The role of morphosyntactic cues on anticipatory sentence processing within a rich visual context: Evidence from eye-tracking

Dato Abashidze & Angela Schmidt
The Leibniz-Centre General Linguistics (ZAS)
Language Development and Multilingualism, Psycholinguistics Laboratory
Berlin, Germany, Email: abashidze@leibniz-zas.de

Zurab Oqropiridze
Department of Theoretical and Applied Linguistics
Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia

Abstract

Eye-tracking research has revealed that people rely more on a recently enacted action-event than consider a plausible future action-event when hearing a sentence referring to a visual scene. When participants encountered a recently enacted action-event, and then listened to a (NP1-Verb-Adv-NP2) past or futuric present tense sentence in German, they inspected the target of the recently seen event more often than that of the equally plausible future target shown, irrespective of sentence tense. These preferential looks towards the recent target persisted even when future events and futuric present sentences were presented with greater frequency within the experiment. The current experiments assessed whether the preferential looks toward recent targets occur in similarly structured English sentences containing earlier, more localized tense markers (auxiliary verbs: will/has), and in Georgian sentences containing an even earlier case marking at the first noun phrase (nominative and ergative case). Can the early morphosyntactic cues eliminate the preferential inspection of the recent target? Results revealed that when participants processed the tense marker in Experiment 1 (English) and 2 (Georgian), the bias towards looking at the recent target was reduced. However, in Georgian, the morphological marker on its own was not able to eliminate the strength of the recent event. In both experiments, participants rapidly started to decrease their looks towards the recent target when exposed to the clear future tense cues at the verbs and this gaze pattern continued into the later word regions. This shows that participants were able to use the tense cues but only partly use the morphosyntactic cues to anticipate sentence referents to the visual context.

Keywords: eye-tracking, sentence comprehension, morphosyntactic cues, split ergativity, anticipatory processing, recent-event preference

Introduction

Psycholinguistic investigation has shown the effect that concurrent visual and linguistic information has on language processing. Results from eye-tracking studies have revealed the rapidity with which language comprehension influences people's eye-movements to visual referents (e.g., Chambers et al., 2004; Tanenhaus et al., 1995). Some studies have shown that linguistic factors, such as tense or morphosyntactic verbal cues, can aid participants in developing expectations about upcoming referents and events (e.g., Altmann & Kamide, 2007; Garrido Rodriguez et al., 2023). At the same time, listeners show a preference for inspecting the target of a recently performed event, when encountering an utterance that might indicate either past or future event (e.g., Abashidze et al., 2011; Knoeferle et al., 2011). Although most authors agree on a close interplay between visual and linguistic contextual cues in language processing (see, e.g., Rigoli & Spivey, 2015, for a review), questions remain regarding the strength of the cues. The current studies examined to which extent the preference for interpreting an utterance towards a recent (vs. future) event holds when comprehenders encounter early localized morphosyntactic cues that might indicate a future event.

Rich visual context cues in language processing

A number of studies have revealed an attentional behavior that has been dubbed the recent-event preference (i.e., Abashidze et al., 2011; Knoeferle et al., 2011; Knoeferle & Crocker, 2007). Findings from these studies, which initially used clipart images and then, in later studies, action events performed by animate human agents, revealed that comprehenders looked more at the target object of a recently performed action vs. that of an equally plausible future event. However, it was noted that the experimental procedure, in which the participants never saw the future action carried out may have caused them to rely more on the recently seen event (Altmann & Kamide, 2009).

Taking this concern into account, Abashidze et al. (2011) and Knoeferle et al. (2011) modified their experimental setup to ensure an equally balanced (50/50) presentation of both past and future action events. Furthermore, the authors modified the visual materials by creating a richer visual context by using real-world events, in contrast to the past research that utilized statically depicted events. The findings of this study (Knoeferle et al., 2011) confirmed that, overall, people preferred to ground utterances in the recently seen event, initially disregarding the cues provided by the linguistic stimuli. Remarkably, in a follow-up study by Abashidze et al. (2019), which also applied a rich visual context by using video-based stimuli, the presentation of future tense and event with greater relative frequency (75 vs. 25% and 88 vs. 12% in Experiments 1 and 2, respectively) was only able to diminish the duration of the recent-event preference but not override it altogether. These results were also found in other studies that used actor gaze cues and gender markers in an attempt to manipulate speedier and more frequent looks to the future event (Abashidze & Knoeferle, 2021; Rodriguez et al., 2015).

The effect of linguistic cues in language processing

Whereas the above studies argue for the strength of the visual scene over linguistic information, other studies show that ut-

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

terance interpretation guides visual attention. For example, the work of Chambers and San Juan (2008) showed that linguistic context and language as a communicative act determined interaction with visual information. Similarly, a study by Altmann and Kamide (2007) revealed that tense cue together with the affordances provided by observed objects influenced the visual selection of the target object mentioned in the sentence. Furthermore, Kamide, Altmann, et al. (2003) found in experiments with English and Japanese that linguistic and non-linguistic information encoded at each sentence constituent is incrementally used to anticipate forthcoming constituents and influences viewing behavior. Further evidence for the use of verbal information in prediction by adult speakers of Dutch and Turkish was provided by the work of Brouwer et al. (2019). In both of these studies, the visual stimuli was presented through static images.

Additionally, numerous studies have investigated the effect that information encoded at the morphosyntactic level has on utterance interpretation. An early study by Kempe (1999) investigating the use of case markers, word-order, animacy cues, and cue interaction in online processing, found differences between the two languages studied, Russian and German. Russian speakers made stronger use of case markers, whereas German speakers tended to rely more on animacy cues. In both languages, they found both the order of appearance and strength of cues to influence processing. Earlier cues tended to be the most relied on unless a stronger cue presented later was provided. Case-marking of noun phrases (nominative vs. accusative) shortly after sentence onset in German had an anticipatory effect echoed in participants´ interpretation of the second NP as being the agent or patient of the utterance (Kamide, Scheepers, et al., 2003). Another study within the visual-world paradigm showed that grammatical gender encoded in the determiner/article was sufficient to quickly direct eye-gaze to the target object before the noun (Huettig & Janse, 2016). A further study of German revealed that comprehenders were able to use verb suffix cues to predict incoming information (Koch et al., 2021).

A recent eye-gaze study by Garrido Rodriguez et al. (2023) investigated a Mayan langauge (word order: VOS) in which rich syntactic information is encoded in the verb at the sentence initial position. They found that comprehenders speedily interpreted the morphosyntactic verbal cues to predict the upcoming referents. The predictive use of morphsyntactic markers was furthermore investigated in a study by Covey et al. (2018). Results of this study revealed that multilingual speakers of Hindi used gender agreement dependency information to rapidly facilitate access to a target noun during online processing. In a study that tested language learners' capacity to exploit recently acquired knowledge about novel morphosyntactic cue patterns, Abashidze et al. (2020) found that learners of Georgian applied both the verb and noun cues presented within the same sentence. Furthermore, another study in Georgian by Skopeteas et al. (2012) investigated the processing of dative subjects and objects and found that case marking plays a more important role in online parsing than word order.

As can be gathered from the evidence presented thus far, both visual and linguistic cues are clearly utilized during sentence comprehension. Both language and visual context inform each other when concurrently perceived. In particular, when they are co-referential. Psycholinguistic findings (see, e.g., Kamide, Scheepers, et al., 2003) have also shown that the sequential integration of linguistic cues during utterance interpretation allows comprehenders to anticipate what may come next and update information seen in a visual display. However, less is known about the role that linguistic cues, especially morphosyntactic cues, play when interpreting utterances that are embedded in a rich visual context. Previous research has shown that when viewers see an action having being performed they use this information as a reliable cue when processing linguistic verbal information. The extent to which this reliance can be diminished is still under debate. Furthermore, the use of pre-verbal information, for example that of morphological markers of a noun phrase preceding a verb in a split-ergative language has not yet been investigated to our knowledge. By conducting two eye-tracking experiments, we aim to fill this gap and expand previous research on incremental sentence processing within a rich visual context. For this purpose, we utilized early tense markers in English and an earlier morphological marker of nominative/ergative cases in Georgian. In Georgian, the morphological case marking of noun phrases depends on a number of mutually independent parameters: such as the tense series of the verbs of the clause; the lexical class of the verb; and the grammatical relation which the noun phrase bears. The verb tense is a most relevant factor in the current Georgian study, as it must align with the morphosyntactic markers at the nouns. When a subject/noun phrase appears in the nominative, marked with -i, the verb in the future tense agrees with that subject; the object then takes the dative case, marked with -s (see Table 1a/a'). When a subject appears in the ergative, marked with -ma, the verb in the aorist tense[1] agrees with that subejct; the direct object in nominative is then marked with -i (see Table 1b/b').

Current Study

Given that morphosyntactic cues have been shown to influence participants' visual attention and generate expectations about upcoming event information, the current two eye-tracking studies examined the effect that early linguistic cues in English and Georgian have on utterance interpretation within a rich visual context. In the English sentences (Experiment 1), an early localized tense marker was applied. In the Georgian linguistic stimuli (Experiment 2), an even earlier morphological marker was introduced. We tested to what extent these early linguistic cues might reduce the preferential observation of the target object of the recent action-event. We examine whether people are able to quickly integrate the in-

[1]Note: The Georgian aorist tense governs the subject of the sentence and assigns it the ergative case.

formation provided by the morphological marker in Georgian and at the auxiliary verb in English during sentence interpretation. The tense cue provided by these structures might enable comprehenders to rapidly decide whether the sentence is referring to the target of a recently seen or equally plausible future event. If, in the current study, the strength of the rich visual context is sensitive to the early localized tense markers, then we should see a decrease in inspections of the recent event target when participants hear the cue words that indicate a future action. Consequently, we might find no overall preference for viewing the recently seen event, in contrast to previous findings (e.g., Abashidze et al., 2019; Knoeferle et al., 2011).

Experiments 1 and 2: Methods

Participants

Twenty-eight English native speakers (mean age=22.4; range 18 to 27; 12 males 16 females, all of them right-handed), all students of the University of Toronto Mississauga, participated in Experiment 1. Twenty-four Georgian native speakers (mean age=19.7; age range 18 to 24; 9 males 15 females, all of them right-handed) from different Universities in Georgia took part in Experiment 2. All participants were paid 10 US Dollars /10 Lari or they received credit points for their participation. They all had normal or corrected-to-normal vision and were unaware of the purpose of the experiment.

Materials and design

The visual materials for both experiments were the same as those used in a previous experiment by Abashidze et al. (2019). In this experiment, pre-recorded action sequences with an average duration of 5015 ms were created for each experiment trial. Each trial comprised two short videos showing an agent and two objects (e.g., a person seated at a table on which a plate with tomatoes and a plate with cucumbers lay). The first video showed the agent interact with one of two objects in the display (e.g., flavoring the cucumbers, Figure 1, a). The second video showed the same interaction with the other object (e.g., flavoring the tomatoes, Figure 1, c). The position of the target objects (right vs. left) was counterbalanced across items. Figure 1 shows a typical order of presentation of the visual stimuli used in the study.

For the linguistic materials, new utterances were recorded in both English or Georgian languages. In Experiment 1 (English), all critical sentences had the structure NP1-Aux-Verb-NP2 in either aorist past or future tense (example: "The experimenter will/has flavor/ed the tomatoes.") and were recorded by a male native English speaker (see English in Table 1). In Experiment 2 (Georgian), all critical sentences had the structure NP1-Verb-Adv-NP2 in either past or future tense (example: Eksp'eriment'at'or-ma daamarila uk've k'it'r-I, 'Experimenter-ERG salted:AOR.3SG recently cucumber-NOM') referring to a recently performed action or (Eksp'eriment'at'or-i daamarilebs male p'omidor-s, 'Experimenter-NOM salt:FUT.3SG soon tomato-DAT) referring to a plausible future action. They were recorded by a female native speaker (see Georgian in Table 1).

Each experimental item consisted of two action videos, a snapshot (i.e., a static photo of the last frame of the first video showing the agent in an inactive position Figure 1, b), and a spoken sentence.

Figure 1: An example of the order of presentation of the visual stimuli in a typical experimental trial.

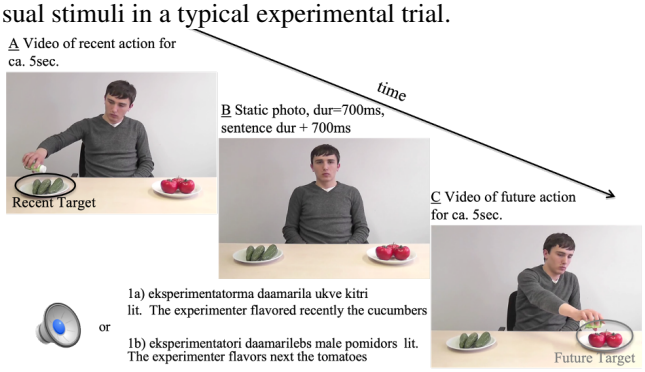


Table 1: Experimental sentences and conditions

Tense, sentence conditions, counterbalancing
English
1a Future: The experimenter will flavor the tomatoes.
1a' Future: The experimenter will flavor the cucumbers.
1b Past: The experimenter has flavored the tomatoes.
1b' Past: The experimenter has flavored the cucumbers.
Georgian
1a Future: Eksperimentator-*i* daamarilebs male pomidor-*s*
Lit. Experimenter-Nom salt:FUT.3SG soon tomato-Dat
1a' Future: Eksperimentator-*i* daamarilebs male kitr-*s*
Lit. Experimenter-Nom salt:FUT.3SG soon cucumber-Dat
1b Past: Eksperimentator-*ma* daamarila pomidor-*i*
Experimenter-Erg salted:AOR.3SG recently cucumber-ABS
1b' Past: Eksperimentator-*ma* daamarila ukve kitr-*i*
Experimenter-Erg salted:AOR.3SG recently tomato-ABS

Participants first saw video 1. Next, they heard the utterance while the static image of the final frame of video 1 remained on the screen (see Figure 1, b), and then they saw video 2. The frequency of referring to recent/future events and the occurrence of future/past tense was counterbalanced. In addition to the 24 critical items in English and in Georgian, 40 filler items were included that comprised a variety of different scenarios or events in either the past or future tense.

Procedure

Upon arrival at the lab, participants were given a printed document with the pertinent details and asked to make sure they understood it. They were informed that the videos would be shown on a screen and sentences heard through a loudspeaker. Next, they were seated in a quiet experimental room, facing a monitor at a distance of approximately 65 cm from the screen. An EyeLink 1000 desktop head-stabilized eye tracker (SR Research Ltd., Canada) was used to capture participants' eye movements in Experiment 1. A Tobii eye tracker (Tobii Fusion Pro, Tobii Technology AB, Sweden) was used to monitor and record gaze data of participants in Experiment 2. The eye-tracking experiment was a slightly modified look-and-listen paradigm in which participants were instructed to understand as much as possible. The gaze of each participant was successfully calibrated prior to beginning the test proceedings using a 9-point calibration procedure, in which an attention-getter appeared in every position of a 3-by-3 grid of calibration points. Each trial started with a centrally located white dot which appeared against a black background for 500 ms, followed by the presentation of the first action-event video. The final frame of this video was shown and silently remained on the screen for 700 ms, which was followed by the onset of a past or future tense sentence. With the offset of the sentence, the static image disappeared, and participants were shown the second action-event video. Participants could take a short break halfway through the eye-tracking experiment. The entire duration of the experiment was approximately 35-40 min.

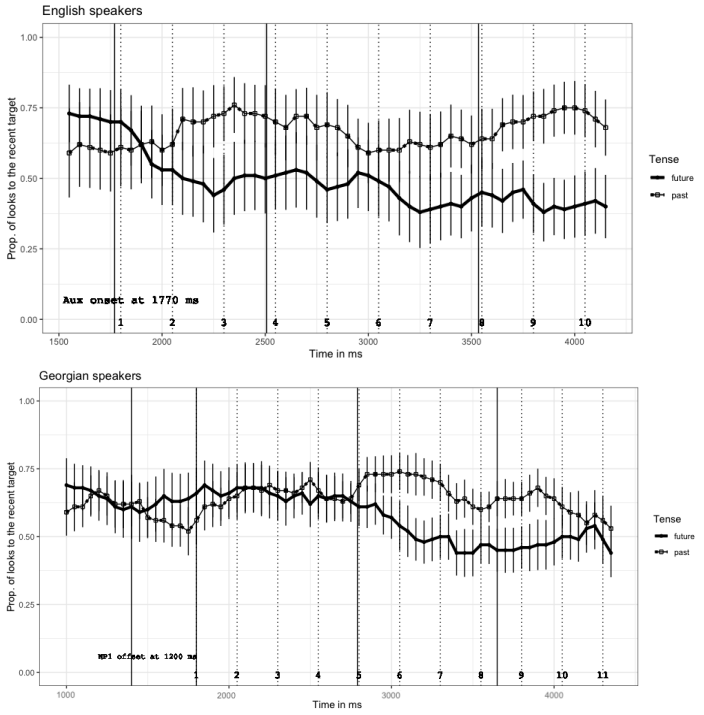
Analyses

For the eye-tracking data, in Experiment 1, we used EyeLink Data Viewer to export the fixation data for the analyses. In Experiment 2, we used the Tobii Pro Lab software to export the fixation data for the analyses. We defined two areas of interest, which comprised participants eye-movements either towards the recent target object (e.g., the cucumbers) or the future target object (e.g., the tomatoes, see Figure 1). Because the target object images of each trial differed in size, rectangular areas of interest were defined around each image. Before conducting the analyses, the raw data were cleaned and checked for validity by using the eyetrackingR package (Dink & Ferguson, 2015). Fixations shorter than 80 and longer than 1500 milliseconds were removed from the data. Furthermore, the trials with more than 25% of trackloss proportion were not included in the analyses.

Next, we computed gaze proportions to the two target objects in each successive 50 ms time slot, starting from the onset of the Auxiliary until the end of the sentence (Aux-Verb-NP2) in Experiment 1, and starting from the first noun phrase until the end of the sentence (NP1-Verb-Adv-NP2) in Experiment 2. The length of the three critical word regions in Experiment 1 are: The auxiliary region (from auxiliary onset until verb onset, mean duration = 735 ms), the verb region (from verb onset until the second noun phrase (NP2) onset, mean duration = 975 ms), and the NP2 region (from NP2 onset until NP2 offset, mean duration = 725 ms). The length of the four critical word regions in Experiment 2 are: The first noun region (from noun onset until verb onset, mean duration = 1646 ms), the verb region (from verb onset until adverb onset, mean duration = 991 ms), the adverb region (from adverb onset until the second noun phrase (NP2) onset, mean duration = 791 ms) and the NP2 region (from NP2 onset until NP2 offset, mean duration = 815 ms). In the first time window/region, fixations were counted for the analyses if they started in that time window.

As for the time-course analyses, the graph of Experiment 1 depicts data of a 2650 ms window of the sentences (with a mean length of 4100 milliseconds), which consist of the 200 ms before auxiliary onset, and an additional 2430 milliseconds after the onset of the auxiliary verb (see Figure 2). The graph of Experiment 2 depicts data of a 3250 ms window of the sentence that consist of the 400 ms before the offset of the first noun phrase and additional 2850 milliseconds after the onset of the verb (see Figure 2).

Figure 2: Mean proportion (with *SE*) of looks to the recent target depending on the sentence tense in Experiments 1 and 2.



Note: the looks to the future target were complementary to the looks to the recent target.

The graphs present the proportion of looks to both the recent and future target objects. The proportions of looks are presented as follows: the dotted lines indicate trials with the past tense sentences, and the solid lines indicate trials with the future tense sentences. The three vertical solid lines indicate the mean *average* onset of the critical word regions. For

Experiment 1, the first vertical line at 1770 ms indicates the onset of the auxiliary verbs (*will* or *has*). The second vertical line at 2507 ms indicates the onset of the main verb. The third vertical line at 3534 ms indicates the onset of the second noun phrase. For Experiment 2, the first vertical line at 1400 ms indicates the offset of the first noun phrase /case marking (*-i* or *ma*). The second vertical line at 1780 ms indicates the onset of the main verb. The third vertical line at 2752 ms indicates the onset of the adverb. The fourth vertical line at 3380 ms indicates the onset of the second noun phrase.

For the statistical analyses, we calculated the empirical logit for the looks to the recent target object, and the analysis was weighted using the procedure of Barr (2008) in both experiments. The lme4 package (version 1.1-32) was utalized to calculate linear mixed-effects models to assess the fixed effects of Tense and Time and their interactions in the critical time windows.

Results

Eye-movement data in the graphs (Figure 2) show an overall preference for looking at the recent target relative to the future target in the early sentence word regions, as looks for both the past and future tense conditions remained at/above the chance level until 3100 ms in Experiment 1 and until 3350 in Experiment 2 (indicating that the recently acted upon object received more inspections than the object acted upon in the future event).

In Experiment 1, when participants listened to the auxiliary verbs, as a function of tense, their looks started to diverge continuously towards the recent and future targets. While the amount of inspections of the recent target in the past tense condition increased from the auxiliary verb region, the inspections in the future tense condition started to decrease continuously until the middle of the main verb region (the differences in this region were also confirmed by the statistical analyses). From the middle of the verb regions, subjects inspected the future target more in the future tense condition than in the past tense condition. The looks towards the recent target in the past tense remained mostly at the same level throughout the last two word regions. The diverging looks towards the recent target in the past and future tense conditions in all three word regions were confirmed by the statistical analyses.

In Experiment 2, when participants listened to the first noun phrase (with morphological marking referring to the recent target in the ergative case and to the future target in the nominative case), they did not reveal an immediate usage of this cues as the auxiliary verb cues in Experiment 1, however as in the first experiment, they showed a clear reduced gaze pattern towards the recent target after encountering the main verb tense information. In the early middle adverb region subjects started to inspect the future event target more in the future tense than in the past condition. The tense effect was reliably in the last two word regions.

The overall gaze pattern in the NP2 region remained similar in both Experiments. As expected, in the past tense conditions preferential observation occurred towards the recent target and in the future tense condition preferential looks appeared towards the future target.

Discussion

Across two experiments, we investigated early morphosyntactic cues within a rich visual context (e.g., Abashidze et al., 2019; Garrido Rodriguez et al., 2023; Knoeferle et al., 2011) and asked to what extent the cues provided by the auxiliary verb in Experiment 1 and the morphological marker in Experiment 2 would facilitate the anticipation of upcoming utterances that referred to a plausible future action event. We predicted that if the early localized tense and morphological marking is a strong anticipatory cue (e.g., Brouwer et al., 2019; Kamide, Altmann, et al., 2003; Koch et al., 2021) within the rich visual context, then participants' overall preferential inspection of the recent target should decrease, or even disappear, upon encountering these cues in the future tense condition.

The eye-movement data in both experiments revealed that when participants heard the early morphosyntactic cues in the future tense condition they initiated a decrease of looks towards the recent target. Despite there being a decrease of fixations towards the recent target, they remained at the chance level until the middle of the main verb in Experiment 1 and above this level (indicating a clear recent event preference) until 200 ms after the verb offset in Experiment 2. However, importantly, in both experiments, the decrease of the gaze pattern towards the recent target occurred earlier than in the previous studies employing the futuric present tense in German (e.g., Abashidze et al., 2011; Knoeferle et al., 2011). Furthermore, an assumption made in previous studies (e.g., Abashidze et al., 2011; Knoeferle et al., 2011), that the verb processing might govern the recent event preference – in other words, that the bias towards the recent target remains strong or even increases during the verb region – could not fully be confirmed with the current data. Importantly, in line with other findings (e.g., Altmann & Kamide, 2007; Kamide, Scheepers, et al., 2003), participants in both experiments clearly utilized the verb tense cue that was supported by the auxiliary verb in Experiment 1 and by the morphological marking in Experiment 2 to anticipate the upcoming event information – they looked at the future event target before it was mentioned. These results are in line with previous findings that investigated the use of suffix markers at the verb and show that people quickly integrated these morphosystactic cues in sentence comprehension (Koch et al., 2021).

Our findings differ from Garrido Rodriguez et al. (2023), who found that morphosyntactic cues encoded in the sentence initial position did elicit anticipatory looks. As the current Georgian data does not show this immediate effect with the morphological marking. One explanation could be that in the Georgian sentences, morphosyntactic information was encoded in the noun phrase, whereas in the Mayan language

it was encoded in a transitive verb (VOS word order). It is possible that, in Georgian, the effect of the pre-verbal morphological marker is weak for two reasons. First, these nominative and ergative markers can appear in both future and past tense sentences depending on the verb class and series. Second, participants might have come to rely on the combined morphological and tense cues and therefore waited until encountering the verb information before deciding which object the sentence was referring to. Evidence supporting this assumption was furthermore found in a Georgian language learning study by Abashidze et al., 2020, in which learners' patterns indicated a comparable reliance on both cues. Perhaps for this reason, inspection of the recent event target persisted only from the end of the verb region, which shows that the tense cue (or the combination of the morphological marking with the tense cue) was utilized in predicting the plausible future target resulting in looks toward the future target before it was mentioned.

An important distinction between the current study and other studies in which morphosyntactic cues were used predictively (for example, the study by Huettig and Janse (2016) in which sentence-initial morphosyntactic cues elicited anticipatory looks) regards the visual materials used. In contrast to the more conventional use of clip-art images, the current study used action videos to create a rich visual context. Furthermore, these videos showing a recent event (in which an actor performed an action) were shown before participants heard the linguistic utterance. Both of these factors might have had the effect that viewers still chose to more often inspect the recent event rather than use the future cue to quickly focus on the not yet acted-upon object.

In conclusion, consistent with earlier findings, listeners overall preferred to inspect recent-event targets, confirmed by the significant intercept. At the same time, the early morphosyntactic markers affected the strength of the recent event bias and comprehenders decreased the preferential inspections towards the recent target. While in Experiment 1, participants, as a result of tense cue information at the auxiliary verb, started to inspect the future event in the middle of the main verb; in Experiment 2, an earlier effect of morphological marking and verb tense cue caused comprehenders to show preferential inspections of the future target object in the middle of the adverb region. This gaze pattern remained throughout the sentence, as there was a measurable decrease in attention towards the recent event target following adverb offset. These results are discussed in relation to language comprehension within a rich visual context and their combined effect on anticipatory sentence processing.

Acknowledgments

We would like to thank the research assistants who helped with data collection and coding: Prabhjot Multani, Marika Gvianidze, Chris Klammer, Gvantsa Rukhiaashvili, Mtvarisa Shainidze, Megi Abashidze, as well as Prof. Rusudan Asatian, Prof. Marina Giorgadze and Prof. Craig Chambers, who provided their laboratories for conducting the studies.

References

Abashidze, D., Carminati, M. N., & Knoeferle, P. (2019). Anticipating a future versus integrating a recent event? evidence from eye-tracking. *Acta Psychologica*, *200*(102916), 1–16.

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* (pp. 1–6). New Bulgarian University Press.

Abashidze, D., McDonough, K., & Gao, Y. (2020). Exploring the effect of eye gaze cues on novel l2 morphosyntactic pattern learning. *Second Language Research*, 1–14.

Abashidze, D., & Knoeferle, P. (2021). Influence of actor's congruent and incongruent gaze on language processing. *Frontiers in Psychology*, *12*, 701742.

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.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of memory and language*, *59*(4), 457–474.

Brouwer, S., Özkan, D., & Küntay, A. C. (2019). Verb-based prediction during language processing: The case of dutch and turkish. *Journal of Child Language*, *46*(1), 80–97.

Chambers, C. G., & San Juan, V. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition*, *108*(1), 26–50.

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), 687–696.

Covey, L., Gabriele, A., & Fiorentino, R. (2018). Can learners use morphosyntactic cues to facilitate processing? evidence from a study of gender agreement in hindi. *Language Acquisition*, *25*(3), 327–337.

Dink, J. W., & Ferguson, B. (2015). Eyetrackingr: An r library for eye-tracking data analysis. *Retrieved from: http://www.eyetracking-r.com*.

Garrido Rodriguez, G., Norcliffe, E., Brown, P., Huettig, F., & Levinson, S. C. (2023). Anticipatory processing in a verb-initial mayan language: Eye-tracking evidence during sentence comprehension in tseltal. *Cognitive Science*, *47*(1), e13292.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93.

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. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from german and english. *Journal of Psycholinguistic Research*, *32*, 37–55.

Kempe, V. (1999). Processing of morphological and semantic cues in russian and german. *Language and Cognitive Processes*, *14*(2), 129–171.

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*(376), 1–12.

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

Koch, E. M., Bulté, B., Housen, A., & Godfroid, A. (2021). Using verb morphology to predict subject number in l1 and l2 sentence processing: A visual-world eye-tracking experiment. *Journal of the European Second Language Association*, *5*(1).

Rigoli, L., & Spivey, M. J. (2015). Real-time language processing as embodied and embedded in joint action. In *Attention and vision in language processing* (pp. 3–22). Springer.

Rodriguez, A., Burigo, M., & Knoeferle, P. (2015). Visual gender cues elicit agent expectations: Different mismatches in situated language comprehension. *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science (EAPcogsci 2015)*, *1419*, 234–239.

Skopeteas, S., Fanselow, G., & Asatiani, R. (2012). Case inversion in georgian: Syntactic properties and sentence processing. *Case, Word Order and Prominence: Interacting Cues in Language Production and Comprehension*, 145–171.

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.