Detecting Event Construal Shifts in Aspectual Coercion

Uğurcan Vurgun[1], Yue Ji[2], Anna Papafragou[1]

Abstract[1]

Aspectual coercion occurs when there is a semantic mismatch between constituents in terms of their lexical aspect. Despite the long psycholinguistic history of this phenomenon, we currently lack direct measures of how people interpret coerced sentences. We introduce a novel method combining aspectual comprehension with event cognition, allowing us to detect changes in how individuals construe events after reading sentences with varying aspectual information. This study involved two experiments where participants read sentences—either telic or atelic, with or without coercion—followed by a video clip related to the sentence. They assessed if the actor completed the task and identified any brief interruptions during the event, located at the midpoint or late points. The focus was on whether coerced sentences altered participants' event construals, impacting their responses. Results uncovered distinct cognitive responses to aspectual coercion and highlighted differences between coercion types. This method advances our understanding of how lexical aspect influences event representation, offering insights into the nuanced effects of aspectual coercion on cognitive processing and event perception.

Keywords: aspect; aspectual coercion; telicity; boundedness; events; sentence processing

Introduction

Language differentiates between telic predicates, which denote bounded events with an inherent, natural endpoint, and atelic predicates, representing unbounded events without such endpoints (Jackendoff, 1991). For instance, (1a) is a telic sentence, denoting a bounded event that culminates with a complete balloon. In contrast, (1b) is an atelic sentence, indicating an ongoing, unbounded activity without a specific endpoint.

(1) a.The girl drew a balloon. (telic)
b. The girl did some drawing. (atelic)

Telicity is computed compositionally and is influenced by various sentence elements (Bach, 1986). Aspectual coercion arises when there is a semantic mismatch between sentence constituents, necessitating an adjustment in the aspectual classification to align with the sentence's temporal constraints (Jackendoff, 1991). Two subtypes of coercion have been widely studied in the literature. *Subtractive* coercion involves removing an inherent endpoint from a telic phrase and transforming it into an atelic one, such as (2). Conversely, *additive* coercion adds an inherent endpoint to the event construal, as in (3). (On the potential of adverbial modifiers to induce coercion, see Piñango, Zurif, & Jackendoff, 1999; Brennan & Pylkkänen, 2008; Paczynski, Jackendoff, & Kuperberg, 2014; Proctor, Dickey, & Rips, 2004).

(2) The girl drew a balloon …
a. in an hour. (telic, no coercion)
b. for an hour. (telic to atelic coercion)

(3) The girl did some drawing …
a. for an hour. (atelic, no coercion)
b. in an hour. (atelic to telic coercion)

Previous research has suggested that aspectual coercion increases cognitive load, as shown by longer reaction times in tasks combining sentence processing with a secondary task such as lexical decision (Piñango, Winnick, Ullah, & Zurif, 2006). This finding has been supported by subsequent studies (Brennan & Pylkkänen, 2008; Husband, Beretta, & Stockall, 2006), and similar patterns were observed across languages and coercion contexts (Bott, 2010). The assumption that aspectual coercion invariably leads to a shift in interpretation and hence increased cognitive load has been debated. Some studies (Pickering, McElree, Frisson, Chen, & Traxler, 2006) did not find such difficulties with coercion and instead proposed that aspectual representation remains fluid, with a definitive interpretation forming only when necessary (*aspectual underspecification account*; see also Dölling, 2014; Egg, 2020). It is important to note that both of these sets of accounts share the assumption that extra cognitive load is a proxy of commitment shifts which would require the parser to incur additional load due to reanalysis or reinterpretation.

Resolving the debate over how and whether aspectual commitments shift in coercion becomes harder because we currently lack measures that can directly detect event construal shifts. This is because reading times and other tasks offer no direct window onto the kind of event representation formed under non-coerced vs. coerced interpretations (For instance, does it include an endpoint? Or is the situation denoted by the sentence a continuous process without a specific endpoint?). To address this issue, we introduce a new method probing *whether* and *how* aspectual coercion shifts event construals.

[1] University of Pennsylvania, [2]Beijing Institute of Technology

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

A New Paradigm For The Study Of Aspectual Coercion

We build on a recent visual detection task that was devised to study aspectual processing but not coercion per se. Vurgun, Ji, and Papafragou (2022) demonstrated that participants' perception of event boundedness could be influenced by the aspectual nature of sentences they read before viewing a video. Participants were presented with either a telic or an atelic sentence. Afterward, they watched a short video to see whether it matched the sentence. They were also asked to identify if there was any interruption ('glitch') in the video; such interruptions would occur either at the midpoint or at a late point. The interruption detection was an attention probe: when observers were paying more attention to an ongoing event, their detection of irrelevant, external distractors such as a momentary stimulus disruption should be lower (see also Huff, Papenmeier, & Zacks, 2012, for further evidence). The study found that telic sentences led to different rates of interruption detection at mid- and late points in the video, indicative of a bounded event construal: viewers were more likely to detect the interruptions at midpoints and miss those at late points (when viewers' attention is called to the highly salient endpoints of the event itself; cf. also Ji & Papafragou, 2022). However, when speakers were exposed to atelic sentences, there was a uniform interruption detection rate across midpoints and late points, suggesting an unbounded event construal. Thus aspect in language shifted the cognitive interpretation of a dynamic event.

Our study adopts this paradigm but includes coerced sentences among the telic or atelic stimuli presented to participants. In two experiments, we assess whether (and how) coercion influences participants' detection of temporal interruptions in videos, and thus shifts the underlying (bounded vs. unbounded) construal of the depicted events. This is a previously unavailable method of probing aspectual commitments.

Experiment 1

In Experiment 1, we asked whether atelic-to-telic (additive) coercion causes a shift in aspectual commitment - as measured by break detection during event processing - compared to control sentences without coercion.

Participants Using data from a pilot study, we did a power analysis to find the smallest sample size for 80% statistical power with the significance level set at 5%. 177 monolingual English speakers recruited from Prolific participated in Experiment 1. Data from 11 additional adults were collected but excluded from the results due to poor performance with the fillers.

Materials/Procedure Participants were presented with a video of a woman performing different everyday actions with household objects and read the following scenario.

(4) This is Ebony. She has just recovered from orthopedic surgery. Now she needs some extra help with fine motor movements and coordination. Ebony's physical therapist gave her a set of timed exercises with household objects to determine how she is doing now. Your task is to watch the videos and see whether she did the exercise. There may be some glitches in the videos because of Ebony's camera. After each video, please also let us know if you notice a glitch.

In each trial, participants first read a sentence that stated the exercise. There were three between-subjects conditions depending on sentence type. A sample set of sentences is given in (5).

(5) a. Atelic (no ADV, no coercion):
Ebony should do some inflating.
b. Atelic+FOR (no coercion):
Ebony should do some inflating for 10 seconds.
c. Atelic+IN (coercion):
Ebony should do some inflating in 10 seconds.

Each sentence was presented for 6.5 seconds and was followed by a video depicting an actor performing an action. There were 15 videos in total, all with the same actor. These videos had been found in that previous study to depict bounded events, in the absence of any external cues. Following each sentence, participants viewed videos, including both fillers, showcasing actions unrelated to the sentences (n = 6), and test items consistent with the sentences (n = 9). For the test items, the videos were either uninterrupted or modified to include a subtle visual interruption, achieved by removing a single frame. This interruption was either at the midpoint (50% of the total video duration) or at a late point (80% of the duration). Each participant was exposed to an equal number of videos with midpoint or late-point interruptions. The missing frames were experienced, if at all, as a slight disruption of the video. The versions for each original video (no interruption, midpoint interruption, late point interruption) were rotated among participants so that each participant saw only one version of each video (cf. Figure 1).

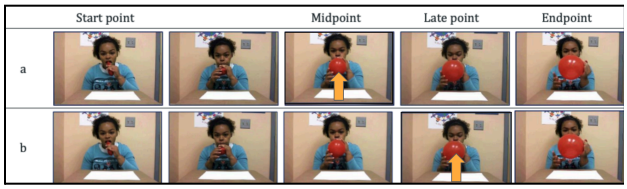

Figure 1: A sample video of a woman inflating a balloon of Experiment 1: (a) a sample video with a 30 ms interruption placed within the 'Midpoint' panel (indicated here by an arrow), (b) a sample video with a late point interruption.

Next, participants had to answer two questions. The first question was about whether the actor had done the exercise (the answer should always be Yes for critical trials, where the sentences for each video were always true of the video, albeit from different aspectual perspectives). The second, most important question was whether there was an interruption (a glitch) in the video. Interruption detection measured event processing and a window onto aspectual (including coerced) construals.

Predictions We expected the transition from atelic to telic meanings to affect participants' interruption detection in videos. We posited that atelic sentences, both in their basic form (no adv, no coercion) and with a non-coercive temporal adverbial (Atelic+FOR), would lead to an unbounded interpretation. This would be evident in uniform interruption detection rates at both midpoints and late points, as there would be no cognitively significant endpoints to focus on. Conversely, in the Atelic+IN condition (coercion), we expected a bounded interpretation due to the adverbial 'in 10 seconds', highlighting the prominence of a specific endpoint. This shift in temporal construal due to the adverbial phrase, despite the verb phrases remaining constant across conditions, should result in decreased interruption detection rates at later stages of the event. These hypotheses are grounded in the established principle that the processing of visual stimuli and the attention allocated to them are influenced by their perceived temporal structure (Ji & Papafragou, 2022; cf. also Huff et al., 2012).

Crucially, we also considered participants' response times to each question as an indicator of cognitive load. While this measure differs from the traditional concept of 'processing cost,' it provides valuable insights into the cognitive efforts required to reconcile aspectual coercion with dynamic visual events. Our analysis, therefore, not only focuses on aspectual commitment shifts but also cognitive engagement during the processing of coerced versus non-coerced sentences. This approach is vital to understanding the complexities of how aspectual coercion impacts cognitive processing.

Results *Verification question*: We analyzed the binary accuracy data with a mixed logit model with fixed effects of Condition (Atelic, Atelic+FOR, Atelic+IN), Interruption (Midpoint, Late point), and their interaction. All factors were coded using centered contrast. Random intercepts were added for participants and items along with random slopes for interruption type and coercion by items. All models were fitted using the *mixed* function of the *afex* package in R. We observed high accuracy across all conditions - Atelic (99%), Atelic+FOR (99%), and Atelic+IN (99%), with no effect of Condition ($\chi^2(2) = 0.0784$, $p = 0.96$). Thus, as expected, all sentences for each video were overwhelmingly accepted as true, regardless of aspectual perspective.

In the RT analysis, outliers in reaction times were excluded using Median Absolute Deviation (MAD). Reaction times were log-transformed, and each participant's MAD and median values for these times were calculated. The dataset was then filtered to remove entries where log-transformed times were more than 2.5 times the MAD from their median. This approach ensures a robust and individualized outlier exclusion. This model revealed a significant effect of Condition ($\chi^2(2) = 50.41$, $p < 0.001$). Post hoc comparisons showed that the Atelic+IN (coercion) condition (mean RT = 3,074 ms) caused significantly longer response times compared to the Atelic+FOR (mean RT = 1,929 ms, $z = 4.049$, $p = 0.0002$) and Atelic (mean RT = 2,463 ms, $z = -7.134$, $p < 0.001$) conditions. These results indicate that participants incurred a larger cognitive load in the coercion condition compared to other conditions in the verification question. Interruption type was not a significant predictor in response times ($\chi^2(1) = 0.0269$, $p = 0.87$). The interaction between Condition and Interruption Type was significant $\chi^2(2) = 11.231$, $p < 0.004$). Following up on the significant interaction, we found that participants in the Atelic+FOR (no coercion) condition responded similarly after videos with Midpoint and Late point breaks (Midpoint = 1,991 ms, Late point = 1,864 ms, odds ratio = 0.936, SE = 0.077, $p = 0.4232$) while participants in the Atelic+IN condition (coercion) had longer response times in videos with Late point interruptions (Late point = 3,387) compared to ones with Midpoint interruptions (Midpoint = 2,754 ms, odds ratio = 1.23, SE = 0.1017, $p = 0.0123$). This result reveals the prominence of endpoints in the coercion (Atelic+IN) condition.

Interruption detection question: The critical results are presented in Figure 2. Here we analyzed the trials in which participants gave correct verification answers. A mixed effects model with random intercepts for participants and items and random slopes for interruption types and coercion revealed an effect of Condition: participants showed different overall accuracy rates in responding to the interruption detection question ($\chi^2(2) = 6.04$, $p = 0.049$).

Participants in the Atelic condition detected midpoint and late point interruptions similarly (odds ratio = 1.080, SE = 0.405, p = 0.8371). Participants in the coercion condition (Atelic+IN) had significantly lower accuracy rates compared to the ones in the no coercion condition (Atelic+FOR, odds ratio = 2.47, SE = 0.946, p = 0.0487). Interruption type (Midpoint vs. Late Point) did not have a significant effect on interruption detection ($\chi^2(1) = 0.0312$, $p = 0.86$). Notably, the interaction between Condition and Interruption was not significant ($\chi^2(2) = 1.337$, $p = 0.512$). This indicates that the aspectual coercion in the sentences did not significantly influence the way participants detected interruptions, as evidenced by the unbounded event construal in the coercion condition (similar detection rates across different time points).

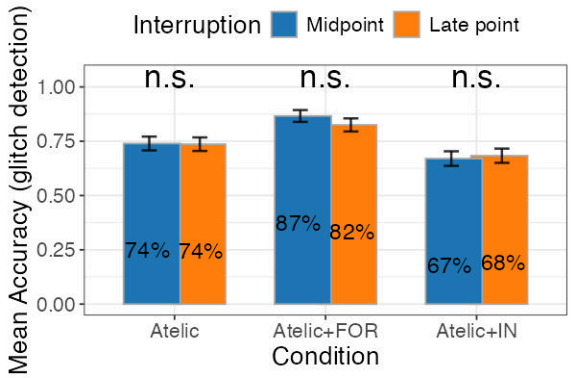

Figure 2: Proportion of correct detections of interruptions in Experiment 1. Error bars represent ± SEM.

The response time analysis for the interruption detection question showed that Condition was not a significant predictor of response times to this question ($\chi^2(2) = 1.71$, $p = 0.43$). Participants in all three conditions answered this question similarly (mean RT (Atelic) = 2,379 ms, mean RT (Atelic+FOR) = 2,386 ms, mean RT (Atelic+IN) = 2,213 ms). Therefore, coercion did not affect how participants answered the interruption detection question.

Discussion Experiment 1 showed that atelic-to-telic coercion did not significantly change participants' aspectual commitments: interruption detection accuracy across different event time points was similar, irrespective of coercion. In other words, participants' event perception patterns were consistent with atelic interpretations of sentences, including those that should be coerced (Atelic+IN). Participants in the coercion condition (Atelic+IN), however, displayed reduced accuracy in detecting interruptions and longer response times to the verification question, suggesting a higher cognitive load from aspectual mismatches in the sentences. Notably, this increased response time was more pronounced with videos featuring late-point interruptions, indicating heightened attention to event endpoints. Yet, this did not translate into shifts in aspectual commitment, as the coerced sentences failed to alter participants' event construals.

Experiment 2

In Experiment 2, we used the same method to test commitments to aspectual coercion in the other direction (telic-to-atelic interpretation, or subtractive coercion).

Participants 175 monolingual English speakers recruited from Prolific participated in Experiment 2. Data from 8 additional adults were collected but excluded from the analyses due to poor performance in the filler trials (Telic = 2, Telic+IN = 3, Telic+FOR = 3).

Materials/Procedure The procedure was identical to Experiment 1 but new sentences were created for each of the three between-subject Condition groups (Telic, Telic+IN, Telic+FOR). A set of sample sentences is given in (4).

(4) a. Telic (no ADV, no coercion):
Ebony should inflate a balloon.
b. Telic+IN (no coercion):
Ebony should inflate a balloon in 10 seconds.
c. Telic+FOR (coercion):
Ebony should inflate a balloon for 10 seconds.

Predictions We hypothesized that shifting from telic to atelic meanings in sentences would affect how participants notice interruptions in videos. We expected that telic sentences, either in their basic form or with a telic adverbial, would create a bounded event interpretation, leading to varying rates of interruption detection between midpoints and late points. This is based on the idea that focusing on visual stimuli

reduces attention to distractions, especially near cognitively significant event endpoints. In contrast, we predicted that coerced atelic sentences would result in an unbounded interpretation, with consistent interruption detection rates throughout the event.

Results *Verification question:* The statistical analysis for Experiment 2 followed the same methodology as Experiment 1. The results indicated high levels of accuracy for the verification question ("Did she do the exercise?") across all conditions: Telic (99.7%), Telic+IN (99.3%), and Telic+FOR (99.3%). The Condition factor was not a significant predictor of accuracy ($\chi^2(2) = 2.08$, $p = 0.3531$). This finding suggests that both the coerced and non-coerced sentences with an adverbial were perceived as congruent with the event, confirming that the same scene could be truthfully described from multiple aspectual perspectives. Response times analysis showed that Condition ($\chi^2(2) = 2.03$, $p = 0.36$), Interruption type ($\chi^2(1) = 0.07$, $p = 0.8$), or their interaction ($\chi^2(2) = 1.64$, $p = 0.44$) were not significant.

Interruption detection question: Here we analyzed the trials in which participants gave correct verification answers. There was no effect of Condition on accuracy rates in this question ($\chi^2(2) = 0.136$, $p = 0.99$). Interruption type (midpoint vs. late-point) was also not significant in predicting interruption detection ($\chi^2(1) = 2.94$, $p = 0.087$). Crucially, the interaction between Condition and Interruption (midpoint vs. late-point) was significant ($\chi^2(2) = 7.1332$, $p = 0.028$). Following up on this interaction, we compared the midpoint and late point detection rates within each Condition to answer our main research question (Figure 3). Surprisingly, participants in the Telic condition detected midpoint interruptions (M = 77%) and late-point interruptions (M = 76%) similarly (odds ratio = 0.952, SE = 0.476, p = 0.92). However, as expected, participants in the Telic+IN condition detected Midpoint (M = 86%) and Late point (M = 71%) interruptions at significantly different rates (odds ratio = 0.272, SE = 0.147, p = 0.0163). Telic+IN interpretations thus point to bounded events that have well-defined endpoints, and attention to these endpoints causes participants to miss irrelevant visual interruptions to the event (see Ji & Papafragou, 2022). Sentences in the Telic+FOR condition, as expected, were aspectually coerced (atelic) sentences, which induced an unbounded event construal with similar detection accuracy for midpoint and late point breaks (Midpoint = 84%, Late point = 73%, odds ratio = 0.361, SE = 0.195, p = 0.06).

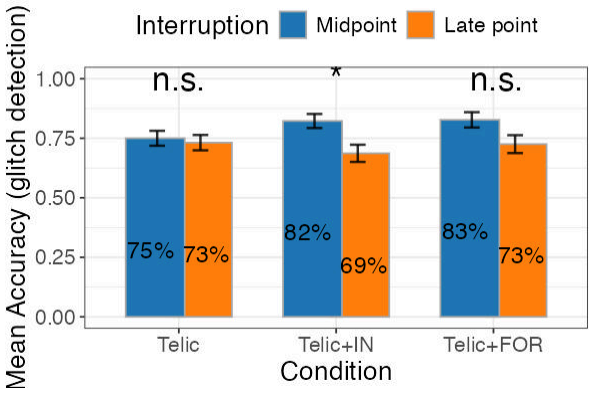

Figure 3: Proportion of correct detections of interruptions in Experiment 2. Error bars represent ± SEM.

Similarly, RT analysis for the interruption question showed that Condition ($\chi^2(2) = 1.92$, $p = 0.38$), Interruption type ($\chi^2(1) = 0.30$, $p = 0.58$), and their interaction ($\chi^2(2) = 1.29$, $p = 0.52$) were not significant.

Discussion Experiment 2 explored whether telic-to-atelic coercion would influence aspectual commitments, as measured by participants' attention to different times within an event. We find that, depending on the presence of an adverbial, what has been considered a telic sentence shifted in terms of participants' aspectual commitments (see Telic+IN, Telic+FOR). Furthermore, these interpretations did not equal increased cognitive load, as response times to questions remained consistent. This finding suggests that aspectual commitment changes are not necessarily linked to heightened cognitive processing demands. Surprisingly, our results showed that in the Telic condition, where sentences lacked adverbial cues, participants detected interruptions similarly at midpoints and late points. That is, comprehenders did not interpret this sentence as truly bounded. This may be because, without explicit adverbials, the inherent endpoint in Telic sentences was less apparent, leading participants to process these events akin to atelic ones, with undifferentiated attention across the timeline. This unexpected finding raises the question of whether the 'coercion' condition in this experiment is an instance of a true interpretive shift (and is consistent with the lack of a cognitive cost for either the IN or FOR adverbial condition). We return to this finding in the General Discussion.

General Discussion

Previous research on aspectual coercion has not offered a direct measure of how aspectual coercion influences the event representations built as comprehenders entertain an aspectual interpretation. In two

experiments, we used a new paradigm that can reliably show how (and whether) aspectual coercion and non-coercion in linguistic input shape ongoing event representations as people compare a sentence to the reference world.

Our results point to two major findings. First, cases of what has traditionally been considered coercion do not always result in shifts in event construals, nor do the shifts coincide with increased processing costs. Specifically, atelic-to-telic (additive) coercion, although associated with a higher cognitive load, does not consistently result in commitment shifts (Experiment 1). Conversely, classic telic-to-atelic (subtractive) coercion can alter aspectual commitments without necessarily increasing cognitive load (Experiment 2; see also Bott, 2010). The lack of increased cognitive load in the subtractive coercion condition in the latter case indicates that commitment shifts do not always occur with increased processing demands. Crucially, our findings question a long-standing assumption in psycholinguistics: that increased cognitive load is a direct indicator of commitment shifts necessitated by reanalysis or reinterpretation of coerced sentences. This assumption, shared by various studies (e.g., Piñango et al., 1999; Brennan & Pylkkänen, 2008; Pickering et al., 2006), is not confirmed by our data.

Second, our findings raise fundamental questions about the computation of telicity. Recall that the basic assumption that telic sentences without adverbials inherently denote bounded events (Egg, 2020) does not mean that participants track the natural endpoint of these events during verification (Experiment 2). This observation suggests that the presence of explicit adverbials may be more crucial for demarcating event boundaries than previously thought, thereby affecting how events are cognitively construed. Our work implies that the link between the traditional understanding of telicity and the verification of telic interpretations in language comprehension requires a more nuanced approach. Our study, while illuminating, also opens up several avenues for further inquiry. Firstly, we consider the suitability of our sentence-to-video verification task for assessing telicity interpretations in natural language use. We believe our approach, focusing on verifying the truthfulness of a sentence, is a valid measure that doesn't distort everyday interpretations (and not an 'additional' task, as in Pickering et al., 2006). Secondly, we aim to explore whether our findings extend beyond the incremental-theme cases (Filip, 2012) used in this study, potentially applying to a broader range of sentence types. Thirdly, we acknowledge that our reaction time analysis reflects later stages of processing and not the immediate comprehension of sentences. Future studies should aim to directly measure processing costs, perhaps through an online test where participants indicate interruptions in videos in real time, offering a more nuanced view of how aspectual coercion affects comprehension. Though several questions remain, our methodology provides a unique perspective on the interplay between coerced aspectual sentences and visual event perception, potentially enriching our understanding of telicity in language processing.

References

Bott, Oliver. 2010. The processing of events, vol. 162. John Benjamins Publishing.

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language*, *106*(2), 132–143.

Egg, M. (2020). Aspectual Composition: "Drinking (A Glass of) Milk". *The Wiley Blackwell Companion to Semantics*, 1-30.

Filip, H. (2012). Lexical aspect. In R. I. Binnich (Ed.), The Oxford Handbook of Tense and Aspect. Oxford.

Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, *20*(7), 848–864.

Husband, E. M., Beretta, A., & Stockall, L. (2006). Aspectual computation: Evidence for immediate commitment. *Talk given at AMLaP*.

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, *41*(1-3), 9–45.

Ji, Y., & Papafragou, A. (2022). Viewers spontaneously represent event temporal structure. In *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. (2006). Underspecification and Aspectual Coercion. *Discourse Processes*, *42*(2), 131–155.

Paczynski, M., Jackendoff, R., & Kuperberg, G. (2014). When events change their nature: The neurocognitive mechanisms underlying aspectual coercion. *Journal of Cognitive Neuroscience*, 26(9), 1905-1917.

Piñango, M. M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax–semantics interface. *Journal of Psycholinguistic Research*, *28*, 395–414.

Piñango, M. M., Winnick, A., Ullah, R., & Zurif, E. (2006). Time-Course of Semantic Composition: The Case of Aspectual Coercion. *Journal of Psycholinguistic Research*, *35*(3), 233–244.

Proctor, A. S., Dickey, M. W., & Rips, L. J. (2004). The time-course and cost of telicity inferences. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 26, No. 26).

Vurgun, U., Ji, Y., & Papafragou, A. (2022). Linguistic aspect constrains event apprehension. Poster at the 44th Annual Meeting of the Cognitive Science Society, Toronto, Canada.