end of the last region (NP2). However in the past tense condition a weak gaze effect was found in the Adverb region but it lasted only until 300 ms before the Adverb offset.

During the last word region (NP2), participants more looked at the recent event target in the past tense sentences in both the gaze and the no-gaze condition. The future target object in the future tense received more inspections in the gaze than in the no-gaze condition.

Importantly, as we mentioned in the beginning, we can see that there is an overall stronger gaze effect in the future tense condition (see light grey dotted line, Fig. 4.2) than in the past tense condition (see darker dotted line, Fig. 4.2). Immediately after the gaze onset the lines started to diverge which means that in the gaze condition participants looked less at the recent event than in the no-gaze condition. This gaze pattern continues until the end of sentence.

The means in Table 4.2 show that there is a general inspection bias in favor of the recent target over the future target object. This pattern can also be seen in the time course graph (see Fig. 4.2). However, we should note that in the Verb region the value for the future gaze condition is lower (.04) than the values (from 1.94

TABLE 4.2: *Grand mean and mean log gaze probability ratios (ln (P(recent target)/P (future target)) by participants as a function of condition and time region for the experiment. Standard errors (SE) in parentheses, Experiment 3*

Regions	*Future NoG*	*Future G*	*Past NoG*	*Past G*	**Grand mean**
Verb	1.94 (.48)	.04 (.46)	1.95 (.36)	2.78 (.38)	1.68 (.24)
Adv	.42 (.33)	-1.75 (.44)	2.11 (.34)	2.87 (.30)	.913 (.23)
NP2	-1.41 (.38)	-2.35 (.34)	2.36 (.41)	3.14 (.38)	.43 (.19)

to 2.78) in the other conditions. In the Adverb region this can clearly be seen by the negative value (-1.75) for the gaze condition in the future tense, while in the no-gaze condition the value (.42) remains positive. By contrast, in the past tense condition the values of the gaze (2.87) and no-gaze (2.11) conditions did not differ much for the Adverb region. The grand mean in all the three regions is positive (however, much less positive in the NP2 than in the Verb and Adverb regions), indicating that people look more at the recent than future target independent of tense and gaze. In the past tense and gaze condition, the 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 tense and gaze conditions 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 4.2). However, the increase of the absolute log ratio in the future tense condition is higher than in the past tense condition, especially in the NP2 region.

After having discussed the descriptive statistics, the following text will describe the inferential statistics. ANOVA results reveal a main effect of tense in all three regions $ps < .008$, however, the gaze effects were reliable only in the Verb and Adverb regions. Importantly, the conclusions emerging from the descriptive analyses above were mostly corroborated by the inferential analyses of the data. The log ratio showed a positive value in all three regions (grand mean, Table 4.2), and the inferential analyses supported this with a significant intercept in all three regions by participants and items $ps < .033$ (see Table 4.3).

For the Verb region, the ANOVA showed a main effect of tense by participants $F1(df{=}1.31) = 8.09$, $p1 = .008$ and by items $F2(df{=}1.23) = 15.52$, $p2 = .001$. However, the gaze effect in the Verb region was only significant in the by items analysis $F2(df{=}1.23) = 5.34$, $p2 = .030$, and only marginal in the by participants analysis $F1(df{=}1.31) = 3.21$, $p1 = .083$. The significant effect of tense in the Verb

TABLE 4.3: *ANOVA analyses for the eye-tracking data by regions: The intercept is also given since a significant intercept indicates that the grand mean is significantly different from 0, Experiment 3*

Regions	Effect	F1 (df=1,31)	F2 (df=1,23)	P1	P2
Verb	Intercept	51.22	103.18	.000	.000
	Tense	8.09	15.52	.008	.001
	Gaze	3.21	5.34	.083	.030
	Interaction	10.31	6.32	.003	.019
Adv	Intercept	19.31	62.30	.000	.000
	Tense	57.82	98.67	.000	.000
	Gaze	4.92	3.55	.034	.072
	Interaction	15.86	7.56	.000	.011
NP2	Intercept	4.97	10.05	.033	.004
	Tense	100.24	163.18	.000	.000
	Gaze	.06	.010	.808	.921
	Interaction	6.37	10.85	.017	.003

region has replicated the findings from Experiment 1. Furthermore, we can see that there is a significant tense effect in the Adverb and NP2 regions.

The Adverb region showed the tense effect in both by participants $F1(df{=}1.31) = 57.82$, $p1 = .000$ and by items analyses $F2(df{=}1.23) = 98.67$, $p2 = .000$. Similarly, analyses of the data for the NP2 region, revealed a tense effect in the by participants analysis $F1(df{=}1.31) = 100.24$, $p1 = .000$ and in the by items analysis $F2(df{=}1.23) = 163.18$, $p2 = .000$. In line with the results in the time course graph, the main effect of actor's gaze remains fully significant in the Adverb region in the by participants analyses $F1(df{=}1.31) = 4.92$, $p1 = .034$ and is only marginal in the by items analyses $F2(df{=}1.23) = 3.55$, $p2 = .072$. (see Table 3.3).

By contrast, no effect of gaze occurred in the NP2 region. In addition, a reliable interaction emerged in all three regions indicating a more pronounced gaze cue effect in the future tense condition than in the past tense condition. Overall results show the tense effect in all three regions and the gaze effect only in the Verb and Adverb regions. This finding suggests that the gaze cue is not as effective in the last (NP2) region as in the prior region (i.e., at the adverb). Similarly, short-lived gaze effects have been reported in previous research (e.g., Knoeferle & Kreysa, 2012).

Furthermore, (recall, the aim of the t-tests was to find out whether the log-ratio means in the future condition were significantly different from zero) the t-test

TABLE 4.4: *One-sample two-tailed t-tests on the log ratio means for the future condition, testing whether these are significantly different from 0 by regions, Experiment 3*

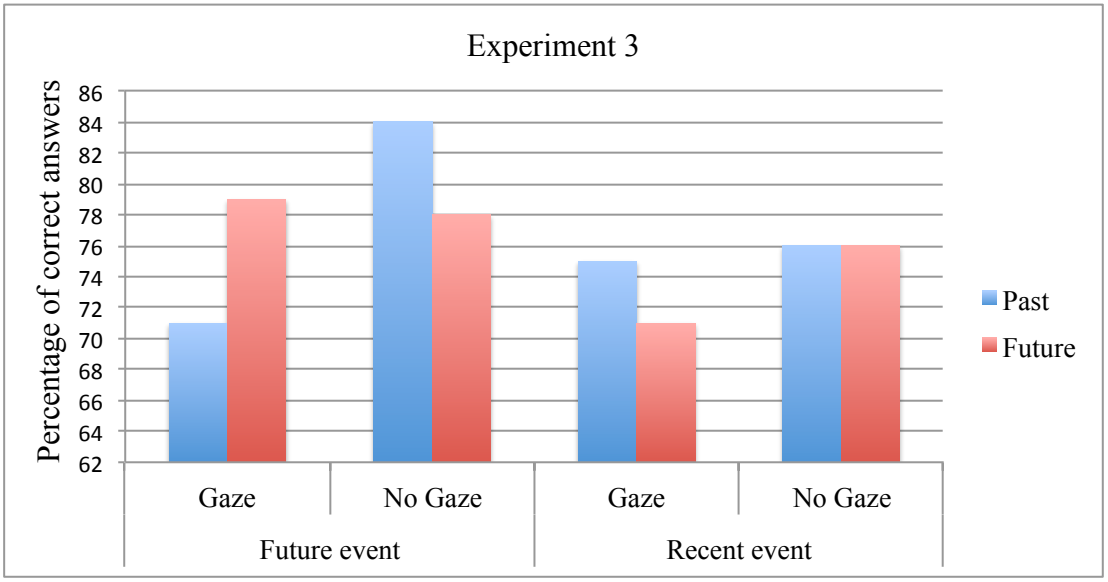
Regions	**Gaze**	**t1 (df=1,31)**	**t2 (df=1,23)**	**P1**	**P2**
Verb	Gaze	-0.32	-0.18	.749	.857
	No Gaze	4.99	4.85	.000	.000
Adv	Gaze	-3.92	-3.80	.000	.001
	No Gaze	1.25	2.34	.200	.028
NP2	Gaze	-6.64	-6.66	.000	.000
	No Gaze	-3.78	-4.52	.001	.001

results in the no-gaze condition replicated the findings of Experiments 1 and 2. The data in the Verb region revealed positive means and reliable effects (indicating the recent-event preference) and similarly the effect was reliable in the Adverb region (by items). The NP2 region showed negative means and a reversed significant effect (indicating more attention to the future than the recent target object) By contrast, the data in the gaze condition in the Verb region did not reveal any effect (no preferential inspection of the recent or future event target). However, the data in the Adverb and NP2 regions showed negative values and reliable effects (meaning that participants tended to look more at the future than at the recent target object). In comparison with Experiments 1 and 2, the present experimental manipulations revealed a preference to inspect the future action target in the Adverb region (only in the gaze condition). This gaze pattern suggests that the situation immediate gaze cue has mitigated but could not completely override the recent-event preference.

Memory Test: Participants gave correct answers to 76% of the questions. Figure 4.3 shows that participants were more accurate in recognizing the future action than the recent action events (78% vs. 74%). The graph further includes results from the gaze and no-gaze condition. Participants were more accurate in recognizing the action events in the no-gaze condition (78.5%) than in the gaze condition (73.5%). Further, they were more accurate in recognizing the future events in the future tense gaze condition than in the no-gaze condition. However, they were better in recalling the recent events in the past tense no-gaze condition and in the past tense gaze condition compared with the future tense conditions. Furthermore, the future action events were recalled better in the past no-gaze condition than in

the gaze condition. Similarly, the recent action events were recalled better in the future no-gaze condition than in the gaze condition.

FIGURE 4.3: *Percentage of correct answers as a function of object and tense*



Results from the LME analysis showed a significant effect of the action event and a marginal effect of gaze (see Table 4.5, p= .060) but no effect of tense. Overall, people were more accurate in identifying the action they had seen after hearing the sentence than before the sentence (which is the equivalent of saying that they were more accurate in recognizing the future than the past action and its target object). There was no reliable interaction between event and tense. Therefore, the future event was recognized better than the past event independent of the tense of the sentence. Moreover, there was no interaction between gaze and tense. However, there was a marginal interaction between event and gaze. This means that the target object was recognized better when the actor shifted gaze to it, than when the actor looked straight ahead - independent of the tense of the sentence.

The results of the memory test, as far as the effect of gaze is concerned, support the findings of the eye-tracking study. The eye-tracking results showed more inspections of the future event target in the gaze than in the no-gaze condition. In the memory test data from the gaze condition, participants correctly identified the future action event in the future tense condition more often than the recent action event in the past tense condition. By contrast, in the no-gaze condition participants performed better in recalling the future action event than the recent action event in both tense conditions. This did not match the eye-tracking findings of the preferential inspection of the recent event target.

TABLE 4.5: *Linear mixed effect model results for the memory test, Experiment 3*

Effects	Coefficient	SE	z-Value	P
Intercept	1.85	0.28	6.37	.000
Object	-.31	0.12	-2.53	.000
Tense	0.03	0.15	0.23	.810
Gaze	-0.42	0.23	-1.81	.060
Object x Tense	-0.11	0.14	-0.82	.410
Object x Gaze	0.26	0.15	1.73	.080
Gaze x Tense	-0.13	0.13	-0.91	.320

In summary:

The data showed an effect of gaze manipulation. Despite the continued strong effect of recent events, the actor's gaze shift was able to direct participants' attention towards the future event target approximately 450 ms after the gaze onset. However, the gaze effect occurred later compared with the gaze cue effect in other experiments, for instance, those by Knoeferle and Kreysa (2012); Staudte et al. (2014). From the end of the Verb region participants followed the actor's gaze to anticipate the target object of the future event in the future tense sentence, which was confirmed by the significant *t*-tests results in the Adverb region. In contrast to experiment 2 by Abashidze et al. (2011) with an equal number of future and recent events as well as future and past tense sentences, the present experiment found an earlier tense effect as a result of the gaze cue manipulation (i.e., in the Verb region). As expected, the effect of gaze was stronger in the future tense than in the past tense condition. The results of the eye-tracking study were only partly supported by the memory test findings, which were that the future action events and their targets in the gaze condition were recalled better than the recent action events and their target objects.

Overall, the results showed an effect of the gaze cue as revealed by more inspection of the future event target in the future gaze condition than inspection of the recent even target in the past gaze condition. However, the results of Experiment 3 replicated the overall recent-event preference from Experiments 1 and 2. One possible reason for not overriding the recent-event preference might be that the actor's gaze shift started late during the Verb region. If the gaze cue were to start at verb onset, perhaps we might see an enhanced likelihood of participants' inspection the future event target more and earlier in the future (vs. past) tense

condition. Would the overall recent-event preference disappear with such an earlier onset of the gaze manipulation? This question was addressed in Experiment 4, and the only change in the current experiment manipulation in comparison to Experiment 3 was that the actor's gaze shift started early - at the verb onset.

4.2 Experiment 4

In the present experiment the gaze shift onset and the verb onset occurred at the same time. Recall that in Experiment 3 the gaze cue was effective in directing listeners' visual attention during spoken sentence comprehension but this effect emerged only approximately 450 ms after the gaze onset. However, it is not clear whether this effect was elicited by the gaze cue alone or by the gaze cue plus the verb information (which would have been fully encountered approximately 150 ms after the onset of the gaze shift).

The manipulation in the present experiment was motivated by the findings from Experiment 3 and other findings (e.g., Driver et al., 1999; Friesen et al., 2005; Habets, Kita, Shao, Özyurek, & Hagoort, 2011). Recall the findings from the second chapter that the reaction time in the gaze cue condition was faster when the gaze cue to a target occurred in 300 ms than in 100 ms stimulus onset asynchrony (e.g., Driver et al., 1999; Friesen et al., 2005). In experiments by Habets et al. (2011), speech and gesture occurred at different stimulus onset asynchronies (from 0 to 360 ms). They found that participants used the gesture cue more when it appeared at the same time as the speech compared to when these two cues did not appear at the same time. Moreover, in a visual world paradigm study an earlier and longer preview of the visual context resulted in earlier attention toward a target object when the language unfolded (e.g., Dahan & Tanenhaus, 2005). Together these results suggest that the time participants have to inspect a visual context and the timing with which different linguistic and non-linguistic cues are present matters for their integration, and thus arguably also for their effects on participants' language comprehension and visual attention.

The present experiment further tested the situational gaze cue against the recent-event preference. However, it used a higher degree of synchrony of the gaze cue and the verb onset. In comparison to Experiment 3, the gaze cue in the present experiment started at the verb onset; in other words, some hundred milliseconds

earlier than in Experiment 3. This earlier onset was chosen in order to make earlier effects possible and eliminate a potential conflict between tense and gaze cue information. We asked whether the earlier gaze cue onset would have an earlier and stronger effect on participants' visual attention to the recent and future target objects than in Experiment 3 - especially in the future tense condition. If the gaze cue effect in the future tense conditions started early and lasted until the end of the sentence, the overall recent-event preference might be eliminated. Alternatively, the recent-event preference might be stronger than the gaze shift, which occurred at verb onset. In this case, we should replicate the overall recent-event preference from our previous findings.

In addition to the eye-tracking experiment, we conducted a gated memory test. In this test, participants were presented with experimental sentences in three stages and they were asked to recall them. At first, they were asked to recall the tense of the verb, then to recall the last noun phrase without seeing the visual information, and in a third step, they were asked to recall the NP2 but with additional visual information as a cue (picture of potential NP2 referents). If the current gaze cue manipulation shows earlier and more looks towards the future target object and if this translates into better memory performance, then we should also see a better recall of the future tense than the past tense sentence. However, if the tense information was stronger in the past tense condition than in the future tense condition -perhaps because the past tense referred to the recent event that comprehenders can immediately verify- then we might find a better recall of the past tense than future tense sentences. This combination might be transferred better into short-term memory than a future tense referring to a plausible future event that comprehenders cannot immediately verify as they hear the verb and adverb.

Previous findings showed that participants performed better in completing sentences in the past about past thoughts than in the future about imagining future thoughts (Anderson & Dewhurst, 2009). Alternatively, participants might recall the future tense sentences better because during the future tense sentence they have not yet seen the future event and they might be exited about the plausible future event and this could lead to a better recollection and recall of the future tense sentences. Moreover, if the gaze cue has an effect on recall of the sentence information, we should see better recollection of the sentence in the gaze than in the no-gaze condition.

4.2.1 Method

Participants

Thirty-two further native German speakers participated in Experiment 4 (Mean age=24.4; range: 19 to 31; 11 males, 1 of them left-handed, 21 females, 1 of them left-handed). Participants from the Bielefeld University community were each paid 7 Euros for their participation. None of them had learned a second language before age 4. All had normal or corrected-to-normal vision, and were unaware of the purpose of the experiment.

Materials, design and procedure

Eye-tracking: Materials and design were identical to those in Experiment 3. The only difference in the procedure was that the gaze cue appeared 380 ms earlier, that is at the same time as the verb onset.

It was expected that when the participants hear 'The experimenter' they will mostly look at the actor. When hearing the verb in the past tense and adverb 'flavored recently', we expect they will identify the target object of the recent event (the cucumbers). Crucially, their eye-movements toward the cucumbers might occur earlier in the gaze condition than in the no-gaze condition (in contrast to Experiment 3 and the findings by Knoeferle & Kreysa, 2012). But when hearing the present tense verb and the temporal adverb with the future meaning 'flavors next', participants should predominantly look at the future target (the tomatoes). Especially early in the verb region, participants' eye-movements should be directed more towards the tomatoes (the future target) in the gaze than in the no-gaze condition. This effect should occur earlier than in Experiment 3. Taking together earlier and more looks to the future target object, the results should decrease the overall recent-event preference or even override it. On hearing the second noun phrase 'the cucumbers', the fixation on the cucumbers should continue since the target object is confirmed by mentioning it (NP2). Here, the gaze pattern might not differ in the gaze and no-gaze condition. When participants hear the second noun phrase 'the tomatoes', their gaze should mostly be directed at the tomatoes. Eye-movements to the future event target might not be made as early as eye-movements to the recent target object. For this reason, the early onset of the situational gaze cue should be particularly helpful in the future condition. For

the future condition, an earlier onset of the gaze cue should help anticipating the future event target.

Gated memory test: In addition to the eye-tracking test participants performed a gated memory test that examined the accuracy in recognizing the past and future tense sentences in three stages. All sentences were taken from the experimental trials and participants were assigned to four counterbalancing lists of eye-tracking sessions and each were shown a randomized order of the list. The first stage showed participants the first noun phrase and the root of a verb and they were asked to recall the tense of the verb. After that they were shown the correct answer and at the second stage, they saw the verb plus the temporal adverb and were asked to recall the second noun phrase (target name). If participants could not answer the question at the second stage correctly, they were additionally shown the target object from the trial together with a distractor object and a third object from an unrelated experimental trial. Participants were to give their response verbally and the experimenter wrote down their response.

We expected participants to perform better in recalling the past than future tense verb at the first stage (see also Anderson & Dewhurst, 2009). When participants were listening to a past tense verb they had previously seen an event involving the target object and they might retain it in their working-memory. By contrast, in the future tense condition the future event had not been seen at the time when participants heard the verb. Thus, the sentence content in the future condition may be less closely tied to the event than the sentence content in the past condition. However, in the gaze condition we expect a better recollection of the future tense sentence, if the gaze cue has early and strong effects; perhaps it helps in integrating and storing the future tense cues and the possible future event in memory. Does the recall of the sentence tense depend on how often participants inspected the correct target in the future and past tense conditions? If so, then we should see a better recall of sentence content in those conditions for which the target received most inspections. After the gated memory test ended, participants were debriefed. The experiment lasted approximately 50-55 minutes.

Analyses

Eye-tracking: The analyses for the study were the same as in Experiment 3. After looking at the descriptive statistics we conducted the inferential analyses.

Gated memory test: The test was analyzed in a similar manner as the memory test in Experiment 3. The percentage of correct answers by conditions for participants and items were calculated separately. The average percentages of three stages (by participants) are illustrated in Figure 4.5. The gated memory test results were then analyzed statistically. For the LME model the predicted outcome was the response at three different stages and the predictors were tense (past vs. future) and gaze (gaze vs. no-gaze). The stages were analyzed separately.

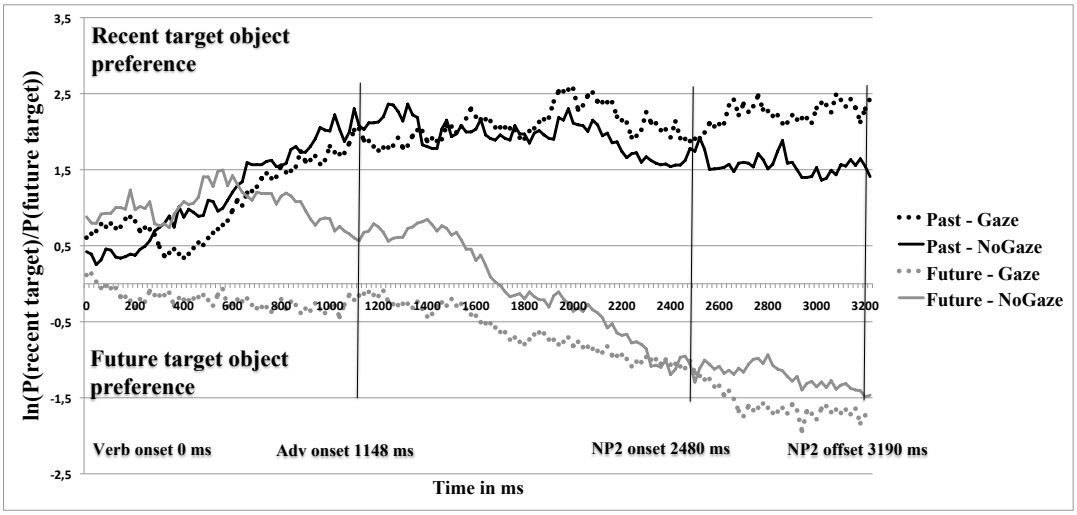
4.2.2 Results and Discussion

Eye tracking: The results of the time course graph are shown in Figure 4.4. The dotted lines represent the gaze condition in both past and future tense condition. In the future tense condition the gaze cue (see light grey dotted line) shows an immediate effect beginning about 100 ms after its onset. In comparison to Experiment 3, the effect of the gaze cue emerges earlier. However, the gaze effect is not strong enough to increase inspection of the future target (evidenced by a negative log-gaze probability ratio) until almost 1500 ms after its onset (see light grey dotted line, Fig. 4.4). From the middle of the Adverb region the gaze effect increased continuously until the end of sentence. The gaze cue is less effective in the past tense condition (see darker lines). The lines are almost at the same level until the middle of the Adverb region. Only in the first 300 ms and then approximately 1900 ms after its onset were participants more likely to look at the recent target object in the gaze than in the no-gaze condition.

In the no-gaze conditions the recent-event preference replicated. Participants preferentially inspected the recent event target until the middle of the Adverb region in the future no-gaze condition (the value becomes negative only from 1700 ms after verb onset). A shift in attention towards the future event target in the future no-gaze condition of the present study emerged approximately 200 ms earlier than in Experiment 3.

In comparison to Experiment 3, there was no recent-event preference in the future gaze condition during the Verb region. Moreover, the overall gaze pattern during the Adverb region in the no-gaze condition showed that the value was much lower in the present experiment .07 (see Table 4.5) than in Experiment 3 (.42 see Table 4.2). But the value of the grand mean in the Adverb region has increased in the

FIGURE 4.4: *Mean log gaze probability ratios (ln(P(recent target/P(future target))) by condition from verb onset, Experiment 4*



present experiment compared to the value in the Adverb region in Experiment 3 reflecting the limited effects of the gaze cue.

TABLE 4.6: *Grand mean and mean log gaze probability ratios (ln (P(recent target)/P (future target)) by participants as a function of condition and time region for the experiment. Standard errors (SE) in parentheses, Experiment 4*

Regions	*Future NoG*	*Future G*	*Past NoG*	*Past G*	**Grand mean**
Verb	1.25 (.33)	-.60 (.34)	2.53 (.35)	1.40 (.36)	1.15 (.20)
Adv	.07 (.36)	-.73 (.24)	3.46 (.40)	3.32 (.35)	1.53 (.18)
NP2	-2.20 (.42)	-2.37 (.36)	2.56 (.40)	3.47 (.39)	.37 (.19)

The following paragraph will give an overview of the inferential statistics. Similarly to Experiment 3, the ANOVA results revealed an effect of tense that was reliable in the analyses by participants and by items in all three regions $ps < .000$. The results further showed the gaze cue effect in the Verb region. However, in comparison to Experiment 3 the gaze effect has been eliminated in the Adverb region and similarly to Experiment 3 no effect was found in the NP2 region. Moreover, the conclusions from the descriptive analyses are mostly supported by the inferential analyses of the data. The log ratio indicated a positive value in all three regions (grand mean, Table 4.6) and the inferential analyses revealed a significant intercept in all three regions by items and by participants except the NP2 region $ps < .068$.

For the Verb region, the ANOVA revealed a main effect of tense in the by participants $F1(df{=}1.31) = 46.12$, $p1 = .000$ and by items analyses $F2(df{=}1.23) =$

42.11, $p2 = .000$. Furthermore, the effect of the gaze cue was reliable in the by participants $F1(df=1.31) = 21.42$, $p1 = .000$ and by items analyses $F2(df=1.23) = 5.67$, $p2 = .041$. The significant effect of tense in the Verb region was replicated from Experiment 3. Moreover, we found a significant tense effect in the Adverb and NP2 regions. However, the present experiment did not reveal a reliable effect in the Adverb and NP2 regions.

The Adverb region revealed the tense effect in the by participants $F1(df=1.31) = 77.33$, $p1 = .000$ and by items analysis $F2(df=1.23) = 93.82$, $p2 = .000$. Similarly, the results for the NP2 region showed the tense effect in the by participants $F1(df=1.31) = 132.14$, $p1 = .000$ and by items analysis $F2(df=1.23) = 155.63$, $p2 = .000$. Interestingly, the gaze cue effect was not shown in the Adverb region even though the time course graph suggested a difference (see Fig. 4.4). Moreover, in comparison to Experiment 3, the results in the present experiment did not reveal a reliable interaction of gaze and tense by participants. By items the interaction was significant in the Verb region and marginal in the Adverb and NP2 regions. The marginal interaction indicates a more pronounced gaze cue effect in the future tense condition than in the past tense condition.

TABLE 4.7: *ANOVA analyses for the eye-tracking data by regions: The intercept is also given since a significant intercept indicates that the grand mean is significantly different from 0, Experiment 4*

Regions	**Effect**	**F1 (df=1,31)**	**F2 (df=1,23)**	**P1**	**P2**
Verb	Intercept	36.06	42.30	.000	.000
	Tense	46.12	42.11	.000	.000
	Gaze	21.42	5.67	.000	.041
	Interaction	.754	12.70	.392	.002
Adv	Intercept	74.00	59.05	.000	.000
	Tense	76.33	93.82	.000	.000
	Gaze	2.59	.47	.118	.500
	Interaction	1.53	3.61	.226	.070
NP2	Intercept	3.57	4.43	.068	.047
	Tense	132.14	155.63	.000	.000
	Gaze	.931	1.913	.342	.180
	Interaction	2.65	4.09	.113	.055

We further present the post-hoc t-test results in the future tense condition (Table 4.8). While in Experiment 3 the gaze condition did not reveal any effect in the Verb region, the present experiment revealed negative means and a reliable preference to inspect the future target by items and a marginal effect by participants.

By contrast, the results in the no-gaze condition showed positive means and significances in the by participants and items analyses (which means that we replicated the recent-event preference in the Verb region). The data of the gaze condition in the Adverb and in the NP2 regions showed further negative values and reliable effects in the by participants and by items analyses (meaning the future event preference). The results in the no-gaze condition revealed reliable effects (i.e., that the log-ratio means for the future condition differed reliably from zero) in the NP2 regions but no effect was found in the Adverb region. Thus, with earlier situation gaze cue onset the present experiment revealed an earlier effect (already at the Verb in the gaze condition), but when the gaze cue was absent similarly to Experiments 1 to 3, the *t*-tests in the present experiment failed to find a reliable effect in the Adverb region (neither a preferential inspection towards the recent target nor towards the future target).

TABLE 4.8: *One-sample two-tailed t-tests on the log ratio means for the future condition, testing whether these are significantly different from 0 by regions, Experiment 4*

Regions	**Gaze**	**t1 (df=1,31)**	**t2 (df=1,23)**	**P1**	**P2**
Verb	Gaze	-1.76	-2.43	.089	.023
	No Gaze	3.81	5.09	.001	.000
Adv	Gaze	-3.07	-3.07	.004	.005
	No Gaze	.204	.401	.840	.692
NP2	Gaze	-6.61	-5.63	.000	.000
	No Gaze	-5.21	-5.02	.000	.000

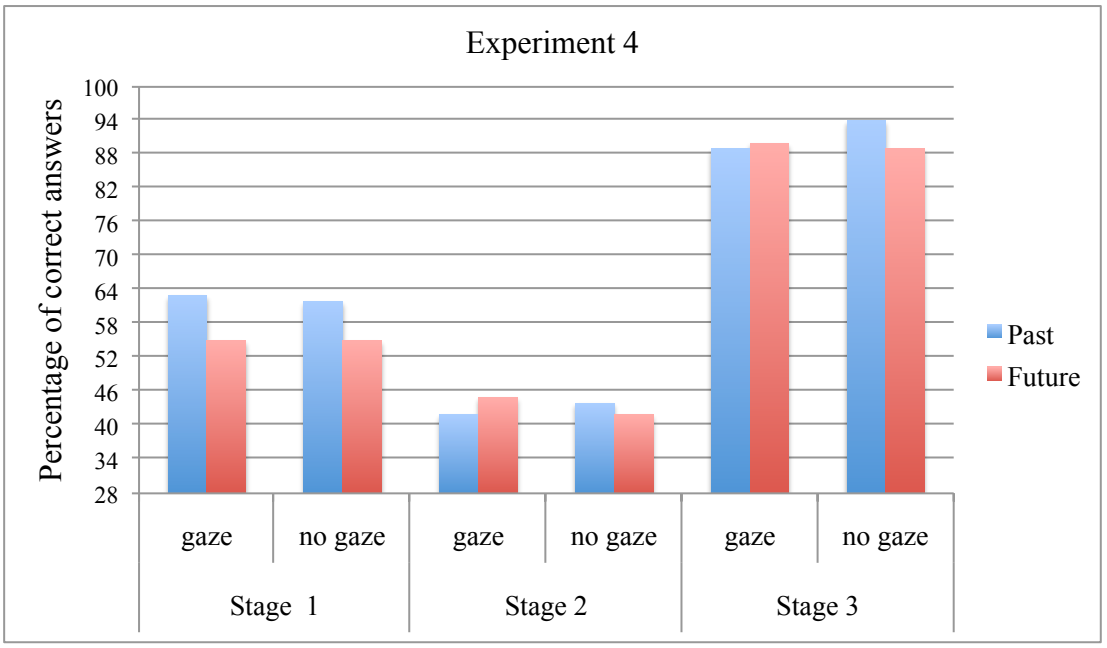
Importantly, the conclusions emerging from the descriptive analyses above are in line with the inferential analyses of the data. The intercepts were significant in all three regions by participants and by items $ps < .047$, except by participants in the NP2 region, which revealed only a marginal effect $ps < .068$ (see Table 4.7). Thus, the current manipulation had an earlier and stronger effect but it could not override the overall recent-event preference. In contrast to Experiment 3, no clearly reliable interaction of gaze and tense emerged in the current experiment. In other words, the gaze cue effect did not differ significantly in the past and future tense condition.

Overall, the results revealed a tense effect in all three regions and a gaze effect only in the Verb region. The current findings support previous findings (from Experiment 3 and others) that the gaze cue had a short-lived effect which did not

show a significant effect in the last two regions of the sentence (the adverb and the NP2 regions).

Gated memory test: Participants overall answered 64% of the questions correctly from all three stages. Figure 4.5 shows that participants correctly recalled 59% of the questions at stage one, which was more than at stage two with 43%. The accuracy was highest at stage three with 90% (see Fig. 4.5, stages 1, 2 and 3).

FIGURE 4.5: *Percentage of correct answers in three stages as a function of tense and gaze*



A further LME analysis for stage 1 and 2 did not show any effect of gaze, however the tense effect was marginal .10 (advantage for the past tense) at stage 1. As we expected, stage 3 revealed an effect of tense ($p < .003$, higher accuracy for past than future tense conditions). The latter result suggests that the past tense sentences about the recently perceived action events are remembered better than the future tense sentences which referred to the plausible future action events.

A main effect of gaze confirmed higher accuracy ($p < .01$, higher accuracy without than with gaze). The low accuracy in the gaze condition might be explained by means of the eye-tracking data in which the gaze cue effect was not significant in the later sentence regions. So, the short-lived gaze cue effect did not support short-term memory of the sentences. Similar effects were found in Experiment 3 and in studies by Knoeferle and Kreysa (2012). Additionally, the data did not show any interaction between gaze and tense.

TABLE 4.9: *Linear mixed effect model results for the gated memory test, Experiment 4*

Stages	Effects	Coefficient	SE	z-Value	P
Stage 1	Intercept	0.46	0.14	3.36	.000
	Tense	-0.17	0.11	-1.61	.103
	Gaze	0.00	0.12	0.06	.950
	Interaction	-0.00	0.09	-0.01	.992
Stage 2	Intercept	-0.27	0.25	-1.09	.273
	Tense	-0.00	0.09	-0.04	.972
	Gaze	0.01	0.12	0.15	.880
	Interaction	0.03	0.09	0.32	.750
Stage 3	Intercept	3.19	0.31	10.26	.002
	Tense	-0.51	0.18	-2.88	.001
	Gaze	-0.47	0.20	-2.34	.013
	Interaction	0.26	0.21	1.24	.227

In summary:

The situational gaze cue manipulation in Experiments 3 and 4 has shown an effect on the listeners' visual attention and has somewhat diminished the recent-event preference. The effect of gaze partly emerged in the post-experimental memory test results. In light of our findings, on the one hand, the recent-event preference seems to be sensitive to situation-immediate cues such as gaze. The gaze cue seemed to modulate the recent-event preference more strongly than the frequency bias that we had established in Experiments 1 and 2. Moreover, Experiments 1 to 3 all revealed a significant intercept in the NP2 region (indicating a preference for the recent target object). However, the intercept was only marginal by participants in the NP2 region of Experiment 4.

On the other hand, except for the NP2 region, the recent-event preference has been replicated in all regions in both experiments. In Experiment 3, when the gaze cue onset started after the Verb onset, the effect of gaze occurred late in the beginning of the Adverb region. But the effect was strong leading to an increased negative log-gaze ratio value in the future tense gaze condition and a significant gaze effect in the Adverb region. In Experiment 4, when the gaze cue appeared at the same time as the Verb onset, the effect of gaze occurred immediately. Participants began to attend to the future target object more than to the recent target in the future gaze condition - as early as 100 ms after the onset of the gaze shift. The early inspection pattern was supported by statistical analyses: the t-test for the future

tense gaze condition confirmed a significant effect of preferential looks towards the future event target by items and a marginal effect by participants in the Verb region.

Taken together, the results provided evidence of the gaze cue manipulation. Gaze had a much stronger effect in enhancing looks to the future than to the recent target. Thus, it did mitigate the recent-event preference but it did not completely override it. The memory effect is not in complete agreement with the eye tracking data: While gaze (vs. no-gaze) was beneficial in enhancing looks to the future target, it was not as beneficial in enhancing memory of future targets in the memory test of Experiment 3. In fact, it was detrimental. This suggests that the actor's gaze is only used on the fly and that its effect on memory is short-lived (note that there is a similar result in the gated memory test of Experiment 4). In the gated memory test, the past tense sentences were recalled better than the future sentences (in agreement with the recent-event preference). This suggests that there might be a linguistic component to this preference (past tense sentences may be better than future ones in grounding an event in memory).

Although the manipulation of gaze has increased the preferential looks towards the future target in the future gaze condition, participants in the no-gaze condition started to inspect the future target object only in the middle of the Adverb region. Thus, we know now that the recent-event preference can be affected at some level by various manipulations but the future event target was inspected nonetheless later than the recent target. What is the cause of the late attention towards the future event target and the overall preference of the recent target? Does the verb guide participants' attention towards the recent event target? Perhaps the fact that the recent action event and the past tense verb are available at an early stage of the sentence and that they can be integrated results in the preferred inspection of the recent event target? We examined these questions in Experiment 5 by changing the stimuli such that the recent event and the past tense verb mismatched.

Chapter 5

Effects of Verb-Action Mismatches during Sentence Processing

Experiment 5 investigated the effects of incongruence between the past tense verb and the recent event during language comprehension. To recall, Experiments 1 and 2 had instantiated a strong frequency bias in favor of the future and against the recent event. Despite the strong frequency bias towards the future event the results replicated the overall recent-event preference. Similarly, Experiments 3 and 4, in which a situated gaze cue was pitted against the recent-event preference, found preferential looks towards the recent event. In all four experiments (except the gaze condition in Experiments 3 and 4) the preferential inspection of the recent event persisted until the middle of the adverb region. In other words it occurred about 1900 ms from the verb onset. While the length of the verb region was only 1148 ms on average, participants not only seemed to ignore the verb tense information but also the temporal information provided by the adverb. Since comprehenders are sensitive to the linguistic input, we conducted a further experiment to assess what might be the cause of people's tendency to seemingly ignore tense cues. Therefore, the present experiment focused on examining incongruence between the past tense verb and the recent event as a potential cause of the the recent-event preference. Do participants miss the future tense verb information only because of an influence of the recently observed action event, which was referred to by the past tense sentence or verb in half of the trials? Alternatively, humans might have a bias towards the recently perceived or described action scene irrelevant of

linguistic information (e.g., Altmann & Kamide, 2009). Furthermore, is it possible that the recent-event preference is elicited by the verb (e.g., Abashidze et al., 2011; Knoeferle & Crocker, 2007)?

These questions motivated the present experiment. We wanted to examine whether participants will preferentially inspect the recent event when the recent action event and the past tense verb mismatch referentially. Recall the reviewed studies in chapter 2 regarding the processing of incongruence during comprehension. In language comprehension research, many studies have employed picture-sentence incongruence and verification as a method. Experiments using this method, have found that participants are sensitive to the incongruence (Carpenter & Just, 1975; Knoeferle, Urbach, & Kutas, 2014; Underwood et al., 2004). Participants were faster to respond to congruent than incongruent picture-sentence pairs (Just & Carpenter, 1971; Underwood et al., 2004). In an eye-tracking study by Knoeferle and Crocker (2005) participants were presented with depicted scenes and either matching or mismatching np1-verb-adv-np2 sentences. The gaze data showed an incongruence effect in the verb and adverb regions. Participants were faster reading the sentence regions in the congruence than in the incongruence condition (see also related findings on gender stereotype effects in a picture-sentence verification task by Rodríguez, Burigo, & Knoeferle, 2015). Other studies used the sentence-picture verification procedure while participants were presented with positive and negated sentences (Glenberg, Robertson, Jansen, & Johnson-Glenberg, 1999). Results showed that pictures matching the presented sentences (even when the sentences were negated) elicited faster responses than pictures mismatching the sentences. In addition, participants were slower in responses when the sentence was negated than positive. In another experiment Stanfield and Zwaan (2001) reported faster responses to matches than mismatches between the orientation of objects depicted in a scene and the orientation described by a sentence. Furthermore, participants showed their sensitivity to the shape of the object. They responded more quickly when the shape of the targets in the scene and the shape described by the sentence mismatched than when it matched (i.e., Zwaan, Stanfield, & Yaxley, 2002). Furthermore, event-related brain potential studies found an N400 and later negativity effect when audio-visual information mismatched (e.g., Liu, Wang, & Li, 2011). These findings and the results of Experiments 1 to 4 made us ask, whether incongruence between the recent event and the past tense verb would affect the recent-event preference. If this is the case, will it decrease participants' attention towards the recent event target, so that the overall

recent-event preference disappears?

If the recent-event preference is sensitive to the incongruence of visual and language information, then we should see an early decrease in looks towards the recent event target and accordingly, more and earlier inspections should occur to the future event target. In comparison with the other experiments of the present thesis, we might observe an earlier gaze shift towards the future event target as a result of the incongruence. However, if there is an epistemic bias towards the recent event target, and the verb (simple past) does not determine the visual attention towards the recent event then we should replicate the recent-event preference even when the verb mismatches the recently seen action.

In the present experiment we used the materials from Experiment 1 for the event-sentence match. For the event-sentence mismatch, we recorded new videos. For the critical items the recent event and past tense verb mismatched, but the future tense sentence and the future event matched. However, the filler items presented some recent event and past tense matches to balance the match/mismatch event-sentence pairs over the experiment. As in Experiment 4 the present experiment conducted a post-experiment gated memory test.

5.1 Experiment 5

The aim of the present experiment was to examine potential effects of the incongruence between the recent action and the past tense verb on the recent-event preference. As in experiment 2 by Knoeferle, Carminati, et al. (2011) recent and future events as well as the sentences in the past and in the future tense were presented equally often. However, in the present experiment, the critical past tense verb always mismatched the recent action event. By contrast, the future tense sentence always matched the future action event. We assessed to what extent the incongruence of the recent event and the past tense verb would reduce the preferential inspection of the recent event target during the verb (i.e., when do people realize that the verb does not match the action they had just seen?).

The preferential inspection of the recent even target should start to decrease from the early verb region. We know that humans are sensitive to incongruence and that they rapidly process visual and language information when they match. Thus, any preferential looks towards the future event should occur in the early sentence

regions. When participants had seen an action event and listened to a past tense sentence, in which the verb mismatches that recent event, they might start to search for another suitable object in the scene that would be a plausible target for the action expressed by the verb. On this account, they might inspect the future event target often in the search for a suitable referent and in response to the incongruence between the past tense verb and the recent event. The incongruence in the past tense condition might strengthen the congruence in the future tense condition (e.g., Glenberg et al., 1999; Stanfield & Zwaan, 2001) such that participants inspect the future event target already in the verb region, when the verb matches the plausible future event. Alternatively, earlier looks towards the future event target might fail to appear. Participants might not immediately respond at the offset of the verb to the incongruence (since for the majority of the filler items the past tense verb did match the visual context).

Alternatively, if the recent event inspection is not guided by the verb meaning, participants might continue to inspect the recent event target irrespective of the verb mismatch. The incongruence of the recent event and the past tense verb might not change participants' behavior in their overall preferential inspection of the recent event. If so, we should replicate our previous findings.

In addition to the eye-tracking study, participants performed a gated memory test as in Experiment 4. The aim of the test was to examine whether participants would recall the future tense sentence (match) better than the past tense sentence (mismatch between the recent event and the past tense verb). If participants spent less time inspecting the recent event target because of the mismatch between the verb and the recent action, then we might find a corresponding reduction in recalling the recent event target in contrast with our previous experiments.

5.1.1 Method

Participants

A further thirty-two native German speakers took part in Experiment 5 (Mean age=24; range: 18 to 32; 13 males and 19 females, all right-handed). All participants were from the Bielefeld University community, and they were each paid 6

Euros for their participation. None of them had learned a second language before age 4. All had normal or corrected-to-normal vision and were unaware of the purpose of the experiment.

Materials and Design

Eye-tracking: The experiment used the same experimental materials as the previous experiments except for one change. We recorded new critical videos (for the recent events), for which we used the same target object but the actions were different. For example, instead of sugaring strawberries, a video showed the experimenter tasting the strawberries (see Fig. 5.1., A). Then, a sentence was presented e.g., *Der Versuchleiter zuckerte gerade die Erdbeeren*, 'The experimenter sugared recently the strawberries', so that the past tense verb never matched the recent action event. By contrast, the corresponding future tense sentence always matched the future event e.g., *Der Versuchleiter zuckert demnächst die Pfannkuchen*, 'The experimenter sugars soon the pancakes' and a video showed the experimenter sugaring the pancakes (see Fig. 5.1., C). In half of the critical trials the sentences were in the past (simple past verb plus past-referring adverb) and the verb mismatched the recent event. In the other half of the trials the sentences were in the future (simple present tense verb plus future-referring adverb) and the sentence matched the future event (shown after the sentence had ended). Thus, the experiment manipulated one factor: sentence tense (past vs future). Past and future events were shown equally often.

The filler items were added to balance congruous and incongruous action-sentence pairs across the trials. They included incongruence between the future tense verb and the future action to ensure verb-action mismatches were balanced across tense within the experiment. In order to decrease participants' potential bias of inspecting the recent target object because it would be mentioned by the second noun phrase, we created a few fillers in the past tense in which the recent event target (e.g., a thermos flask) was not mentioned in the past tense sentence.

We expected a decrease of attention towards the recent event target when participants process the incongruence. When hearing the first noun phrase 'The experimenter', participants will mostly look at the experimenter. The gaze pattern towards the experimenter should not differ between future and past tense sentence conditions. When hearing the verb 'sugared' in the past tense condition during that period, it should become clear to listeners that the verb does not match

the recent event. Therefore, from that time onwards eye-movements towards the recent event location should reduce.

On hearing the verb 'sugars' in the present tense referring to the plausible future action event, we should see an increase of looks towards the future event target. After the mismatch experience from the past tense condition, participants might react faster for verb-action matches and they might inspect the future event target more often and earlier in the present experiment than in previous experiments (Knoeferle, Carminati, et al., 2011, experiment 2). Thus, an increase in eye-movements towards the future event target should appear at the end of the future tense verb region.

On hearing the temporal adverb in the past tense condition 'recently' the eye-movements should continue to reflect a reduced recent event target inspection. By contrast, when participants hear the temporal adverb in the future tense condition 'soon' they should look more at the future event target than at the recent event target.

When the second noun phrase of the past tense sentence 'the strawberries' is encountered, the eye-movements towards the strawberries might increase. But it is plausible that the object might be inspected later, since it is the target of an action that did not match the verb. By comparison, when the second noun phrase in the future tense sentence 'the pancakes' is heard, the fixations to the pancakes should clearly continue to increase until the end of the sentence. This should even result in significantly more looks to the future event target in the future tense than in the past tense sentence. Crucially, to the extent that the incongruence of the recent action event and the verb of the past tense is effective, it should decrease the recent-event preference and it might even override its overall bias.

The experimental and filler items were combined to form 4 lists using a Latin square design. Each list contained every critical item (N=24) in only one condition and all fillers (N=40). Before the experiment, lists were pseudo-randomised and each participant saw an individually randomised version of one of the four experimental lists.

Gated memory test: In addition to the eye-tracking session, participants performed a gated memory test that was identical to the one in Experiment 4. They were asked to recall the past and future tense sentences in three stages. We wanted to find out whether the incongruence of the recent action event and the past tense

verb would affect participants' memory in recalling the critical sentences. Our expectation is that in the gated memory test we might find higher accuracy for the recall of the future tense sentence (match) than for the past tense sentence (mismatch).

If the congruence of verb and action plays an important role, we should see better accuracy in recalling the future tense sentence than the past tense sentence at the first stage. Similar results should occur at the second and third stages. However, if participants encode the incongruence for the past tense sentences, an alternative possibility is that they will recall the mismatching past tense sentences better than the matching future tense sentences. If this is the case, we might find that participants are more accurate in recognising the past tense sentences than the future ones at all stages of the memory test.

After the gated memory test, participants were debriefed. The experiment lasted approximately 40-45 minutes.

Procedure and Analyses

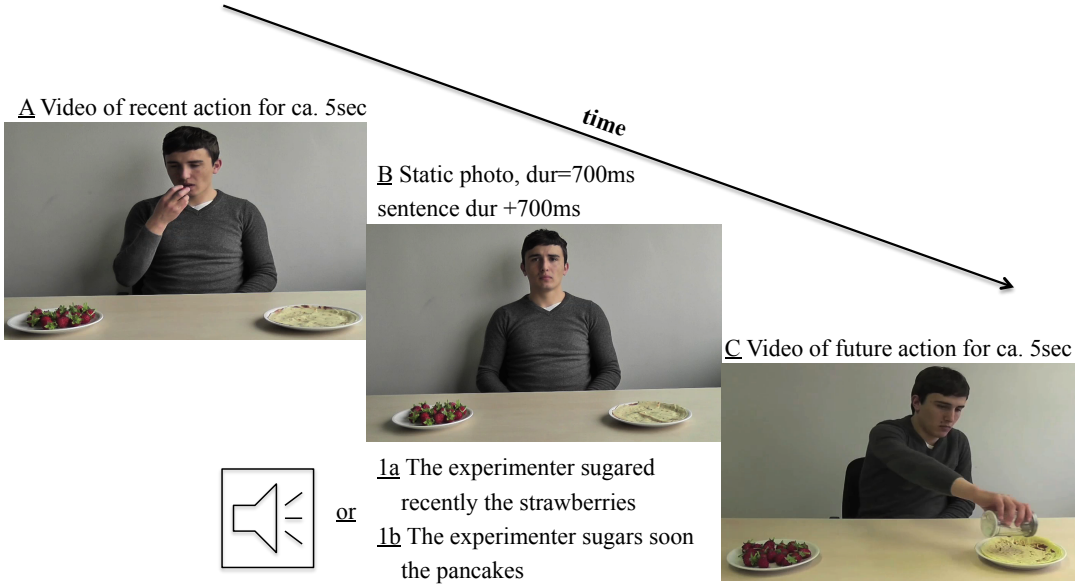
Eye tracking: The procedure and analyses were similar to Experiment 1. Participants saw the videotaped actor perform one action before the sentence (e.g., tasting strawberries, Fig. 5.1., A) and then they saw a static photo (see Fig. 5.1., B). 700 ms after the onset of the static photo, a German sentence was presented via speaker, and the sentence was either in the past tense condition e.g., 'The experimenter sugared recently the strawberries' (literal translation of the German original) or in the future tense condition e.g., 'The experimenter sugars soon the pancakes' (see Fig. 5.1., literal translation of the German original). After the sentence, a second video was played in which the actor performed an action involving the other object (Fig. 5.1., C).

Gated memory test: The procedure and analyses in the present experiment were identical to the one in Experiment 4.

5.1.2 Results and Discussion

Eye tracking: We first report the descriptive analyses as time course graph of the eye-movements in Figure 5.2 and then the inferential analyses of the word regions in Table 5.2. Figure 5.2 presents the time course of looks to the recent target and

FIGURE 5.1: *Sequence of events of a typical experimental trial, Experiment 5*



future target objects from the Verb onset until the end of the sentence. Table 5.1 displays the log gaze probability ratios of the individual regions in past and future tense conditions, including the grand mean. Before turning to the results, let us first briefly recall the incongruence manipulation of the experiment. In the past condition, the Verb of the past tense and the recent action event always mismatched. But the Verb of the future tense sentence and the future action event always matched.

The gaze data in Figure 5.2 shows that early during the Verb region attention towards the recent event target in both sentence tense conditions was relatively low. In the middle of the Verb region listeners looked more at the recent target object in the past tense than in the future tense condition. But at the end of the Verb region there was no difference in looking at the recent target in the past compared to the future tense condition. While the amount of visual attention towards the recent target object decreased at the end of the Verb region, the log-gaze probability ratio value remained above zero, meaning more looks went to the recent (vs. future) target object. Surprisingly, the positive value continued in the Adverb region in both sentence conditions. However, in the present experiment the past tense Adverb region value .44 (see Table 5.1) is substantially lower than the past tense adverb region value 2.13 from Experiment 1 (see Table 3.2). In

contrast, the future tense Adverb value .28 (see Table 5.1) is very similar to the adverb value .35 from Experiment 1 (see Table 3.2).

The eye-movement record for the Adverb region for future tense sentences in Experiment 5 was very similar to the corresponding gaze pattern in the experiments that manipulated event frequency (Experiments 1 and 2). Furthermore, the fixation pattern in the first half of the Adverb region gives the impression that the recent target object is inspected equally often in the past and future tense conditions. Afterward the looks towards the recent target start to increase in the past tense condition. It is plausible that participants learned that in the past tense condition -in spite of the mismatch of action and verb- the sentence often correctly mentioned the object of the event. Perhaps this is the reason for increasing looks towards the recent target in the NP2 region. Moreover, the few items in the past tense fillers for which the NP2 mismatched an object in the visual context may not have been sufficient to prevent participants from inspecting the recent target object even before its mention. By contrast, in the future tense condition the eye-movements towards the recent event target decreased at the end of the Adverb region and from the onset of the NP2 participants inspected the future target object more in the future tense sentence than in the past tense sentence and vice versa.

The time-course graph data suggests that the incongruence manipulation did reduce the recent-event preference in the Verb region in Experiment 5 (compared to previous experiments in which even more looks went to the recent target object during the verb). However, the gaze pattern in both past and future tense conditions replicated the recent-event preference (both lines are above zero) until 2600 ms after the Verb onset. Despite the mismatch, a clear increase of fixations to the future event target occurred only at the beginning of NP2. Moreover, the grand mean was positive in all three regions (but less positive in the NP2 region, see Table 5.1), which means the overall recent-event preference has been replicated in all regions.

After presenting the descriptive statistics, we present now the inferential statistics. The data of the inferential analyses are shown in Table 5.2. The conclusions from the descriptive analyses above were mostly confirmed by the inferential analyses of the data. In the time course graph, the log ratio showed a positive value in all three regions and the inferential analyses supported this with a significant intercept in all three regions by participants and items $ps < .000$ (see Table 5.2).

FIGURE 5.2: *Mean log gaze probability ratios (ln (P(recent target/P(future target))) by condition from verb onset, Experiment 5*

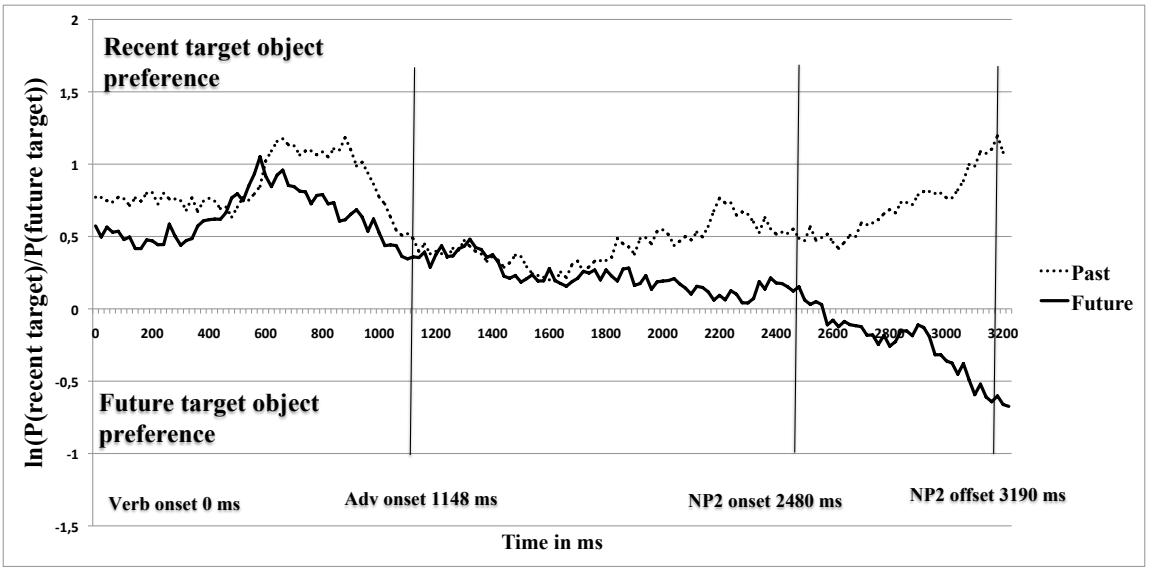


TABLE 5.1: *Grand mean and mean log gaze probability ratios (ln (P(recent target)/P (future target)) by participants as a function of condition and time region. Standard errors (SE) in parentheses, Experiment 5*

Regions	Future tense	Past tense	Grand mean
Verb	.76 (.19)	.92 (.21)	.84 (.17)
Adv	.28 (.25)	.44 (.15)	.36 (.12)
NP2	-.21 (.13)	.82 (.13)	.30 (.09)

We present the inferential analyses for the Verb region. The ANOVA showed a marginal tense effect by items $F2(df{=}1.23) = 3.41$, $p2 = .078$ but not by participants $F1(df{=}1.31) = 23.21$, $p1 = .447$. The absence of a tense effect differs from previous experiments (e.g., Experiment 4), and can be explained as a result of the incongruence manipulation (not being as strong as the other experimental manipulations). During the Verb participants' eye-movements towards the recent event target did not differ between past and future sentence conditions. The absence of tense effects in the Adverb region might have the same cause as in the Verb region. However, we found a significant tense effect in the NP2 region in the by participants $F1(df{=}1.31) = 44.33$, $p1 = .000$ and by items analyses $F2(df{=}1.23) = 22.87$, $p2 = .000$ (see Table 5.2).

Overall, results revealed a marginal effect of tense in the analyses by items for the Verb region and a fully significant effect by participants and items in the NP

TABLE 5.2: *ANOVA analyses for the data by region: The intercept is also given because in this case a significant intercept indicates that the grand mean is significantly different from 0, Experiment 5*

Regions	Effect	F1 (df=1,31)	F2 (df=1,23)	P1	P2
Verb	Intercept	.597	40.21	.000	.000
	tense	23.21	3.41	.447	.078
Adv	Intercept	8.91	14.46	.006	.001
	tense	1.08	2.38	.306	.136
NP2	Intercept	7.97	9.82	.008	.005
	tense	44.33	22.87	.000	.000

2 region. More inspection to the recent target object than future targets objects occurred in the past tense than in the future tense.

Moreover, the post-hoc *t*-test on the log ratios from the future tense condition are presented in Table 5.3. The data in the Verb region revealed positive means and reliable effects in the by participants and items analysis (meaning we replicate the recent-event preference). Interestingly, the results in the Adverb region revealed a reliable effect also in favour of the recent event. This outcome for the Adverb region differs from the *t*-test results obtained in Experiments 1 to 4 for which no reliable effect was observed (except in the gaze conditions). Surprisingly and in contrast with Experiments 1 to 4, the present experiment failed to reveal a significant effect of the future event preference in the future tense condition (even in the NP2 region). These results suggest that the incongruence manipulation seems to have a comparatively weak effect on the recent-event preference.

TABLE 5.3: *One-sample two-tailed t-tests on the mean log ratios by word region for the future condition, Experiment 5*

Regions	t1 (df=1,31)	t2 (df=1,23)	P1	P2
Verb	3.96	3.97	.000	.001
Adv	2.01	2.36	.053	.027
NP2	-1.54	-1.29	.135	.212

Gated memory test: Participants correctly answered 61% of the questions when averaging across all three stages. Figure 5.3 displays the data showing that participants correctly recalled 54% of questions at stage one, which is more accurate than at stage two with 40%. Importantly, the highest accuracy is revealed at stage

three (89%, see Fig. 5.3, stages 1, 2 and 3). As can be seen from the graph, subjects recalled the future tense sentence (match) better than the past tense sentence (mismatch). The current findings are in line with the findings of Glenberg et al. (1999).

FIGURE 5.3: *Percentage of correct answers as a function of tense and stages*

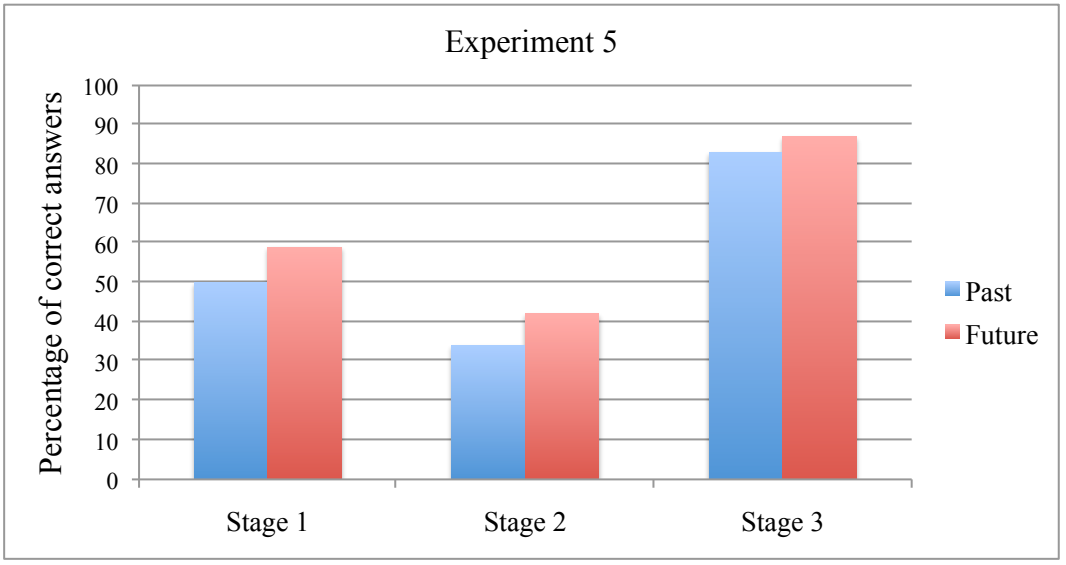


Table 5.3 presents the LME results from the gated memory test. Linear mixed effect model analyses showed a marginal tense effect at stage 1 $p < .080$ and a fully significant tense effect at stage 2 $p < .020$, indicating a higher accuracy for the future than for the past tense sentence condition. Even though stage 3 descriptively shows a future tense sentence preference, this was not confirmed by the statistical analyses. These findings contrast with the findings in Experiment 4 where the marginal tense effect in the verb was in favour of the past (vs. future) tense condition. While in Experiment 4 the tense effect was significant at stage 3, the present experiment did not show significance for the tense effect at stage 3.

TABLE 5.4: *Linear mixed effect model results for the gated memory test, Experiment 5*

Effects	Stages	Coefficient	SE	z-Value	P
Intercept	1	0.16	0.09	1.68	.091
Tense	1	0.19	0.11	1.73	.080
Intercept	2	-0.53	0.27	-1.96	.041
Tense	2	0.26	0.12	2.25	.020
Intercept	3	2.20	0.14	15.62	.000
Tense	3	0.25	0.16	1.55	.120

Summary

The present experiment revealed that participants decreased their attention towards the recent event target from the middle of the Verb region onwards (e.g., in comparison with Experiment 1). However, they inspected the recent target object more than the future target object in the Verb and Adverb regions independent of tense (see Fig. 5.2). In this respect, our findings showed on the one hand that the recent-event preference is sensitive to the incongruence (motivated by descriptive comparisons across the experiments) but on the other hand, they underscore the robustness of the recent-event preference. In comparison to Experiment 1, the log-gaze probability ratios value of the adverb regions in both tense conditions are lower than in the present experiment. By contrast, the analysis of the present experiment revealed an overall recent-event preference. Thus, the present findings replicated the findings of the other experiments (see, chapters 3 and 4). Interestingly, the gated memory test showed that participants were more accurate in recognizing the future tense sentences (match) than the past tense sentences (mismatch).

Taken together, these findings provide evidence of an influence of incongruence on the recent-event preference during spoken sentence comprehension. However, the incongruence effect appeared to be more beneficial for the later memory test than for the immediate sentence comprehension. This is an interesting finding that needs further examination.

Chapter 6

General Discussion

The present thesis examined the importance of recent action events for language comprehension. The reported research situates itself in the context of extant findings, supporting the view that both the visual and linguistic context can rapidly contribute towards language processing. In the following we first motivate the research question on the present thesis and then summarize the experiments and the key results.

Recall first the findings reported in the last section in Chapter 2 which motivated us to examine the recent-event preference (an inspection preference, suggesting a recent action event is prioritized over the anticipation of a future event target during language comprehension). The other findings reported in Chapter 2 mainly addressed issues of visual context effects during language comprehension. In addition, we reviewed early studies directed at examining referential context effects in situated language comprehension (e.g., Chambers et al., 2004; Eberhard et al., 1995; Spivey et al., 2002; Tanenhaus et al., 1995). Further, we reported recent studies that found an influence of recently perceived events on language comprehension (Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011), and we discussed findings of incremental visual context effects. For instance, comprehenders can rapidly integrate information about the shape, size, and color of an object into language comprehension (e.g., Dahan & Tanenhaus, 2005; Huettig & Altmann, 2011). Following this overview, we have reviewed studies that examined the role of linguistic cues during language processing. For example, we saw that tense information provided by a verb or the meaning of a preposition can help comprehenders to predict an upcoming noun

phrase referring to an object in a scene (e.g., Altmann & Kamide, 1999, 2007; Chambers et al., 2002; Sedivy et al., 1999). Moreover, we discussed the role of depicted actions that are able to modulate listeners' eye-movements in incremental language comprehension (Knoeferle et al., 2005; Knoeferle & Crocker, 2006). In addition, we reviewed literature on the role of attention in the visual and linguistic context and its effect on the ensuing recall of information in memory tests (e.g., Luck & Vogel, 1997; Richardson et al., 2003; Trueswell & Papafragou, 2010; Zelinsky et al., 2011).

Taken together, all of the reviewed findings provide strong experimental evidence of non linguistic and linguistic context effects in language processing. Recall some findings in support of the 'anticipation account'; for instance, when participants viewed a scene depicting objects and listened to a verb that related to a particular object in the scene, their anticipatory eye-movements towards the object emerged before the object was named (e.g., hearing 'eat' made it possible to anticipate an edible object Altmann & Kamide, 1999, 2007; Kamide, Scheepers, & Altmann, 2003). Similar to exploiting the verb, participants were able to anticipate a target object based on the meaning of an adverb or of a proposition (e.g., Chambers et al., 2002; Sedivy et al., 1999). Furthermore, comprehenders anticipated objects in the scene through the immediate depicted action; when they listened to a sentence about characters involved in depicted actions, the verb-mediated actions permitted them to inspect a target character early, before it had been mentioned (Knoeferle et al., 2005; Knoeferle & Crocker, 2006).

However, how strong the effect of an action event is during language processing has been less clear. Previous findings revealed that participants process ongoing events faster than completed events, when they read sentences presented in the past perfective and in the past imperfective alongside with events (Madden & Therriault, 2009). Moreover, a description of an event modulates participants' eye-movements even if the description of the event conflicts with the actual scene depiction (Altmann & Kamide, 2009). Comprehenders preferentially inspect a recently seen event target over a plausible future target (Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011). In the latter experiments, participants always saw an action event before the sentence and another action event after the sentence and they listened in half of the trials to a sentence referring to the recent event target and in the other half of the trials to a sentence referring to the plausible future event target. Their preferential

inspections towards the target of the recent event appeared from the beginning of the sentence and that gaze pattern continued until the middle of the second noun phrase even in the future tense condition (Abashidze et al., 2011; Knoeferle, Carminati, et al., 2011, experiment 2).

The present thesis addressed the issue of the recent-event preference by introducing manipulations such as a strong frequency bias in favour of the future event; in addition, we examined the effects of a cue that was part of the immediate situation (the actor's gaze towards the future action target), and of incongruence between the recent event and the past tense verb (to see if a mismatch would reduce the inspection of the recent event target). Using the visual-world paradigm method (recorded short-videos using real-objects) we investigated the recent-event preference during spoken sentence comprehension. Participants saw short-video sequences before and after the sentence and during the sentence they saw an actor (who had just performed an action event) in a static position and two related objects (e.g., tomatoes and cucumbers) were located on the table. All experimental sentences had a NP1-Verb-Adv-NP2 order and they were either in a past or in a future tense condition. Participants' eye-movements were monitored while they listened to sentences e.g., 'The experimenter flavoured recently the cucumbers' in the past tense condition with a simple past tense verb plus a past adverb referring to the recently perceived action event; or they listened to another sentence such as 'The experimenter flavours soon the tomatoes' in the future tense condition with a present tense verb plus an adverb with a future meaning referring to the plausible future event (original German sentences are displayed in Table 1). In the sentences, the verb plus time adverb denoted an action, which was equally plausible in terms of verb-related knowledge for the recent and future events.

The present experimental manipulations tested the robustness of the recent-event preference and were motivated by previous findings discussed (see Chapter 2). A first set of findings related to the frequency of linguistic input. For instance, words that are more frequent are processed faster than comparatively infrequent words (e.g., Arnon & Snider, 2010; Dahan, Magnuson, & Tanenhaus, 2001; Wells et al., 2009). In addition, the frequency of visual context modulated the time course of eye-movements in another experiment (e.g., Knoeferle et al., 2011, experiment 2). Drawing on the insight that more frequent linguistic and visual cues can modulate language processing, Experiments 1 and 2 in the present thesis increased the number of future events and future tense sentences in order to test the robustness of

the recent-event preference. Encountering more future than recent events, might bias comprehenders to make more eye-movements towards the plausible future event target. With the frequency manipulation we expected an early gaze shift to the future event target after the tense information became available. The anticipatory eye-movements in the future tense condition should emerge earlier than in experiment 2 by knoeferle et al. (2011) and accordingly a reliable tense effect should be revealed in the early sentence regions.

The second manipulation in the present thesis has previously been reported as being very effective in directing visual attention during language comprehension (e.g., Hanna & Brennan, 2007). Studies examining the effects of a speaker's gaze have shown that participants followed the gaze cue shortly after its onset to resolve an ambiguity in a language context (see Chapter 2). The gaze shift affected an interlocutor's attention even when the speaker was not facing the listener (e.g., Knoeferle & Kreysa, 2012). Furthermore, participants followed the gaze cues of a robot to accelerate anticipation of a target object during an ambiguous sentence (Staudte et al., 2014). In order to apply this manipulation in Experiments 3 and 4, we pitted a situation-immediate cue (an actor's gaze) against the recent-event preference. The actor's gaze cue onset appeared 480 ms after the verb onset in Experiment 3 and at the verb onset in Experiment 4. As the previous findings had revealed a strong and immediate effect of the gaze cue, we expected an immediate gaze effect against the recent-event preference. We especially thought the gaze effect might show its impact in the future tense condition more than in the past condition (in the past tense condition, participants already prefer to inspect the recent event target). Thus, with this added situation-immediate cue, the overall recent-event preference should disappear.

The third factor that we pitted against the recent-event preference was the incongruence between the recent action and the sentential verb. Previous studies that have examined the effects of scene-sentence incongruence in the verification task found that participants were faster in processing a matching than mismatching picture-sentence pair (e.g., Carpenter & Just, 1975; Knoeferle, Urbach, & Kutas, 2011; Underwood et al., 2004). In Experiment 5 with the verb-action incongruence (arguably rendering the recent action irrelevant), we expected an early decrease of looks towards the recent event target. On the other hand the match of the future tense sentence and the future action event should elicit an anticipatory gaze pattern towards the future action target at an early point in the sentence (e.g., from

the end of the verb region).

In addition to eye-tracking experiments, memory or gated memory tests were conducted at the end of each experiment. The aim of the memory tests was to find further evidence that might support or refute the gaze data in relation to the recent-event preference. Experiments 1 to 3 asked participants to indicate which action (out of two seen action events) they had seen before and which action after the sentence. Experiments 4 and 5 asked participants to recall the linguistic content of the past and future tense sentences.

In the following section we first summarize and discuss the key findings of the present experiments. Furthermore, we compare the experimental findings across the five studies and interpret the results. Moreover, we discuss the sensitivity of the recent-event preference to the experimental factors, as well as its robustness (section 6.1). On the one hand, the overall recent-event preference was replicated despite the powerful experimental biases and cues that we pitted against it. On the other hand, the robustness of the recent-event preference seems to have its limit when pitted against the present experimental manipulations as highlighted in sections 6.1.1, 6.1.2, and 6.1.3.

Summary of Experimental Findings

The results from Experiments 1 and 2 showed that a strong frequency bias towards future events did decrease the preferential inspection of the recent event target. The time-course data in Figures 3.3 and 3.5 shows that when the verb was presented, the preferential inspection of the recent event target replicated as in experiment 2 by Abashidze et al. (2011). However, the lines representing the gaze data in the past and future tense condition began to diverge from the end of the verb region, meaning decreased attention towards the recent event target. Once the adverb had been encountered, the preferential looks towards the future event target in the future tense occurred approximately 700 ms (Experiment 1) and 800 ms (Experiment 2) after the adverb onset. Thus, the future target inspection emerged approximately 1000 ms earlier than in Abashidze et al. (2011, Exp. 2) when future and recent actions were equally frequent. The increasing inspections to the future event target in the future tense condition continued until the end of the sentence. As expected, the gaze data in the NP2 region showed more inspections of the future event target in the future tense condition than in the past tense condition and vice-versa. Furthermore, the frequency effect has

been confirmed by the statistical gaze data analyses. It revealed a reliable tense effect in the adverb and NP2 regions, whereas the effect was marginal in the verb region (Experiment 1). Overall, the early appearance of the future event target inspections in the future tense and the tense effect revealed that the frequency distribution in Experiments 1 and 2 affected the recent-event preference. However, surprisingly, we replicated the overall recent-event preference in both experiments.

Experiments 3 and 4 aimed to test the overall recent-event preference and the increasing inspections of the recent event target in the verb region by pitting the actor's gaze shift as a strong object-based cue starting in the verb region against the recent-event preference. The present experiments further extended and generalised the recent-event preference findings for the first two experiments in the no-gaze condition. Experiments 3 and 4 revealed an effect of the actor's gaze shift, which crucially was earlier and stronger in the future than in the past condition. In the gaze condition, findings from Experiment 3 demonstrated that the effect of the gaze cue occurred approximately 450 ms after its onset. This was evidenced by the finding that the eye-movements towards the future event target in the future gaze condition began to increase already from the end of the verb region. By contrast, in the no-gaze condition the future event target inspection occurred later in the middle of adverb region. The effect of the early gaze manipulation in Experiment 4 provided even earlier support for the increase in looks to the future event target, beginning immediately after the onset of the gaze shift. However, effects of the actor's gaze shift in the future tense condition were not fully reliable until the end of the verb region. From the beginning of the adverb region inspections of the future event target began to increase in the future gaze condition and this continued throughout the sentence in both experiments. Similarly to Experiment 3, inspections to the future event target in the no-gaze condition began to increase only from the middle of the adverb region in Experiment 4 (approximately 1000 ms later than in the gaze condition, see Figures 4.2 and 4.4).

These findings suggest that the actor's gaze shift has had an immediate and strong effect. In addition, it modulated the gaze data in the future tense condition more strongly than in the past tense condition. However, while the immediate situational gaze cue onset occurred early during the sentence (in the verb region) and enhanced inspections to the future event target, the overall recent-event preference persisted in all three word-regions of both sentence tense conditions. Furthermore, we should note that the gaze cue effects began to disappear in the last region of

the sentence, suggesting that actor gaze effects are not long-lasting (Knoeferle & Kreysa, 2012; Kreysa & Knoeferle, 2013, for further discussion).

Experiment 5 further explored whether findings from Experiments 1 to 4 (of the recent-event preference and rising eye-movements towards the recent event target in the verb region) generalise even when pitted against verb-action incongruence, effectively rendering the recent action event irrelevant when the verb mismatches. To recall, participants saw a recent action event (e.g., an actor tasting strawberries) and then they listened to a sentence such as 'The experimenter sugared the strawberries' in which the verb mismatched the recent action event. By contrast, the sentence in the future tense condition matched the future event always. In fact, results showed that the incongruence reduced the recent-event preference. We saw a decrease in attention to the recent event target from the end of the verb region. However, the decreased attention was not strong enough to reverse the preferential looks towards future event targets. Despite the incongruence, preferential inspections to the recent event target occurred until the onset of the second noun phrase. The preferential looks towards the future event target in the future tense condition began only from the beginning of the second noun phrase. Thus, the gaze pattern did not show as strong an incongruence effect as we had expected. However, the incongruence manipulation revealed inspection of the future event target approximately 400 ms earlier than in the experiment 2 by Abashidze et al. (2011). Furthermore, the significant tense effect in the adverb region disappeared in comparison with the experiment by Abashidze et al. (2011, Exp. 2), suggesting that reduced eye-movements were made to the recent event target. Similar to Experiments 1 to 4, we replicated the overall recent-event preference. Thus, the grand mean (i.e., the mean of both conditions) was positive in all sentence regions. Experiments 1 to 5 all revealed participants' overall preference for the recent event target irrespective of tense, which was confirmed by a significant intercept in all word regions of the sentence (significant intercept = grand mean significantly different from zero, see chapters 3, 4 and 5).

At this point we can compare the present experimental findings to see how different manipulations dealt with the recent-event preference. While Experiment 1 with the frequency distribution of 75/25% was able to reverse the recent-event preference by the middle of adverb region in the future tense condition, Experiment 2 with an even stronger bias towards future events and future tense sentences failed to provide evidence for an even earlier effect of future event target inspections.

Furthermore, Experiment 3 in the gaze condition revealed an effect of the actor's gaze shift; however, it was not strong enough to get an early effect of the gaze cue in the verb region. The early gaze onset in Experiment 4 had an early effect but it did not change the overall bias towards the future event target compared to Experiment 3. While in Experiment 3 the gaze onset occurred approximately 480 ms later than in Experiment 4, both experiments elicited a clear gaze cue effect at the beginning of the adverb region in the future tense condition. Interestingly, the no-gaze condition data revealed a time course of eye-movements to the recent event target that was highly similar to that observed in Experiments 1 and 2. The analyses of the gaze data from Experiments 1 to 4 (considering only the no-gaze conditions), revealed that in the future tense conditions participants started to inspect the future event target from the middle of the adverb region, which is approximately 1800-2000 ms after the verb onset. Despite the fact that we observed a decrease in attention to the recent event target even in the past tense condition at the end of the verb region, the effect of the incongruence manipulation in Experiment 5 seemed to be less intense than the effects of the frequency biases and of the gaze cue manipulations: While the preferential inspection of the future event target in the future tense condition began already in the middle of the adverb region in Experiments 1-4, this gaze pattern emerged only from the onset of the NP2 region in Experiment 5.

Perhaps the mismatch of only the verb and the action was not sufficiently strong to completely eliminate the recent event target preference in the adverb region. Further studies are needed to test whether a more extended mismatch (e.g., including verb and NP2 in the sentence and the acted upon action and its object in the visual context) could decrease the bias towards the recent event target and whether participants will start to anticipate the future event target in the future tense condition earlier than observed in Experiment 5, viz during the adverb region.

The memory test results provide some support for the view that the recent-event preference may have a functional basis (encoding recently seen action events in memory through subsequent reference). A better recognition of the recent event than the future event would be compatible with the overall recent-event preference. However, a better memory for the recent action event emerged in Experiment 2 but not in Experiment 1. The findings suggest that seeing more future events than recent events did not lead to an improved recognition of future action events.

However, the inspected recent events and its object were recalled better, at least in Experiment 2, which is in agreement with previous related findings (e.g., Zelinsky et al., 2011) of improved recognition for re-inspected (vs. not re-inspected) target objecta in a forced-choice task. By contrast, in Experiments 3 and 4, while analyses of the data in the gaze conditions revealed a strong effect of comprehenders looking earlier and more to the target objects in both future and past tense conditions, the memory test data did not show a better task performance in the gaze condition. Interestingly, in Experiment 3 the action events from the no-gaze condition were recognised better than the action events from the gaze condition, which was confirmed by the statistical analyses (see Table 4.4). Likewise, the linguistic information in Experiment 4 from the no-gaze condition was recalled better than from the gaze condition.

These findings suggest that participants rapidly used the gaze cue in directing visual attention during language comprehension; however, the effect does not seem to last as long (which can be seen in the time course graphs Figures 4.2 and 4.4). Moreover, the short-lived gaze cue effect in the gaze data agrees with the memory test results. In addition to overall better memory in the no-gaze condition, the sentences in the past tense condition were recalled better than sentences in the future tense condition, which is in line with the recent-event preference. While in Experiments 1 to 4 the manipulations provide evidence for the recent-event preference in the gaze data, by contrast, analyses of the memory test data revealed no strong evidence for corresponding effects in the memory data (i.e., recognition of recent actions was mostly no better than of future actions). Furthermore, the results of Experiment 5 showed a different response in the memory test. Participants performed better in the future tense condition than in the past tense condition, which was confirmed by statistical analysis.

All in all, the present thesis results clearly contribute new evidence regarding the recent-event preference. Even though we knew about the rapid influence of visual context (e.g., depicted events) on incremental language processing, we expected that our strong experimental manipulations might essentially mitigate, abolish, or even reverse the recent-event preference. The present findings show that this is mainly not the case and we interestingly replicated an overall recent-event preference in all experiments. However, we did observe an earlier inspection of the future event target and a significant tense effect compared with the previous experiment that used a balanced number of recent and future events and that did

not pit other factors against the recent-event preference (Abashidze et al., 2011, Exp. 2).

Recall that the previous findings discussed in Chapter 2 provided evidence of language and visual context integration in syntactic ambiguity resolution. In line with previous findings and our expectations we pay close attention once again to the temporal cues in the present experiments of the past and future tense condition and to the importance of non-linguistic cues at the verb and temporal adverb 'flavoured recently'/'flavours soon'. It can be assumed that if one had recently inspected an action performed upon an object, that the representation of that object and action is somewhat active and present in our working memory. Thus, on hearing a sentence convey temporal information such as the 'The experimenter flavored recently ...' rapid access of the mental representation of the recently seen object should not be surprising. In addition to being supported by world knowledge (e.g., we flavor tomatoes) the process is supported by our visual short-term experience. However, not only short-term experience, but also visual and linguistic experience of statistical regularities (e.g., Abashidze et al., 2011, Exp., 2: Kaschak et al., 2006) and of other situational cues (Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012) as well as incongruence (Carpenter & Just, 1975; Knoeferle, Urbach, & Kutas, 2011) can modulate language comprehension (see Chapter 2).

Thus, with the present manipulations it is possible that participants could have anticipated the object of a plausible future event at an early stage of sentence processing when they listen to 'The experimenter flavors soon ...'. When the tense cue becomes available at the end of the verb and during the adverb region, participants should rapidly interpret the linguistic cues about the future and anticipate future event possibilities. Moreover, we saw already other findings that revealed faster processing of the ongoing (not completed) events, which in our case could have been the upcoming future event. Additionally, there are arguments as to why the inspection of the future event target could happen early during sentence processing. One argument in favour of why participants might inspect the future event target early is that they had already seen one object acted upon (in the recent action event). As a result of their experience, they might realise (over the course of the experiment) that the next action would often involve the other object (the scene presented only two objects). Furthermore, this experience might be heightened if they see a higher percentage of future events performed in the

experiment (e.g., Experiments 1 and 2) and experience a high percentage of sentences about future events. In support of our experimental manipulations, for instance, information established through repetition could generate a prediction, viz. that people can quickly utilize information from a small amount of examples (e.g., Britt, Mirman, Kornilov, & Magnuson, 2014). While statistical experience was a powerful cue in Experiments 1 and 2, the situation-immediate gaze cue could have given participants a good reason to anticipate the future target object early and perhaps even prefer to anticipate it overall (Experiments 3 and 4). In Experiment 5, finally, the sentences referring to the recent action event never matched the action such that participants had good reason to anticipate the future event target.

However, on the other hand, our alternative expectations stated were that we would observe a recent event influence on the comprehension of sentence about a plausible future event (see Chapter 2, also Knoeferle, 2015a, 2015b), which was confirmed by a significant overall recent-event preference. For instance, one of the first visual world paradigm experiments revealed that participants' eye-movements were modulated by different visual contexts (one vs two referent context) during language processing (Tanenhaus et al., 1995). Furthermore, other studies by Knoeferle et al., (2005, 2006) have revealed strong visual context (depicted action) effects in guiding participants' gaze pattern towards a character during language comprehension. Other results showed a strong effect of recently seen clipart events on language comprehension (Knoeferle & Crocker, 2007). The latter finding was further confirmed by other experiments that have used real-world action events (with real objects Abashidze et al., 2011).

Moreover, we speculate on the possible causes for the present project findings. For instance, the so-called 'anchoring hypothesis' has been discussed in a study by Tversky and Kahneman (1974). The anchoring hypothesis states that humans are biased toward a first experienced reference point i.e., an anchor, and that they adjust to that reference when they evaluate subsequent information (Tversky & Kahneman, 1974). On the one hand, this account could support findings of a preferential inspection of the first reference target until the name of the second plausible target is encountered (e.g., Knoeferle, Carminati, et al., 2011). On the other hand, this bias could change when the frequency, the actor's gaze cue or the incongruence are manipulated. Most importantly, the anchoring mechanism

should have changed when the incongruence of the recent event and the past tense verb did not support the creation of a reliable prior reference point.

Another possible reason for the observed robustness of the recent-event preference might be the well-known 'recency effect'. For a recency effect, the assumption would be that people are directly influenced by the recently-seen action, and that their delayed anticipation of an equally plausible future action target needs to compete with recent information in their working memory. Because of the strongly activated representation of the recent action and its target, the anticipation of the plausible future event target emerges very late. A recency account is compatible with the gaze preference for the recent action target as exhibited by participants during and after the verb region. Humans might prefer to pay more attention to a recently seen event target, because they have a recent experience of a particular target object or location of the object. For example, in a visual search task participants directed their eye-movements towards a target quickly when they had inspected it in the prior scene (e.g., Körner & Gilchrist, 2007).

Next we also speculated whether the later inspection of the future event target could be caused by our experimental materials. One reason for why participants looked at the recent event target more might have been a possible expectation of a second action on the same object. For example, when we flavored one object (e.g., a cucumber) we might next taste this object instead of flavoring another object (e.g., a tomato) as a second activity. In some cases it could be possible that we would want to try the first flavored object before we flavor the second object. Since some of our experimental trials have included action events on an object for which the possibility of a second action would have been high, we decided to compare the gaze pattern for these trials to trials for which we estimated the probability of a second action as low (e.g., opening an empty bottle). The post-hoc analyses did not show any reliable difference between the gaze pattern for these two sub-groups of the items (see graphs on the gaze data of the experimental items included in the Appendix D for Experiment 1).

Another possible cause of the recent-event preference behaviour might have been the fact that for the present experimental sentences in German, the future tense condition was marked with a present tense verb. Thus, during the verb there was an ambiguity between the recent and future target object. An ongoing research project (Abashidze & Chambers, 2016) has investigated this issue with English sentences (such as *The experimenter will/has sweeten/sweetened the strawberries/the*

pancakes), that used a similar sentence structure as in the current studies. Even though the tense cue occurred early in the sentence via auxiliary verbs (*has/will*) and the main verb, the study on English replicated our present project findings. During the auxiliary verb and the main verb participants mostly inspected the recent target objects in both tense conditions and they started to look at the future event target only during the NP2 region.

These present findings suggest that we might process a recently seen event longer and that it has a strong influence on language comprehension, even when listeners encounter conflicting tense information in the sentence. If this is the case then it is not clear to what extent the gaze pattern in the past tense condition is elicited by the recent event cues in the visual context and to which extent it is mediated by the past tense sentence. Another possibility for accommodating the increasing attention towards the recent event target in the verb region might be an invariant mechanism. Whatever a (subject-object-verb) sentence that refers to the recently perceived event target or to a future event target mentions, comprehenders prefer to ground verb information in the recent action event. Further research is needed to address this issue.

In this section we discussed the experimental findings and we compared the time course graphs of those findings between the experiments in order to specify which of the experimental manipulations was most effective against the recent-event preference. In addition, we discussed the memory test findings in relation to the gaze data. Furthermore, we pointed out that an earlier inspection of the future event target would have been plausible in light of our experimental manipulations, and noted the strength of the visual context during language comprehension. We further outlined possible causes for participants' gaze pattern in the present experimental findings. The next section will highlight the robustness of the recent-event preference.

6.1 Robustness of the Recent-Event Preference

Previous language processing research on events has reported a range of results. For example, if we recall, a word-by-word reading study by Madden and Therriault (2009) has found that participants processed a past imperfect sentence referring to an ongoing event faster than a past perfective sentence referring to a completed

event. On the other hand, eye-tracking findings provided evidence of a strong influence of recently perceived action events on language comprehension. When participants saw a recent event and then they listened to a sentence referring to the recent event or to a plausible future event, the effect of the recently perceived event emerged very early in the sentence (verb region) and lasted throughout sentence. Participants paid more attention to the recent event and its target independent of the sentence tense (Abashidze et al., 2011; Knoeferle & Crocker, 2007). Furthermore, a narration of an event affected participants' eye-movement in the later narrated location of a target object (Altmann & Kamide, 2009). Moreover, depicted action events modulated participants' eye-movements during language comprehension (Knoeferle et al., 2005).

At the outset of the present thesis project, we did not know much about the strength of the recent-event preference and its robustness during language processing, especially when the action events used real-world objects. In order to investigate the robustness of the recent-event preference we instantiated a frequency bias in favour of future action events; we furthermore introduced a situation-immediate cue (the actor's gaze) and pitted it against the recent-event preference, and the final manipulation concerned the incongruence between the recent event and the past tense verb. With these manipulations we expected an early decrease of the influence of the recent event and fewer eye-movements towards the recent event target compared to previous findings by Knoeferle, Carminati, et al. (2011, Exp. 2).

Experimental manipulations in other research, such as statistical regularities, the gaze of an interlocutor, and picture-sentence incongruence, have all revealed a rapid effect on language processing. Indeed, these factors in the present thesis did modulate the recent-event preference, but they did not override the preference early on (e.g., during the verb region). Thus, the post-hoc one-sample t-tests for the future tense condition in the verb region revealed a reliable effect of preferential looks to the recent event target in all experiments (except the gaze condition in Experiments 3 and 4, see Tables 4.4 and 4.8). Moreover, we replicated an overall recent-event preference in all experimental findings (which means that we found a reliable intercept with a positive value in all three word-regions of both sentence tense conditions, and this was confirmed by a significant intercept in analyses by participants and items in all experiments).

In summary, we now know that the overall recent-event preference can not be completely overridden by the factors that were applied in the thesis project. At the same time, our findings provide evidence that our experimental manipulations and biases were effective against the recent-event preference. These effects were mainly shown in the later word-regions by the post-hoc *t*-test analyses (of which more below). Thus, the recent-event preference is not invariant and key findings on its sensitivity to context will be summarised in the following sections.

6.2 Sensitivity of the Recent-Event Preference

Although all thesis experiments found that the recent-event preference is robust and cannot be easily overridden by our experimental manipulations, its robustness has clear limitations. Furthermore, we should note that its sensitivity differed between the experimental manipulations. For instance, in the previous experiment by Abashidze et al. (2011, experiment 2, 50/50%) the preferential inspection of the recent event target lasted until the middle of the NP2 region. By contrast, in the thesis experiments the inspection bias towards the recent event target decreased early. The inspection of the future target in the future tense sentence started in the adverb region in the present experiments 1 to 4 (except the gaze condition), which is approximately 1000-1200 ms earlier. Moreover, in the future gaze condition the reliable gaze pattern towards the future event target occurred even early at the beginning of the adverb region. Comparatively, less sensitivity was found by the incongruence in the last experiment. In the next sections we closely look at the sensitivity of the recent-event preference in response to our experimental manipulations.

6.2.1 Frequency Distribution

The effects of frequency bias in Experiments 1 and 2 revealed that the recent-event preference is sensitive to the statistical regularities. When participants have more often observed future events and heard sentences in the future present, their preferential looks towards the recent event target has decreased compared with the effects reported in previous experiment (Abashidze et al., 2011, experiment 2). The sensitivity in the gaze data can be seen in the time course graphs (see Figures 2.5 and 3.3).

In addition to descriptive analyses, the statistical analyses of the sentence tense revealed significantly less inspection of the recent event target in the future tense condition than in the past tense condition. Moreover, the one-sample *t*-test for the adverb region in the future tense condition did not reveal a reliable effect for the recent event targets (except by items in Experiment 2, Table 3.8), which means that the frequency manipulation eliminated the recent-event preference effect found in Experiment 2 by Abashidze et al. (2011s). The *t*-test revealed a significant effect in the NP2 region (a reversed preferential inspection) by subject and by items. The results of the *t*-tests suggest that participants in the future tense condition inspected the future event target significantly more than the recent event target during the second noun phase in both experiments.

The recent-event preference was not only sensitive to the frequency manipulations but also to the situational gaze cue. In the next section we will outline key findings concerning the effects of an actor's gaze as one situation-specific object-directed cue.

6.2.2 Actor Gaze Cue

Experiments 3 and 4 showed further sensitivity of the recent-event preference, which was caused by pitting a situational actor's gaze against it. This manipulation revealed a comparable early effect against the recent-event preference, which was observed in the reduced eye-movements towards the recent event target early in the sentence. In the future gaze condition, participants shifted their visual attention to the future target object after a few hundred milliseconds of the actor's gaze cue onset. This means that the actor's gaze cue was strong enough to reduce the bias towards the recent event early in the verb region. Experiment 4 which featured an earlier gaze cue onset revealed also earlier effects of the gaze cue on the recent-event preference than Experiment 3. However in both experiments the future target inspections in the gaze condition appeared approximately 1800-2000 ms earlier than in a previous experiment (Abashidze et al., 2011, Exp. 2, 50/50%)

Furthermore, the *t*-tests results for the future tense gaze condition revealed an effect of preferential looks towards the future event target during the adverb in the analyses by participants and by items. The results of the *t*-test suggest that participants during the adverb inspected the future event target significantly more

than the recent event target when they had heard an adverb indicating the future and when the actor's gaze cued the future event target.

The *t*-test results during the NP2 region showed a reliable effect in both gaze and no-gaze future tense conditions. These results suggest that when participants listened to the second noun phrase, they inspected the future event target more than the recent event target in the future tense. Thus, the effectiveness of the gaze cue manipulation against the recent-event preference was confirmed by the statistical analyses in both experiments.

6.2.3 Verb-Action Incongruence

The sensitivity of the recent-event preference was found not only in Experiments 1 to 4 but also in Experiment 5 that changed the fit of the recent event and the past tense verb such that action and verb always mismatched. When the recent event mismatched the past tense verb, inspections towards the recent event target decreased. As an effect of incongruence when participants listened to sentences in the future tense condition they reduced their looks towards the recent event target starting from the middle of the verb region. Furthermore the inspections of the recent event target decreased not only in the future but also in the past tense conditions. This was confirmed by a non significant tense effect in the adverb region. Furthermore, the preferential inspections of the future event target in the future tense emerged at the beginning of the NP2 region, which is several hundreds ms earlier than in the previous experiment (Abashidze et al., 2011, experiment 2 with 50/50%).

The post-hoc *t*-test in the future tense condition failed to find a reliable effect of an inspection preference for the future event target. Even though the time course graph in the last noun phrase region showed a clear bias of participants looking at the future event target, this was not confirmed by statistical analyses. This suggests that the incongruence manipulation had a relatively weak effect against the recent-event preference. In sum, we can argue that there is a strong bias in favour of the recent event during language processing; however this preference is sensitive to the modulating influence of at least the three factors that we examined (i.e., within-experiment frequency biases, situation specific gaze cues, and action-verb incongruence).

6.3 Conclusion

In the above sections we have discussed our findings on the robustness of the recent-event preference, and its sensitivity to the experimental manipulations. In what follows, we highlight previous research on visual context effects and anticipation during language processing and discuss the reported findings in the context of existing research results.

Researchers found that while participants listened to a sentence about a plausible target in the scene they could anticipate the target before it was mentioned (e.g., Altmann & Kamide, 1999, 2007; Kamide, Scheepers, & Altmann, 2003). Furthermore, other authors stated that "... listeners use a wealth of linguistic as well as visual information to disambiguate different sentence structures and to predict the upcoming linguistic input. Language-mediated eye movements reflect neither only linguistic processing nor only processing of the visual scene, but they reflect continuously updated mental representations based on information derived from both the linguistic and the visual input" (Huettig et al., 2011, p. 156.). The authors of the above articles and the later statement (of " ... a wealth of linguistic as well as visual information ...") potentially underestimate the importance of visual context influences on language comprehension (see also Altmann & Mirković, 2009).

Indeed, other findings (e.g., Abashidze et al., 2011; Knoeferle & Crocker, 2007; Knoeferle, Carminati, et al., 2011) and the results of the present projects clearly indicate that a recently seen action event can have a strong effect on language comprehension and on participants' anticipation of a plausible future event target. Thus, when participants saw a recent event and future event equally often and similarly they listened to a sentence about a recent or future event, their eye-movements towards the target object of the future event occurred only at the end of the target name presentation (Abashidze et al., 2011, Exp. 2). By contrast, looks to the target of the recent event occurred earlier, as the verb referenced the action. In addition, the immediate action effect can be seen in the time course graphs in all of the thesis experiments (the increasing gaze pattern towards the target of the recent event from the verb onset in both past and future tense conditions). These results suggest that understanding the verb is guided by the immediate visual context and this despite the strong biases in favor of the future event. By contrast to previous findings, our project experiments did not reveal a clear prediction of the upcoming linguistic information at the end of the verb region

for future tense sentences (such predictive inspection occurred 700 ms after the verb offset, considering only the no-gaze conditions). Our findings support the view that the visual context exerts a strong influence during language comprehension and suggest that the prediction of future event targets does not always occur as rapidly, even when cues about event time support it (see also Knoeferle, Carminati, et al., 2011). Instead, experiencing, for example, action events as is the case in our everyday interaction, may exert a stronger role in guiding our visual attention and ongoing language comprehension.

Previous research on the influence of a recently seen event on language comprehension is scarce. In this thesis project we presented five experiments that stress-tested the recent-event preference (how a recently seen event influences language processing about a plausible future event). The results of the present experiments support and extend previous findings of visual-context effects on incremental language comprehension. First, two experiments established a strong frequency bias against the recent-event preference with the goal of inducing early inspections of the future event target. In particular, at the end of the verb the bias towards the recent event target began to decrease and we observed anticipatory eye movements towards the future event target in the adverb region. In addition, frequency cues in the experiment increased the likelihood that the plausible future event target would soon be mentioned. The results suggested that the frequency bias was effective, leading to an earlier increase of looks to the future action target in sentences about future events. However, at the same time the preferred inspection of the recent even target persisted. The preferential inspections were supported by the post-experimental memory test (participants were better able to recall the recent action events than the future action events)

A further two experiments investigated the issue of the recent-event preference by pitting another situation cue against it. When combined with the present tense verb form and the future adverb, the situation cue (the actor's gaze) elicited an early and strong effect against the recent-event preference. However, the overall recent-event preference persisted. Furthermore, the gaze cue effect was not found in the memory test. By contrast, participants had higher accuracy in recognising the action events and recalling sentences in the no-gaze than gaze condition.

The last experiment examined the incongruence between the recent event and the language input. We identified an effect of the incongruence beginning at the end of the verb; however the effect was not strong enough to elicit an early gaze

preference towards the future event target. This experiment also replicated the overall recent-event preference. By contrast the memory bias changed such that participants were better in recalling the future tense sentences.

In summary, our findings clearly show how an immediate non-linguistic cue (the recent event) influences participants' language comprehension. In all experiments we replicated the preferential looks toward the recent event target. When participants heard the verb they looked at the recent event target irrespective of the sentence tense. Interestingly at the onset of the verb (in the past and future sentence condition) participants incrementally increased their inspections of the recent event target. The preferential looks towards the recent event target lasted until the middle of the adverb region in Experiments 1 to 4 (considering only the no-gaze conditions) and until the second noun onset in Experiment 5. We suggest that in a visual context, when we have an experience of an action event and then listen to the verb in a sentence, we might integrate the verb with the recent event. Furthermore, the memory test results support the findings of the eye-movement data in experiments that applied a frequency distribution with a better memory of the recent events. However, the gaze cue included in the experiments did not benefit later memory but only had an immediate effect on the listeners' visual attention to the future target. The memory test in the last experiment revealed better recall in the future tense than the past tense sentence condition.

Even though we replicated an overall recent-event preference in all experiments, our experimental manipulations did affect the recent-event preference. We found an approximately 1000 ms earlier inspection of the future event target (Experiments 1 to 4, considering only the no-gaze conditions) compared with earlier studies (e.g., Knoeferle et al., 2011, Exp. 2). For the gaze conditions, looks to the future event target emerged even earlier. The action-verb incongruence in Experiment 5 did not show as strong an effect as the other experiments (the future target inspection started at the beginning of the NP2 onset). Overall, the key findings of the thesis are the robustness of the recent-event preference during language comprehension; however, at the same time that preference shows sensitivity to various other factors such as the frequency of future events and sentences, the immediate gaze cue of an actor, or action-verb incongruence.

Furthermore, in line with these findings we suggest that the recent-event preference might reflect an epistemic bias of the human mind. The recent target inspection is based on stronger evidence of truth (of an event in the world), and while the recent

event during the past tense sentence can be verified, the same can not happen for the future event - at least not until it has really occurred. During the future tense sentence, the realisation of the future event remains uncertain (see MacFarlane, 2003; Staub & Clifton Jr, 2011).

In summary, future research efforts are obviously necessary to understand the exact nature of the mechanism underlying the recent-event preference. We argue that when participants have experienced a recent action event, their attention towards the target of the recent event increased during the verb independent of the sentence tense. The increased eye-moments towards the recent target has occurred in all experiments (see the time course graphs). Moreover the pattern has been replicated in ongoing research when the tense of the sentence was conveyed by an auxiliary verb preceding the matrix verb (*will/has*). In order to examine whether the recent event bias develops during the verb one might examine this issue in a language in which the tense cue unfolds before the verb e.g., at the first noun phrase. It would be interesting to find out whether the recent-event preference might already begin to increase during the first noun phrase and if this is the case, how long this preference would persist even when the future tense cue occurs before the lexical verb (referring to the action)? Another possible way to examine the recent-event preference further is to create new materials. This could involve showing a first action event for which it is clear that the action has been completed (e.g., an empty glass after the actor drank the juice). A change in stimuli would permit us to examine whether showing one event as completed might increase participants' expectations that another event would happen on another object. This should increase the probability of an expectation that a second action event will be performed on the other object.

6.4 German Summary

Die vorliegende Doktorarbeit beschäftigt sich mit visuellen Ereignissen und Sprachverstehen. Wir haben den Einfluss dieses Ereignisses auf das Sprachverstehen erforscht. In unserer täglichen Interaktion benutzen wir sprachliche und nichtsprachliche Mittel. Jedoch wissen wir wenig über den möglichen gegenseitigen Einfluss visueller Informationen (z.B., eines unmittelbaren Ereignisses) auf das Sprachverstehen. Um die Forschungsfragen der Doktorarbeit zu motivieren und diese anschliessend zu beantworten, haben wir in erster Linie die relevanten früheren Ergebnisse im situierten Sprachverstehen detailliert diskutiert (siehe Kapitel 2).

Frühere Experimente haben bewiesen, dass nichtsprachliche Informationen wie z.B. visuelle Kontexte, dargestellte Ereignisse und die Form eines Objekts unmittelbar mit dem Sprachverstehen interagieren (z.B., Dahan & Tanenhaus, 2005; Knoeferle & Crocker, 2007; Tanenhaus et al., 1995). Ferner hat ein Experiment von Madden und Therriault (2009) das Sprachverstehen während einer laufenden oder abgeschlossenen Handlung untersucht. Die Autoren verwendeten eine Wort-für-Wort Leseaufgabe bei der die Teilnehmer Sätze im Perfekt und im Imperfekt lasen. Die Teilnehmer haben die Sätze im Imperfekt schneller gelesen, wenn sie während des Satzes ein Bild einer laufenden Handlung (z.B., einen aufgeklappten Laptop) sahen als wenn sie einen Satz im Perfekt mit einem Bild einer beendeten Handlung (z.B., einen geschlossenen Laptop) sahen. Dieser Befund deutet auf eine schnellere Verarbeitung des Satzes in Bezug auf eine laufende Handlung gegenüber einer abgeschlossenen Handlung hin.

Des Weiteren wurden in einer anderen Reihe von Eye-Tracking Experimenten antizipatorische Augenbewegung bei der Verarbeitung eines Satzes gezeigt. Wenn Versuchspersonen eine Szene mit mehreren Objekten auf einem Bildschirm sahen und dazu einen Satz hörten der sich auf ein bestimmtes Objekt bezog, haben sie das Objekt antizipiert und den Blick auf das Objekt gerichtet noch bevor der zugehörige Name gehört wurde (z.B., Altmann & Kamide, 1999, 2007). Die Versuchspersonen nutzten dabei das Verb und den visuellen Kontext, bzw. ihr Weltwissen, um mögliche Geschehnisse in der Szene vorherzusagen.

Im Gegensatz zu den oben beschriebenen Befunden zeigte ein Experiment von Knoeferle und Crocker (2007, Experiment 3) keine antizipatorische Augenbewegung während des Satzes. In diesem Experiment sahen Versuchspersonen eine

Interaktion mit einem Objekt in einer Szene. Danach hörten sie einen Satz, entweder im Präteritum oder im futurischen Präsens. Der Satz in der Vergangenheitsform bezog sich auf die unmittelbare Interaktion mit dem Objekt. Der Satz in der Zukunftsform bezog sich auf die mögliche Interaktion mit einem anderen Objekt. In beiden Satzformen wurde gezeigt, dass vorzugsweise das Objekt betrachtet wurde mit dem vorher in der Szene interagiert wurde auch dann, wenn das Gehörte (der Satz in der Zukunftsform) nicht mit dem Gesehenen übereinstimmte. Ein Grund für die Abwesenheit einer antizipatorischen Augenbewegung zum Objekt der möglichen, zukünftigen Handlung könnte sein, dass dieses Experiment keine zukünftige Handlung nach dem Satz zeigte.

Aus diesem Grund hat ein weiteres Experiment von Abashidze et al. (2011, Experiment 2) diese Fragestellung untersucht. Bei dem Experiment wurde eine ähnliche Methode wie bei Knoeferle und Crocker (2007) verwendet. Dabei wurden beide Handlungen (eine vor dem Satz und eine zweite Handlung nach dem Satz) präsentiert. Trotz der Handlungen vor und nach dem Satz zeigten die Ergebnissen ähnliche Blickmuster wie in dem Experiment von Knoeferle und Crocker (2007). Das Objekt, das bei der zuvor gesehen Handlung in Fokus stand, wurde in beiden Satzformen bevorzugt angeschaut. Bei der Zukunftsform wurde keine antizipatorische Blickbewegung gezeigt. Der Blick wurde nur während der Benennung auf das Objekt gelenkt. Dieses Phänommen, dass der Blick vorzugsweise auf demjenigen Objekt, welches im Fokus der vorherigen Interaktion stand, blieb, nannten wir "recent event preference" (REP).

Infolgedessen konzentriert sich meine Doktorarbeit auf diesen Aspekt, nämlich den kürzlichen visuellen Effekt in der Sprachverarbeitung, der im Zusammenhang mit der Forschung in der Psycholinguistik bisher keine ausreichende Aufmerksamkeit gefunden hat. Darüber hinaus soll diese Doktorarbeit helfen, zu erklären wie kognitive Mechanismen in Verbindung mit in den Experimenten verwendeten Manipulationen und dem Sprachverstehen interagieren, wenn sich die Sprache auf eine zukünftige visuelle Aktivität bezieht. Erstens haben wir erforscht, ob die Häufigkeit der Handlung und die Zeitform des Satzes die sogenannte REP beeinflusst. Zweitens haben wir die REP mit einer weiteren Manipulation überprüft, bei der der Versuchsleiter mit seinem Blick auf das genannte oder zu nennende Objekt hinweist. Drittens haben wir die Inkongruenz zwischen der vorherigen Handlung und dem Verb im Präteritum untersucht. Ausserdem haben wir jeweils nach jedem Eye-Tracking Experiment einen Gedächtnistest eingesetzt, um zusätzliche

Informationen über die Befunde zur Augenbewegung zu gewinnen.

In den Kapiteln 3 bis 5 päsentieren wir die fünf Eye-Tracking-Experimente, in denen eine ähnliche Methode wie in den Experimenten von Abashidze et al., (2011) und Knoeferle und Crocker (2007) angewendet wurde. Die Versuchspersonen sahen ein kurzes Video von 5 Sekunden in dem eine Person (als ein zweiter Versuchsleiter) eine Handlung durchführt (z.B., Gurken salzen). Danach sahen sie ein statisches Bild in dem die Person am Tisch sass und zwei Objekte auf dem Tisch lagen. Währenddessen hörten die Versuchspersonen einen Satz entweder im Präteritum (z.B., *Der Versuchsleiter würzte kürzlich die Gurken*), der sich auf die vorherige Handlung bezog oder einen Satz im futurischen Präsens (z.B., *Der Versuchsleiter würzt demnächst die Tomaten*), der sich auf eine andere zukünftige Handlung bezog. Nach dem Hören des Satzes sahen sie ein zweites Video, in dem die Person eine zweite Handlung (z.B., Tomaten salzen) durchführte.

Als eine erste Bedingung haben wir die Häufigkeit des Ereignisses und der Satzform verändert. Frühere Befunde über Frequenzveränderungen haben gezeigt, dass die Frequenz in den Bereichen der Sprachverarbeitung und der visuellen Wahrnehmung eine wichtige Rolle spielt. In den Experimenten 1 und 2 haben wir die Anzahl der Sätze mit zukünftigen Ereignissen erhöht und dementsprechend die Sätze in der Vergangenheitsform mit den dazugehörigen Handlungen reduziert. Der Anteil der Handlungen, die mit einer Zukunftsform beschrieben wurden, betrug im ersten Experiment 75% und im zweiten Experiment 88%. Die Ergebnisse der Experimente zeigten, dass die Manipulation der Häufigkeit einen Effekt auf die REP hatte. Die antizipatorischen Augenbewegungen traten beim Hören des Satzes in der Zukunftsform beim Adverb auf. Im Vergleich zu dem Experiment von Abashidze et al. (2011, Experiment 2) haben diese aktuellen Experimente gezeigt, dass die Blickbewegung auf das Objekt der möglichen zukünftigen Handlung circa 1000-1200 Millisekunden früher geschah.

Obwohl wir einen Effekt durch die Häufigkeit der Manipulation gezeigt haben, hat sich der Einfluss der zuvor gesehenen Handlung weiter in den aktuellen Experimenten bestätigt. Die bevorzugte Blickrichtung auf das Objekt der unmittelbar zuvor gesehenen Handlung wurde von den Ergebnissen aus dem Gedächtnistest bestätigt. Die Versuchspersonen konnten sich besser an die Sequenzen erinnern, die vor dem Satz gezeigt wurden als nach dem Satz. Die Aufgabe in den Gedächtnistests war, sich an die Handlungen zu erinnern, die entweder vor dem Hören

des Satzes oder nach dem Hören des Satzes in den Eye-Tracking Experimenten durchgeführt wurden.

Eine weitere Manipulation wurde in den Experimenten 3 und 4 angewendet, in denen das Blicksignal des Handelnden im Video gegen die REP eingeführt wurde. Frühere Experimente, in denen eine Person mit einem Blick auf das Objekt in der Interaktions-Szene deutete, haben einen starken Einfluss dieses Blicks auf die Verarbeitung der Sprache gezeigt (z.B., Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012). Das Blicksignal trat in Experiment 3 während des Verbs und in Experiment 4 am Anfang des Verbs auf. Das Blicksignal richtete sich immer auf das Objekt, welches in dem Satz mit der zweiten Nominalphrase benannt wurde. In beiden Experimenten verweilte das Blicksignal bis zum Ende der letzten Nominalphrase auf deren Referenten.

Diese Manipulation hat eine schnelle Unterstützung in der antizipatorischen Augenbewegung bei der Sprachverarbeitung gezeigt. Wenn die Versuchspersonen das Verb im Satz in der Zukunftsform hörten und dazu das Blicksignal wahrnahmen, begannen sie einige hundert Millisekunden später, das Objekt der zukünftigen Handlung zu antizipierten. In beiden Experimenten betrachten die Versuchspersonen das Objekt im Satz in der Zukunftsform vom Ende des Verbs an. Wenn der Versuchsleiter hingegen kein Blicksignal gab, traten die antizipatorischen Augenbewegungen zum zukünftigen Objekt erst zur Mitte des Adverbs hinauf. Insgesamt hatten die Manipulationen einen starken Einfluss auf die REP. Jedoch, wie auch Experimente 1 und 2, zeigten die aktuellen Experimente den bevorzugten Blick der Probanden auf das Objekt aus der vorherigen Handlung. Darüber hinaus zeigten die Ergebnisse aus den Gedächtnistests einen besseren Abruf von Handlungen und Satzformen in der Bedingung in der keine Blicksignale des Handelnden vorhanden waren.

Nachdem die Experimente 1 bis 4 die überwiegende Blickrichtung zum Objekt der kürzlich gesehenen Handlung gezeigt haben, hat das Experiment 5 eine andere Methode, nämlich eine Inkongruenz gegenüber der REP bedient. Dieses letzte Experiment der Dissertation hat insbesondere die vermehrte Blickrichtung zum Objekt des unmittelbaren Ereignisses während des Verbs untersucht. In diesem Experiment stimmten die kürzlich betrachtete Handlung und das Verb im Präteritum nicht überein. Dahingegen stimmte der Satz in der Zukunftsform mit der zukünftigen Handlung immer überein. Die Ergebnisse haben gezeigt, dass die

Blickbewegung zum Objekt der vorherigen Handlung am Ende des Verbs in beiden Sätzen weniger wurden. Jedoch erschienen die meisten Blickbewegungen zum Objekt der vorherigen Handlung bis zum Anfang der zweiten Nominalphrase. Das Objekt der zukünftigen Handlung wurde nur während der zweiten Nominalphrase häufiger in der zukünftigen Zeitform als in der Vergangenheitsform des Satzes angeschaut. Wie in den Experimenten 1 bis 4 hat dieses Experiment einen allgemeinen, bevorzugten Blick zum Objekt der unmittelbaren Handlung bestätigt. Allerdings zeigten die Ergebnisse des Gedächtnistests ein anderes Muster als die Gedächtnistests in den Experimenten 1 bis 4 und zwar erinnerten sich die Versuchspersonen besser an die Sätze mit der Kongruenzbedingung (der Zukunftsform des Satzes), als an die Sätze in der Vergangenheitsform (Inkongruenz). Insgesamt zeigte die Manipulation der Inkongruenz einen stärkeren Effekt auf den später durchgeführten Gedächtnistest als auf die Augenbewegung während der Sprachverarbeitung.

Zusammenfassend lagen unsere Befunde nahe, dass die REP nicht einfach eliminiert werden kann. Trotz der starken Manipulationen gegen die REP haben alle Experimente eine bevorzugte Aufmerksamkeit auf die vorherige Handlung und ihr Objekt gezeigt. Auf der anderen Seite wurde mithilfe der experimentellen Manipulationen eine schnellere /frühere Blickbewegung zum Objekt der zukünftigen Handlung als in dem vorherigen Experiment (Experiment 2 Knoeferle, Carminati, et al., 2011) gezeigt.

In dem letzten Kapitel der Dissertation diskutieren wir unsere Literaturrecherche noch einmal. Weiterhin präsentieren wir unsere Ergebnisse und diskutieren inwieweit unsere Fragestellungen beantwortet wurden. Ferner vergleichen wir die Befunde zwischen den Experimenten und spekulieren über mögliche Gründe für die Ergebnisse in Bezug auf den starken Einflusses des visuellen Kontextes. Im Anschluss schlagen wir weitere Modelle für Experimente vor, mit deren Hilfe das untersuchte Phänomen weiter erforscht werden kann.

Appendix A

Item Sentences

All item sentences for Experiments 1 to 5 are listed below. Sentences (1a) and (1b) are from the future tense sentence condition (for both target objects). (2a) and (2b) are from the past tense sentence condition (for both target objects). The snapshots of the events are listed below (see Appendix B and C for the snapshots).

Item 1

(1a) Der Versuchsleiter zuckert sogleich die Erdbeeren.

(1b) Der Versuchsleiter zuckert sogleich die Pfannkuchen.

(2a) Der Versuchsleiter zuckerte soeben die Erdbeeren.

(2b) Der Versuchsleiter zuckerte soeben die Pfannkuchen.

Item 2

(1a) Der Versuchsleiter mixt sogleich den Milchshake.

(1b) Der Versuchsleiter mixt sogleich den Cocktail.

(2a) Der Versuchsleiter mixte soeben den Milchshake.

(2b) Der Versuchsleiter mixte soeben den Cocktail.

Item 3

(1a) Der Versuchsleiter buttert sogleich die Croissants.

(1b) Der Versuchsleiter buttert sogleich die Brotscheiben.

(2a) Der Versuchsleiter butterte soeben die Croissants.

(2b) Der Versuchsleiter butterte soeben die Brotscheiben.

Item 4

(1a) Der Versuchsleiter bewässert nachher die Kresse.

(1b) Der Versuchsleiter bewässert nachher die Tulpe.

(2a) Der Versuchsleiter bewässerte unlängst die Kresse.

(2b) Der Versuchsleiter bewässerte unlängst die Tulpe.

Item 5

(1a) Der Versuchsleiter poliert nachher die Kerzenständer.

(1b) Der Versuchsleiter poliert nachher die Wassergläser.

(2a) Der Versuchsleiter polierte unlängst die Kerzenständer.

(2b) Der Versuchsleiter polierte unlängst die Wassergläser.

Item 6

(1a) Der Versuchsleiter studiert nachher den Beipackzettel.

(1b) Der Versuchsleiter studiert nachher den Buchtitel.

(2a) Der Versuchsleiter studierte unlängst den Beipackzettel.

(2b) Der Versuchsleiter studierte unlängst den Buchtitel.

Item 7

(1a) Der Versuchsleiter öffnet demnächst die Saftflasche.

(1b) Der Versuchsleiter öffnet demnächst die Schuhkiste.

(2a) Der Versuchsleiter öffnete kürzlich die Saftflasche.

(2b) Der Versuchsleiter öffnete kürzlich die Schuhkiste.

Item 8

(1a) Der Versuchsleiter würzt demnächst die Gurke.

(1b) Der Versuchsleiter würzt demnächst die Tomate.

(2a) Der Versuchsleiter würzte kürzlich die Gurke.

(2b) Der Versuchsleiter würzte kürzlich die Tomate.

Item 9

(1a) Der Versuchsleiter salzt demnächst die Aubergine.

(1b) Der Versuchsleiter salzt demnächst die Zucchini.

(2a) Der Versuchsleiter salzte kürzlich die Aubergine.

(2b) Der Versuchsleiter salzte kürzlich die Zucchini.

Item 10

(1a) Der Versuchsleiter schlürft baldigst die Apfelschorle.

(1b) Der Versuchsleiter schlürft baldigst die Limonade.

(2a) Der Versuchsleiter schlürfte vorhin die Apfelschorle.

(2b) Der Versuchsleiter schlürfte vorhin die Limonade.

Item 11

(1a) Der Versuchsleiter schüttelt baldigst die Sojamilch.

(1b) Der Versuchsleiter schüttelt baldigst die Sprühsahne.

(2a) Der Versuchsleiter schüttelte vorhin die Sojamilch.

(2b) Der Versuchsleiter schüttelte vorhin die Sprühsahne.

Item 12

1.(a) Der Versuchsleiter verrührt baldigst den Kräutertee.

(b) Der Versuchsleiter verrührt baldigst den Milchkaffee.

(2a) Der Versuchsleiter verrührte vorhin den Kräutertee.

(2b) Der Versuchsleiter verrührte vorhin den Milchkaffee.

Item 13

(1a) Der Versuchsleiter pudert sogleich den Gugelhupf.

(1b) Der Versuchsleiter pudert sogleich den Obstkuchen.

(2a) Der Versuchsleiter puderte soeben den Gugelhupf.

(2b) Der Versuchsleiter puderte soeben den Obstkuchen.

Item 14

(1a) Der Versuchsleiter kostet sogleich den Whisky.

(1b) Der Versuchsleiter kostet sogleich den Rotwein.

(2a) Der Versuchsleiter kostete soeben den Whisky.

(2b) Der Versuchsleiter kostete soeben den Rotwein.

Item 15

(1a) Der Versuchsleiter besprüht sogleich die Pflanze.

(1b) Der Versuchsleiter besprüht sogleich die Blume.

(2a) Der Versuchsleiter besprühte soeben die Pflanze.

(2b) Der Versuchsleiter besprühte soeben die Blume.

Item 16

(1a) Der Versuchsleiter zählt nachher die Münze.

(1b) Der Versuchsleiter zählt nachher die Banknote.

(2a) Der Versuchsleiter zählte unlängst die Münze.

(2b) Der Versuchsleiter zählte unlängst die Banknote.

Item 17

(1a) Der Versuchsleiter schärft nachher die Messerklinge.

(1b) Der Versuchsleiter schärft nachher die Küchenschere.

(2a) Der Versuchsleiter schärfte unlängst die Messerklinge.

(2b) Der Versuchsleiter schärfte unlängst die Küchenschere.

Item 18

(1a) Der Versuchsleiter probiert nachher die Traubensaft.

(1b) Der Versuchsleiter probiert nachher die Orangensaft.

(2a) Der Versuchsleiter probierte unlängst die Traubensaft.

(2b) Der Versuchsleiter probierte unlängst die Orangensaft.

Item 19

(1a) Der Versuchsleiter pfeffert demnächst die Pizza.

(1b) Der Versuchsleiter pfeffert demnächst die Pasta.

(2a) Der Versuchsleiter pfefferte kürzlich die Pizza.

(2b) Der Versuchsleiter pfefferte kürzlich die Pasta.

Item 20

(1a) Der Versuchsleiter betrachtet demnächst die Zeitschrift.

(1b) Der Versuchsleiter betrachtet demnächst die Anzeige.

(2a) Der Versuchsleiter betrachtete kürzlich die Zeitschrift.

(2b) Der Versuchsleiter betrachtete kürzlich die Anzeige.

Item 21

(1a) Der Versuchsleiter säubert demnächst die Zitrone.

(1b) Der Versuchsleiter säubert demnächst die Birne.

(2a) Der Versuchsleiter säuberte kürzlich die Zitrone.

(2b) Der Versuchsleiter säuberte kürzlich die Birne.

Item 22

(1a) Der Versuchsleiter mischt baldigst die Farbe.

(1b) Der Versuchsleiter mischt baldigst die Karte.

(2a) Der Versuchsleiter mischte vorhin die Farbe.

(2b) Der Versuchsleiter mischte vorhin die Karte.

Item 23

(1a) Der Versuchsleiter schrubbt baldigst die Möhre.

(1b) Der Versuchsleiter schrubbt baldigst die Rübe.

(2a) Der Versuchsleiter schrubbte vorhin die Möhre.

(2b) Der Versuchsleiter schrubbte vorhin die Rübe.

Item 24

(1a) Der Versuchsleiter knetet baldigst den Teig.

(1b) Der Versuchsleiter knetet baldigst den Ton.

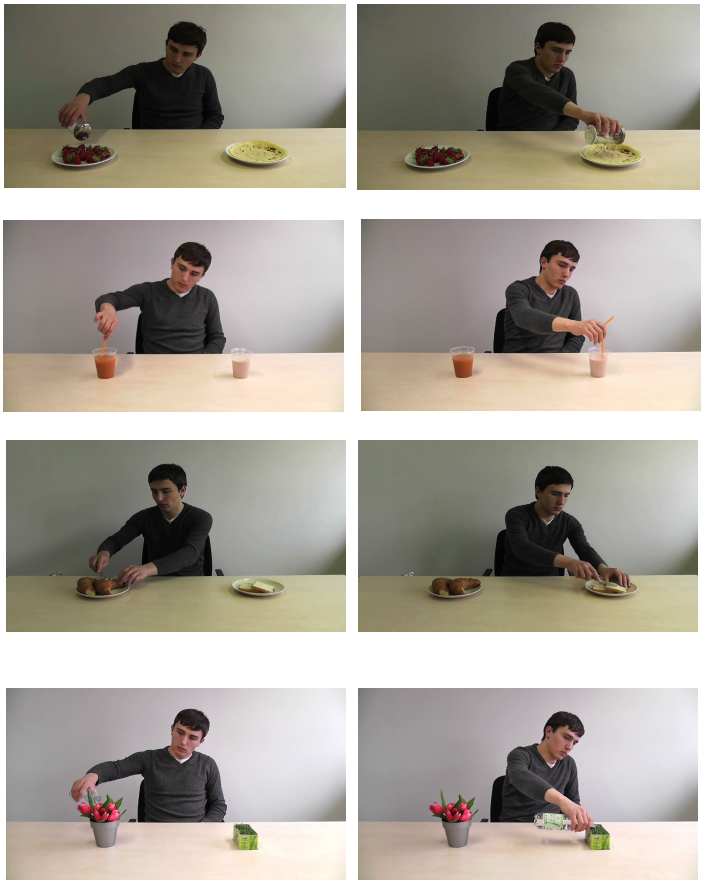
(2a) Der Versuchsleiter knetete vorhin den Teig.

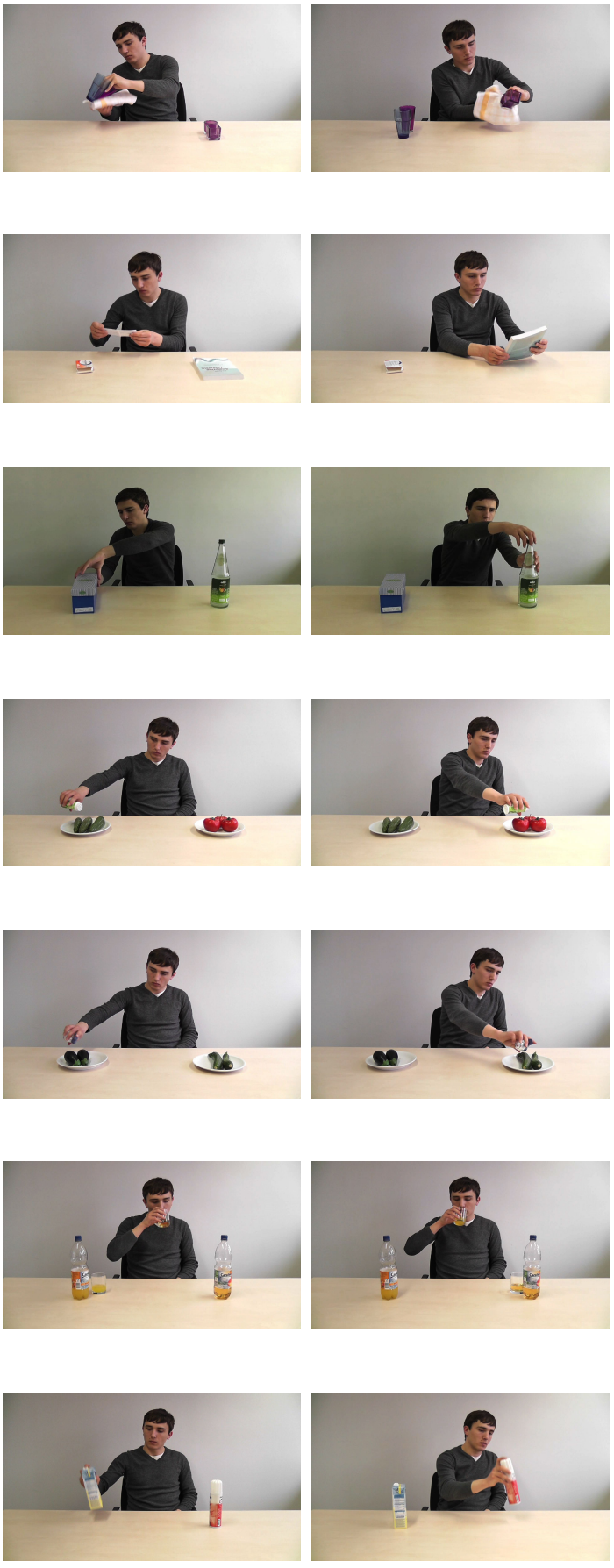
(2b) Der Versuchsleiter knetete vorhin den Ton.

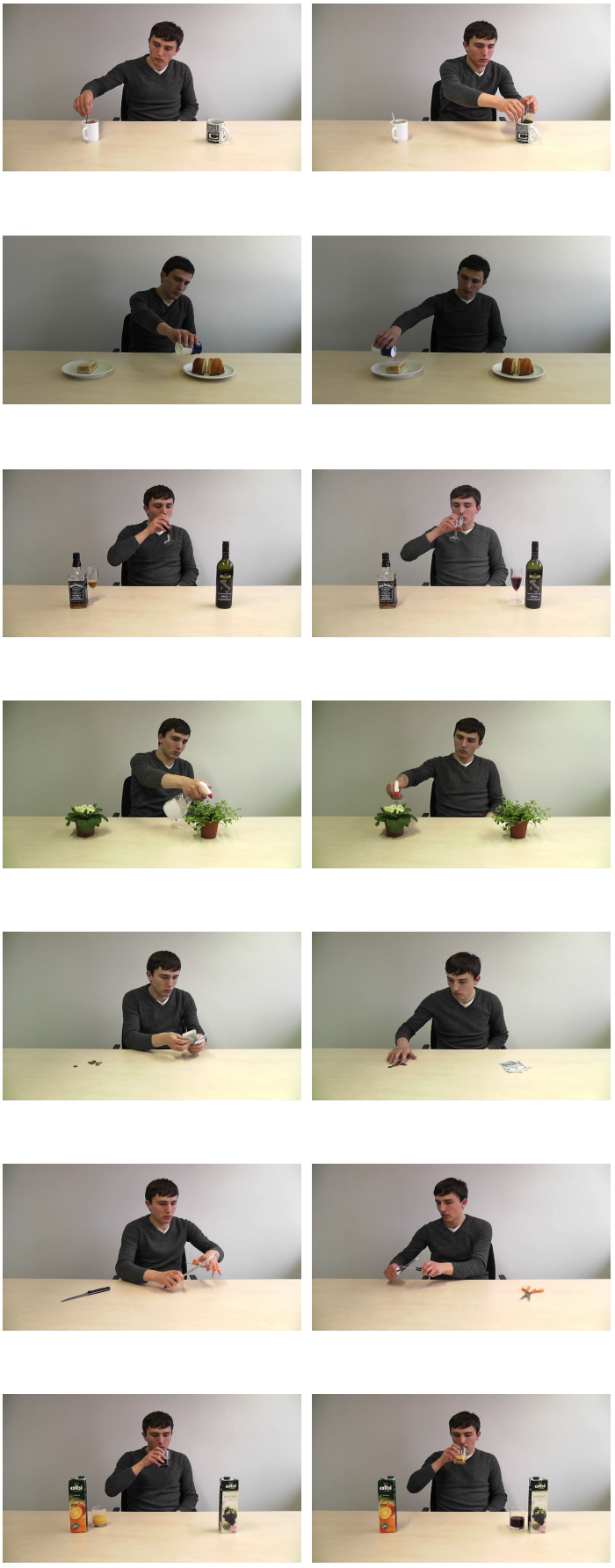
Appendix B

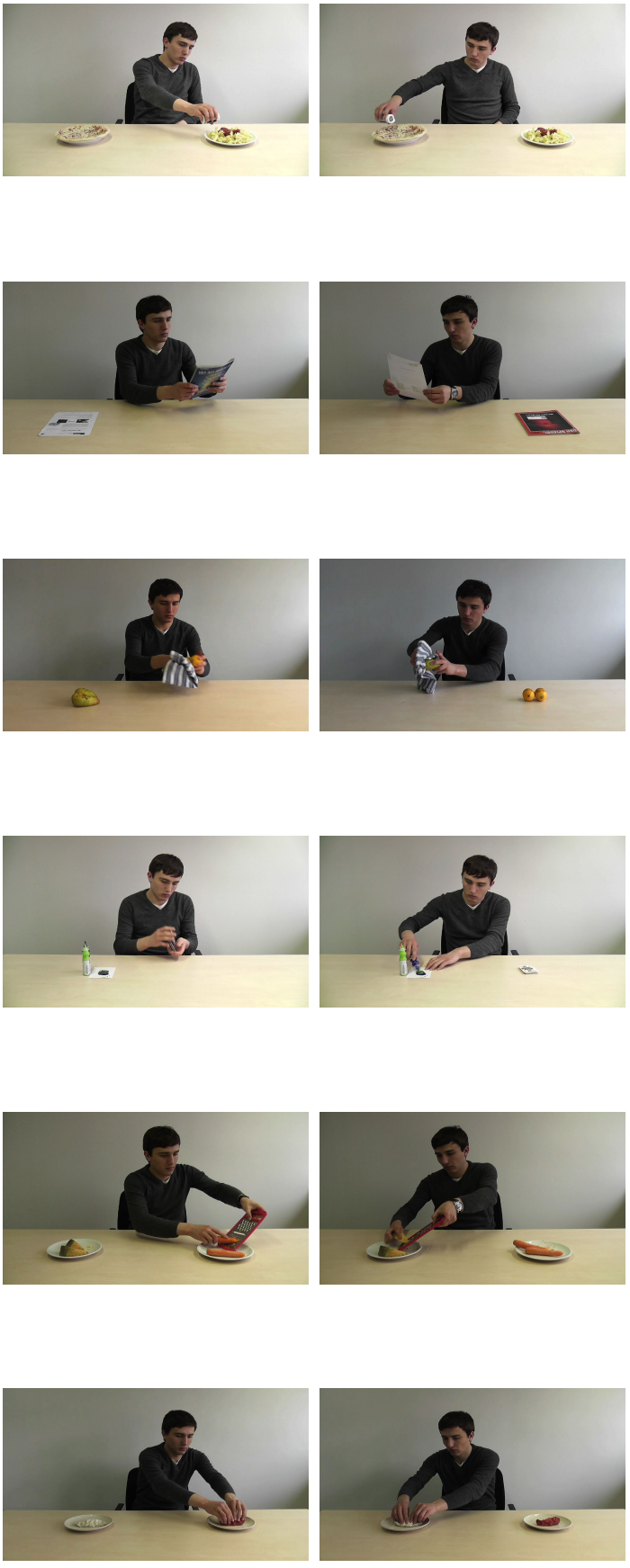
Snapshots of Video-Events, Experiments 1 to 4

The snapshots of the experimental video events (Experiments 1 to 4) for the sentences are listed below. The order of video events is presented according to the sentence order.





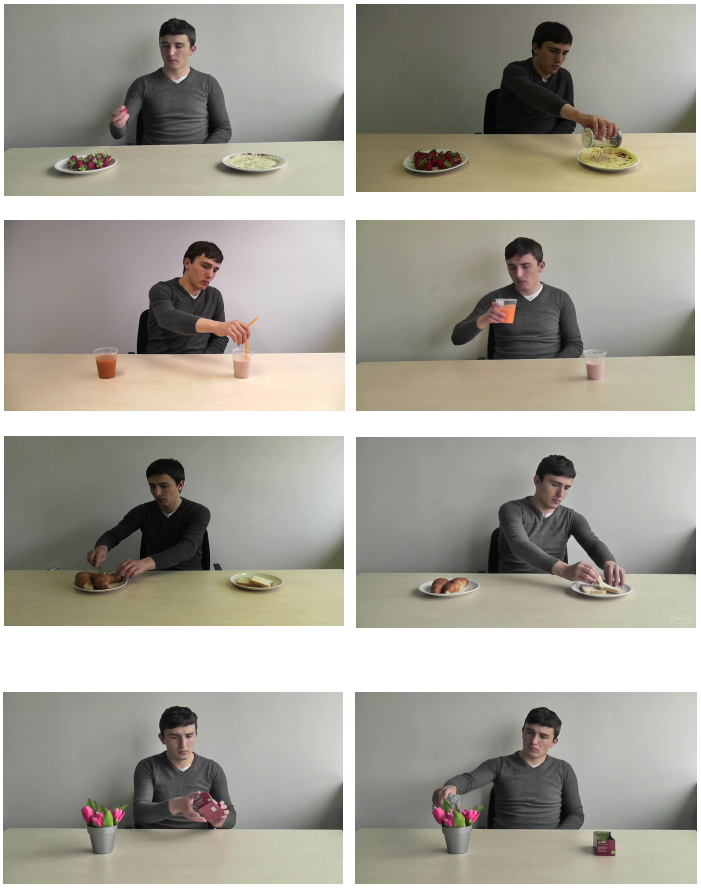


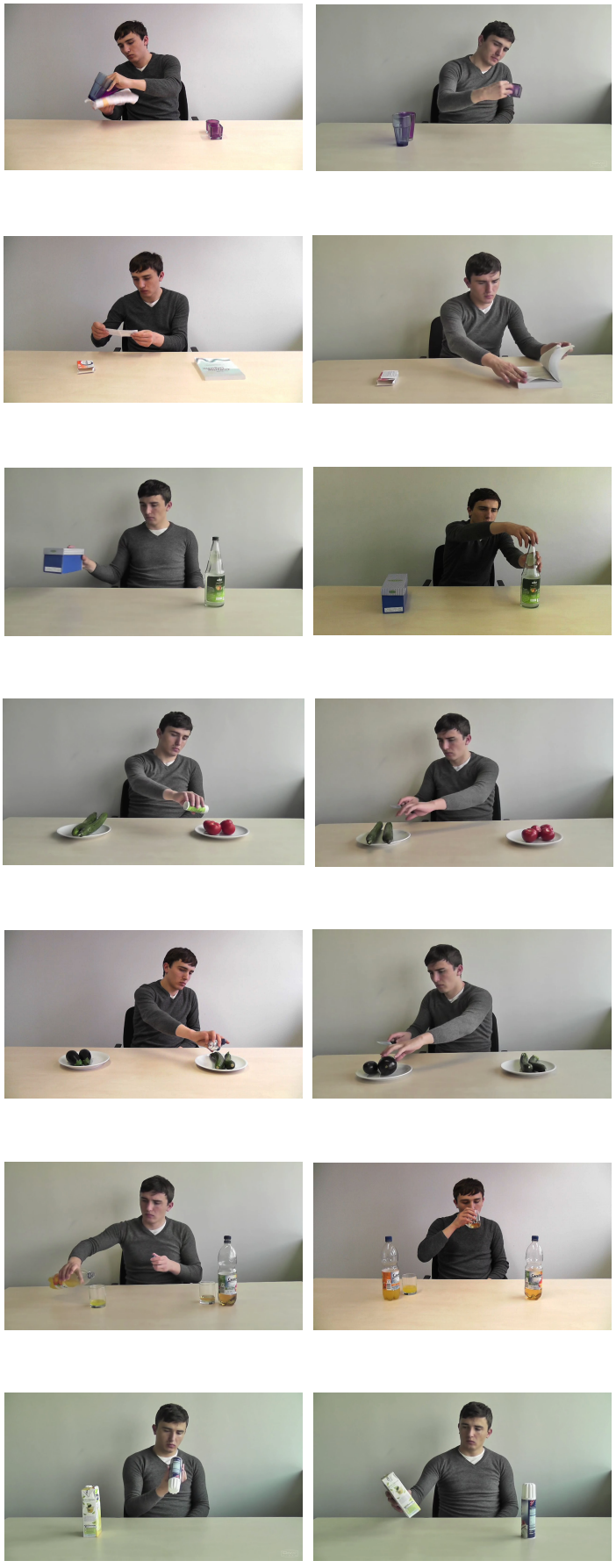


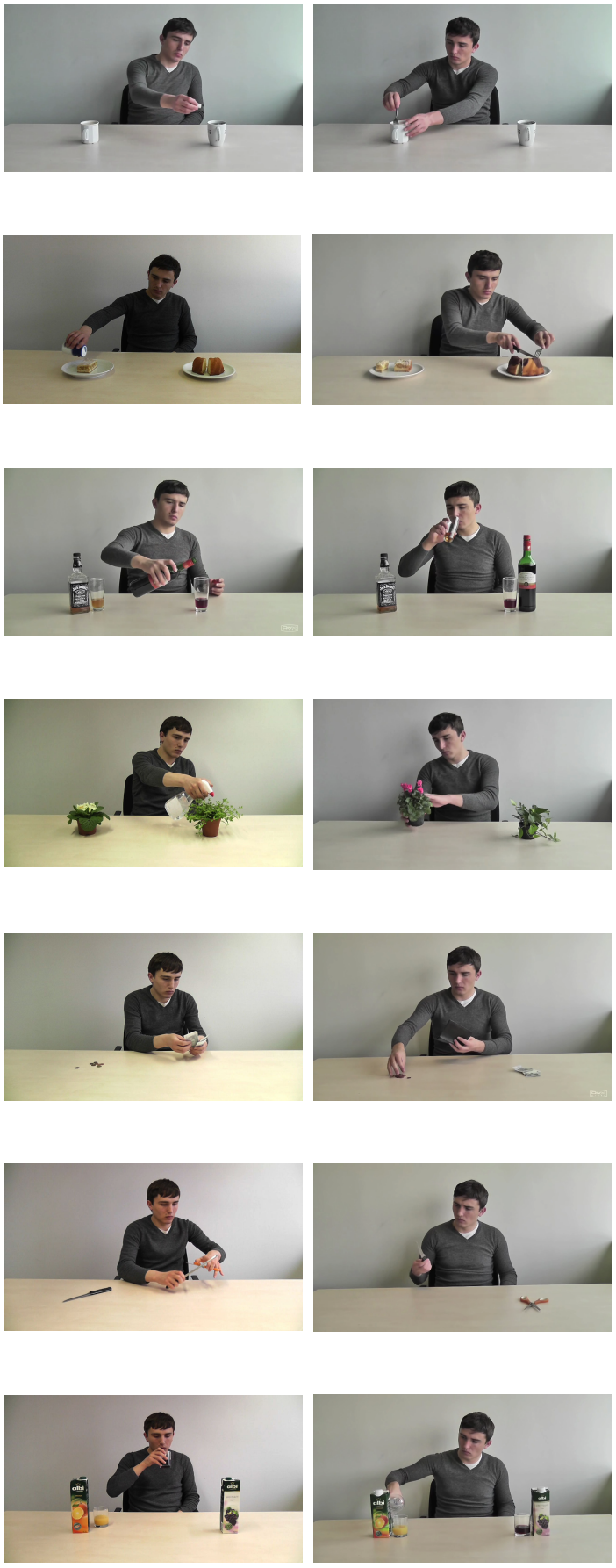
Appendix C

Snapshots of Video-Events, Experiment 5

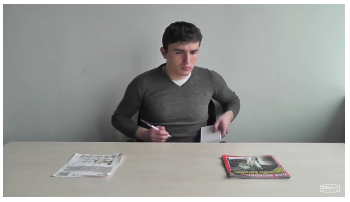
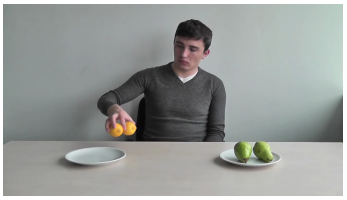
The video snapshots of the experimental video events (incongruence for the recent event) are listed below. The order of video events is presented according to the sentence order.

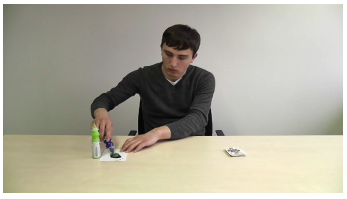
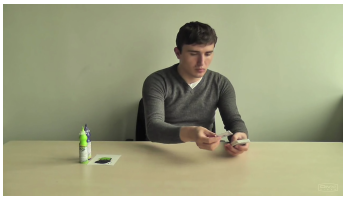


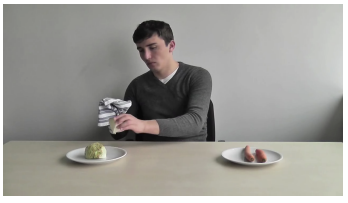
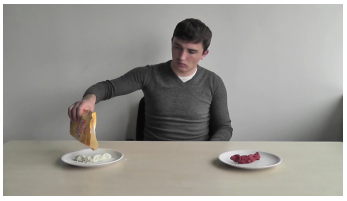







Appendix D

Time Course Graphs as a function of the estimated probability (high vs. low) of a second action on the same object

The Graph D1a presents the gaze data of 11 items for which the probability of a second action on the same object has been estimated as being low. The Graph D1b presents the gaze data of 13 items for which the probability of a second action on the same object has been estimated as being high (the gaze data in both graphs are from Experiment 1).

FIGURE D.1: D1a

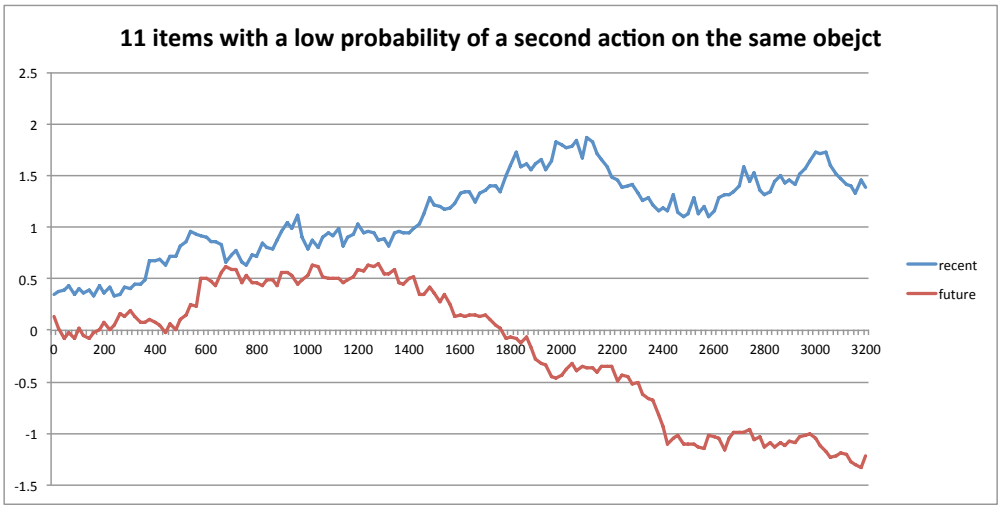
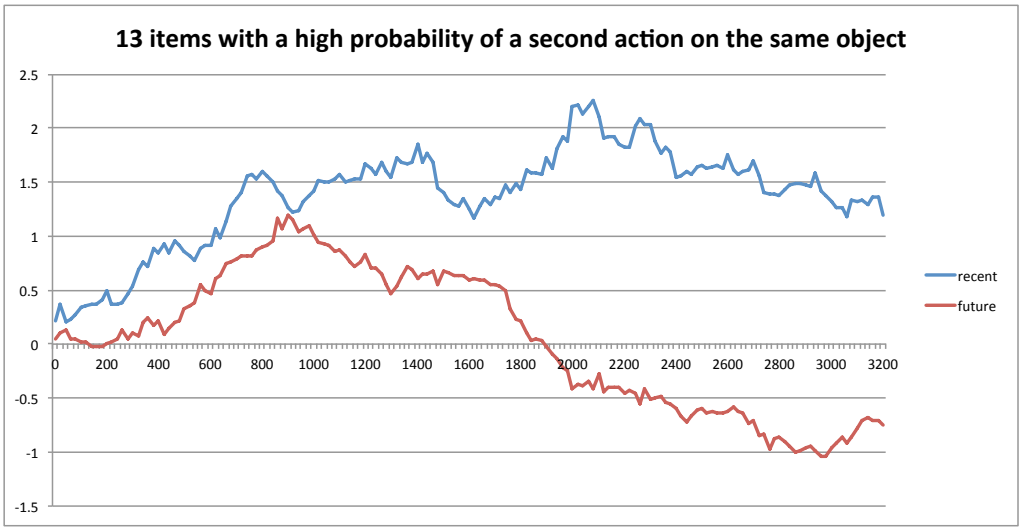


FIGURE D.2: D1b



Bibliography

Abashidze, D., & Chambers, C. G. (2016). The role of early linguistic cues in the recent event preference. In *Proceedings of the Conference on Architectures and Mechanisms for Language Processing.* BCBL, University of the Basque Country.

Abashidze, D., Knoeferle, P., Carminati, M. N., & Essig, K. (2011). The role of recent real-world versus future events in the comprehension of referentially ambiguous sentences: Evidence from eye tracking. In B. Kokinov, A. Karmiloff-Smith, & N. J. Nersessian (Eds.), *Proceedings of the European Conference on Cognitive Science.* New Bulgarian University Press.

Adams, R. B., Pauker, K., & Weisbuch, M. (2010). Looking the other way: The role of gaze direction in the cross-race memory effect. *Journal of Experimental Social Psychology*, *46*(2), 478–481.

Altmann, G. (2004). Language-mediated eye-movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, *93*, B79–B87.

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

Altmann, G., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502-518.

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

Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*(4), 583–609.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111.

Anderson, R. J., & Dewhurst, S. A. (2009). Remembering the past and imagining the future: Differences in event specificity of spontaneously generated thought. *Memory*, *17*(4), 367–373.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.

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

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1521), 1235–1243.

Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, *18*(4), 438–462.

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Observing shared attention modulates gaze following. *Cognition*, *120*(2), 292–298.

Boland, J. E. (2005). Visual arguments. *Cognition*, *95*(3), 237–274.

Bos, L. S., Hanne, S., Wartenburger, I., & Bastiaanse, R. (2014). Losing track of time? processing of time reference inflection in agrammatic and healthy speakers of german. *Neuropsychologia*, *65*, 180–190.

Britt, A. E., Mirman, D., Kornilov, S. A., & Magnuson, J. S. (2014). Effect of repetition proportion on language-driven anticipatory eye movements. *Acta Psychologica*, *145*, 128–138.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*(1), 45.

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory and Cognition*, *25*(4), 438–446.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language*, *47*(1), 30–49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 687.

Chapman, C. S., Gallivan, J. P., Wood, D. K., Milne, J. L., Culham, J. C., & Goodale, M. A. (2010). Reaching for the unknown: multiple target encoding and real-time decision-making in a rapid reach task. *Cognition*, *116*(2), 168–176.

Coco, M. I., & Keller, F. (2014). The interaction of visual and linguistic saliency during syntactic ambiguity resolution. *The Quarterly Journal of Experimental Psychology*, *68*(1), 46–74.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317–367.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin and Review*, *12*(3), 453–459.

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and Cognition*, *13*(4), 844–858.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509–540.

Dumitru, M. L., Joergensen, G. H., Cruickshank, A. G., & Altmann, G. T. (2013). Language-guided visual processing affects reasoning: The role of referential and spatial anchoring. *Consciousness and Cognition*, *22*(2), 562–571.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*, 409–436.

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

Friesen, C. K., Moore, C., & Kingstone, A. (2005). Does gaze direction really trigger a reflexive shift of spatial attention? *Brain and Cognition*, *57*(1), 66–69.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694.

Gilead, M., Liberman, N., & Maril, A. (2013). The language of future-thought: an fmri study of embodiment and tense processing. *Neuroimage*, *65*, 267–279.

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

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Cognitive Systems Research*, *1*(1), 19–33.

Habets, B., Kita, S., Shao, Z., Özyurek, A., & Hagoort, P. (2011). The role of synchrony and ambiguity in speech–gesture integration during comprehension. *Journal of Cognitive Neuroscience*, *23*(8), 1845–1854.

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

Haskell, T. R., Thornton, R., & MacDonald, M. C. (2010). Experience and grammatical agreement: Statistical learning shapes number agreement production. *Cognition*, *114*(2), 151–164.

Hawkins, J., & Blakeslee, S. (2007). *On intelligence*. Macmillan.

Huettig, F., & Altmann, G. T. (2011). Looking at anything that is green when hearing "frog": How object surface colour and stored object colour knowledge influence language-mediated overt attention. *The Quarterly Journal of Experimental Psychology*, *64*(1), 122–145.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*(4), 460–482.

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

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics*, *40*(6), 431–439.

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*(3), 244–253.

Kako, E., & Trueswell, J. C. (2000). Verb meanings, object affordances, and the incremental restriction of reference. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 256–261).

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

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*, 37–55.

Kaschak, M. P., Loney, R. A., & Borreggine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition*, *99*(3), B73–B82.

Knoeferle, P. (2015a). Characterising visual context effects: active, pervasive, but resource-limited. Retrieved from https://www.researchgate.net

Knoeferle, P. (2015b). Language comprehension in rich non-linguistic contexts: combining eye-tracking and event-related brain potentials. *Cognitive Neuroscience of Natural Language Use*, 77.

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

Knoeferle, P., & Crocker, M. (2005). Incremental effects of mismatch during picture-sentence integration: Evidence from eye-tracking. In *Proceedings of the 26th annual conference of the cognitive science society* (pp. 1166–1171).

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science*, *30*(3), 481–529.

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

Knoeferle, P., Crocker, M. W., Scheepers, C., & 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*(1), 95–127.

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

Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from erps and picture-sentence verification. *Psychophysiology*, *48*(4), 495–506.

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

Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin and Review*, *14*(5), 846–851.

Kreysa, H., & Knoeferle, P. (2013). Reference-related speaker gaze as a cue in online sentence processing. In D. Kluge A. & Söffker (Ed.), *Mensch Teams Systeme und Automaten.* Duisburg: DuEPublico.de.

Kreysa, H., Knoeferle, P., & Nunnemann, E. M. (2014). Effects of speaker gaze versus depicted actions on visual attention during sentence comprehension. In *Proceedings of the 36th annual meeting of the cognitive science society.*

Langton, S. R., O'donnell, C., Riby, D. M., & Ballantyne, C. J. (2006). Gaze cues influence the allocation of attention in natural scene viewing. *The Quarterly Journal of Experimental Psychology*, *59*(12), 2056–2064.

Liu, B., Wang, Z., & Li, J. (2011). The influence of matching degrees of synchronous auditory and visual information in videos of real-world events on cognitive integration: an event-related potential study. *Neuroscience*, *194*, 19–26.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

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

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

Madden, C. J., & Therriault, D. J. (2009). Verb aspect and perceptual simulations. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1294–1303.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory and Cognition*, *31*(5), 663–672.

Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition*, *10*(1), 7–14.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165.

Nappa, R., Wessel, A., McEldoon, K. L., Gleitman, L. R., & Trueswell, J. C. (2009). Use of speaker's gaze and syntax in verb learning. *Language Learning and Development*, *5*(4), 203–234.

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science*, 0956797611429800.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*(4), 369–378.

Pusch, M., & Loomis, J. (2001). Judging another person's facing direction using peripheral vision. *Journal of Vision*, *1*(3), 288–288.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition*, *14*(3), 191–201.

Reali, F., & Christiansen, M. H. (2007). Word chunk frequencies affect the processing of pronominal object-relative clauses. *The Quarterly Journal of Experimental Psychology*, *60*(2), 161–170.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*(5), 767–780.

Rodríguez, A., Burigo, M., & Knoeferle, P. (2015). Visual gender cues elicit agent expectations: different mismatches in situated language comprehension. In *Proceedings of the Euroasianpacific Joint Conference on Cognitive Science (EAPCogsci 2015)* (Vol. 1419).

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, and Psychophysics*, *77*(3), 720–730.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–148.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context

on syntactic ambiguity resolution. *Cognitive Psychology*, *45*(4), 447–481.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156.

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

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

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

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64–82.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes*, *6*(4), 303–338.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience*, *30*(33), 11177–11187.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Underwood, G., Jebbett, L., & Roberts, K. (2004). Inspecting pictures for information to verify a sentence: Eye movements in general encoding and in focused search. *Quarterly Journal of Experimental Psychology Section A*, *57*(1), 165–182.

Vanyukov, P. M., Warren, T., Wheeler, M. E., & Reichle, E. D. (2012). The emergence of frequency effects in eye movements. *Cognition*, *123*(1), 185–189.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological Bulletin*,

133(2), 273.

Zelinsky, G. J., & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception and Psychophysics*, *67*(4), 676–690.

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

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.