4-1-2018

Stage Salience and Situational Likelihood in the Formation of Situation Models During Sentence Comprehension

David Townsend
Montclair State University

Available online at www.sciencedirect.com

ScienceDirect

Lingua 206 (2018) 1–20

Lingua

www.elsevier.com/locate/lingua

Stage salience and situational likelihood in the formation of situation models during sentence comprehension

David J. Townsend *

Montclair State University, United States

Received 4 July 2017; received in revised form 2 January 2018; accepted 11 January 2018
Available online 17 February 2018

Abstract

Two experiments examined the relation between event structure, situational likelihood and eye fixation time while reading predicate modifiers in isolated sentences. Experiment 1 used activity predicates and preparatory process predicates (*climbed a mountain*), which make salient the process that leads to a culmination. Preparatory process predicates increased first pass time on durative modifiers (*for several years*) and decreased total time on frequency modifiers (e.g., *every year*). Situational likelihood was associated with fixation times on frame modifiers (*last year*) but not with fixation times on durative or frequency modifiers. Experiment 2 used activity predicates and result state predicates (*halted a class*), which make salient the result that follows from a culmination. Result state predicates had no effect on fixation times on durative modifiers and decreased total time on frequency modifiers. Situational likelihood was associated only with total time on durative modifiers. These results demonstrate that readers use the meanings of predicates and modifiers to form an initial model of a sentence and that the likelihood of the reported situation is related to reading time relatively late. The results are discussed in terms of type coercion theory and situation models in sentences and narratives.

Keywords: Eye movements; Event structure; Situational likelihood; Situation model; Sentence processing; Type coercion theory

Two theories have dominated research on temporal interpretations in language processing. Situation model theory (Zwaan, 1999, p. 15) explains how readers use gaps in time and other properties of text as signals of event shifts and revisions of situation models, that is, mental representations of what a text is about (see also Johnson-Laird, 1983; Radvansky, 2012; Van Dijk et al., 1983). Type coercion theory (Jackendoff, 1997; Pustejovsky, 1993; Pylkkänen and McElree, 2006) explains how readers use incompatible meanings between phrases to revise an aspectual interpretation, that is, the when, for how long and how often of an event. The present research demonstrates that the meanings of verbs rapidly elicit a likely representation of what a sentence is about that readers revise according to temporal modifiers and evaluate according to background knowledge. These sentence-level models are the foundation of the spatial-temporal-causal gaps that signal event shifts in narratives.

Situation model theory proposes that readers maintain mental representations of what a text is about (Radvansky, 2012; Zwaan and Radvansky, 1998; Zwaan, 1999). Gaps in time, causal chains, protagonists, goals, and locations lead to

* Correspondence to: Department of Psychology, Montclair State University, Upper Montclair, NJ 07043, United States.
E-mail address: townsendd@montclair.edu.

https://doi.org/10.1016/j.lingua.2018.01.002
0024-3841/

revisions of the model. For example, modifiers such as *a few days later* or *the next morning* signal the end of an event and lead to revision of the model (Radvansky, 2012; Zwaan and Radvansky, 1998). Revising a situation model can increase cognitive effort and reading time (Radvansky and Copeland, 2010; Rinck and Weber, 2003). In some cases, reading time may increase because of the reader's surprise at receiving unexpected information rather than the cognitive effort of revising a model. For example, compared to a foreshadowing context sentence such as *She would have to call it a night soon*, a neutral context sentence such as *She really liked this song* produces longer reading time for a target sentence such as *The next morning she got up* (Pettijohn and Radvansky, 2016). The foreshadowing context sentence apparently activates a script that includes sleeping until morning, increasing the likelihood that the narrative will report that she got up in the morning and reducing reading time for a sentence that reports that event.

Maintaining and revising situation models occurs in sentence processing as well (Johnson-Laird, 1983). Evidence that revisions of situation models occur in sentence processing comes from observations that the initial word of *Before/After the psychologist submitted the manuscript, the journal changed its policy* produces distinct patterns of brain activity (Münte et al., 1998; Zwaan, 1999) and behavior (Clark, 1971; Townsend, 1983; Townsend and Ravelo, 1980). Just as readers keep track of causal–temporal relations between sentences in text, they keep track of these relations between the clauses of isolated sentences (Townsend, 1983; Zwaan, 1999). Moreover, the imperfective form *was delivering* and the perfective form *delivered* produce behavioral differences that suggest interpretations of background information vs. event boundaries respectively in narratives (Magliano and Schleich, 2000) and in isolated sentences (Yap et al., 2009). These studies demonstrate that the effects of processing situation models in text appear as well in isolated sentences.

The situation model of a sentence includes how long an event lasts, whether it has an inherent end, and how often it occurs. This model depends on the event structure of the predicate of the sentence. For example, *crossed a street* and *watched a street* differ in event structure (Smith, 1997). The telic predicate *crossed a street* has an inherent end in which the state of the protagonist changes to being on the other side of the street. We can demonstrate its telicity by noting that *was crossing a street* does not mean that the protagonist definitely crossed a street: while the protagonist was crossing the street, some event may have prevented reaching the other side. The atelic predicate *watched a street* does not have an inherent end. We can demonstrate its atelicity by noting that *was watching a street* means that the protagonist did watch a street regardless of what other events may have occurred (Vendler, 1957). As with perfective form, telic sentences such as *The next morning she got up* move narrative time forward and are likely to produce an event shift because of the temporal gap that *the next morning* denotes (Pettijohn and Radvansky, 2016) or because of the change of state that *she got up* denotes. As with imperfective form, atelic sentences such as *she adjusted the volume* do not move narrative time but instead may describe the setting for upcoming events (Dowty, 1986; Hinrichs, 1986; Madden and Ferretti, 2009). The event structure of predicates contributes to the situation models of sentences and texts.

The preferred model for a sentence can change as the reader recognizes more words and phrases (Johnson-Laird, 1983; Moens and Steedman, 1988). For example, the reader initially may prefer an atelic model of *John walked*. If the phrase that follows *John walked* is *to school*, the preferred model shifts to telic because *to school* establishes a culmination of walking and a change in the protagonist's state. If the phrase that follows *John walked to school* is the durative modifier *for several days*, the preferred model now shifts to an atelic (iterative) interpretation of a series of walk-to-school events. These observations suggest that studying the effects of phrases on reading time in sentences can reveal the mechanisms of forming situation models. A phrase that is consistent with the working situation model confirms the model. One that is inconsistent leads to a revision.

Studies of aspectual interpretation in sentences often explain re-interpretations in terms of ‘‘type coercion’’ (Jackendoff, 1997; Pustejovsky, 1991; Pylkkänen and McElree, 2006). According to type coercion theory, a temporal modifier is a function that takes a predicate as its input (Moens and Steedman, 1988). When the aspectual type of the predicate does not match the type that the modifier requires, the modifier coerces the predicate into the type that the modifier requires. For example, the durative modifier *for several days* requires an atelic predicate as in *watched a street*. As noted earlier, when the predicate is telic as in *walked to school*, the modifier *for several days* coerces *walked to school* into the atelic interpretation of a series of walk-to-school events. Neurological and behavioral studies demonstrate that aspectual incompatibility between predicates and modifiers can increase cognitive effort (Bott, 2010; Brennan and Pylkkänen, 2008; Kuperberg et al., 2010; Paczynski et al., 2014; Pickering et al., 2006; Proctor et al., 2004; Stockall and Husband, 2014; Todorova et al., 2000; Townsend, 2013).

Telic predicates can differ in how salient or accessible they make different stages of an event. The prototypical event has a preparatory process that leads to a culmination that usually corresponds to a change of state (Caudal, 2005; Moens and Steedman, 1988). The preparatory process is salient in *crossed a street*, consisting of walking from one side of the street until just before the last step onto the other side. That last step is the culmination of the preparatory process and it changes the protagonist's state to being on the other side, setting up the possibility of an event shift in the narrative-level situation model.

On the other hand, the result of the culmination is salient in *opened a window* (Caudal, 1999; Pinón, 1999). As with imperfective forms and atelic predicates, predicates in which the result of a culmination is salient makes a durative

unchanging situation readily accessible. Accordingly, result state predicates may be more likely to describe the setting for an upcoming event and less likely to produce an event shift. For example, *opened a window for several hours* may lead the reader to prefer a model that specifies the duration of the state of being open that follows from the culmination of opening a window (Pinón, 1999). As with imperfective form and atelic predicates, a durative result state interpretation does not advance narrative time, and so is less likely to produce a revision of sentence-level or narrative-level situation models.

In addition to verb form and meaning, background knowledge has a role in sentence processing and in narrative processing (Ferretti et al., 2001, 2007, 2009). Activated scripts may lead to expectations of upcoming events and allow inferences that are needed to interpret narratives and sentences. Inferences that rapidly, effortlessly and unconsciously assign interpretations to the ambiguous words in *The pilot put the plane into a stall* are ''implicit'' (Johnson-Laird, 1983). Implicit inferences may occur in aspectual interpretation as well. As noted earlier, *John walked to school for several days* produces an iterative model of a series of walk-to-school events. The frequency modifier *every day* also produces an iterative model: we generally interpret *John walked to school every day* to mean that there were multiple walk-to-school events over some relevant period of time. The iterative interpretation that occurs with *every day* requires an inference about the relevant period of time for the series of walk-to-school events (Filik et al., 2009; Husband and Stockall, 2015; Landman, 2011; Rothstein, 1995). Readers infer that the period of time for a series of walk-to-school events is more likely the time when John was a child rather than his entire life. This inference appears to follow rapidly, effortlessly and unconsciously from activation of an attending-school script.

Readers may also use background knowledge in a slower, more effortful and more conscious way in sentence processing. Note that even though durative and frequency modifiers are similar in producing iterative interpretations of telic predicates, these modifiers differ in meaning and in the processes they induce. Durative modifiers describe the time period of an on-going activity. They require as input an atelic predicate because its lack of a culmination denotes an ongoing activity. On the other hand, frequency modifiers describe how often distinct events occurred. They require as input a telic predicate because its culmination makes it a bounded, distinct and countable event. Unlike durative modifiers, type coercion theory does not predict that reading times on *every day* are longer on telic predicates but rather, that they are longer on atelic predicates. When *every day* appears with an atelic predicate such as *watched a street*, it shifts the interpretation to a series of distinct watch-a-street activities. These watch-a-street activities are distinct because of boundaries that the reader imposes (De Swart, 1998), perhaps because of an explicit inference that is based on a slow, effortful, and conscious search for the protagonist's goal in watching a street (Deo and Pinango, 2011; Deo et al., 2012; Johnson-Laird, 1983). Inferences about the protagonist's goals may help the reader establish boundaries for distinct and countable instances of *watched a street*, satisfying the requirement that a frequency modifier applies to countable events. These observations about frequency modifiers suggest that background knowledge such as scripts has a role in forming situation models of isolated sentences (see also Baggio et al., 2008).

Background knowledge has a role in forming situation models in narratives as well. As noted earlier, recent studies suggest that we can attribute some increases in reading time in narratives to surprise at receiving information that departs from an activated script rather than to the cognitive effort of revising a situation model. Pettijohn and Radvansky (2016) demonstrated the relationship between unexpected information and sentence reading time by assessing unexpectedness with ratings of the expectedness or likelihood of a situation in a narrative. These ratings assess the typicality of the event in an activated script. If unexpectedness increases cognitive effort within sentences because of the surprise it induces, reading time will increase as situational likelihood decreases. If it is unexpectedness that explains the increases in reading time, controlling for situational likelihood will eliminate the effects of the meanings of verbs and modifiers on reading time. If it is revision of a situation model that explains increases in reading time, reading time will depend on interactions between the meanings of verbs and modifiers. These interactions predict that reading times for durative modifiers will increase following telic predicates that make the preparatory process salient, but not following telic predicates that make the result state salient. Moreover, reading times for frequency modifiers will increase more slowly following atelic verbs because of an explicit search for plausible boundaries of a durative event.

1. Experiment 1

One aim of Experiment 1 was to determine whether mismatches between predicates and modifiers increase eye fixation time on the modifier when the predicate makes salient the preparatory process leading up to a culmination. Mismatches between predicate and modifier appear in *climbed a mountain for several days* (telic predicate and durative modifier) and *admired a mountain every day* (atelic predicate and frequency modifier). Mismatches do not appear in *admired a mountain for several days* or *climbed a mountain every day*. If type mismatches increase cognitive effort, fixation times on durative modifiers (*for several days*) will be longer for telic predicates (*climbed a mountain*) than for atelic predicates (*admired a mountain*) and fixation times on frequency modifiers (*every day*) will be longer for atelic predicates than for telic predicates. Moreover, if setting boundaries to a durative event requires an explicit inference based on slow

and effortful search, the effects of predicate telicity on fixation time on frequency modifiers will appear only in total time, not in first pass time. Because no type mismatch occurs with frame modifiers as in *climbed a mountain yesterday* or *admired a mountain yesterday*, type coercion theory predicts no effect of predicate telicity on reading time on frame modifiers.

A second aim of Experiment 1 was to determine whether variation in situational likelihood explains differences in reading time. If situations that are more likely are easier to read, reading time will decrease as situational likelihood increases. If reading time differences are due to situational likelihood rather than to shifts in predicate type, interactions between predicate type and modifier type will not appear when situational likelihood is included as a variable.

1.1. Method

1.1.1. Participants

Forty-eight undergraduate students (19 males, 18–31 years, mean = 19.7 years) from the Psychology subject pool at Montclair State University received course credit for participating. All participants had normal or corrected to normal vision, all were native speakers of English and all were naïve about the purpose of the experiment. The experiment lasted about 40 min.

1.1.2. Materials

Each of 24 sets of two-clause sentences had six conditions depending on Predicate Type and Modifier Type in the initial clause. The sentences for one set appear in Table 1. The complete set of sentences appears in Appendix A.1. The initial predicate was either telic (e.g., *climbed a high mountain*) or atelic (e.g., *admired a high mountain*). The modifier in the initial clause was a frame modifier (e.g., *last year*), a durative modifier (e.g., *for several years*), or a frequency modifier (e.g., *every year*). The combination of Predicate Type and Modifier Type produced six versions of each item: Telic predicate with a durative modifier, a frequency modifier, or a frame modifier, and atelic predicate with a durative modifier, a frequency modifier, or frame modifier.

Telicity. The classification of predicates as telic or atelic depended on independent judgments on Dowty's (1979) *in X time* test. A second set of thirty native speakers of English who received course credit at Montclair State University read pairs of sentences in an on-line test. The materials consisted of 24 pairs of test sentences and one filler pair. The participants' task was to judge which of the two sentences in a pair sounds better. A sample test pair was *The hiker climbed a high mountain in the Rockies in six hours* vs. *The hiker admired a high mountain in the Rockies in six hours*. The filler pair was *John painted a picture in an hour* and *John walked in an hour* (Dowty, 1979). Two lists of materials controlled for the order of presentation of telic vs. atelic sentences within a pair. Participants received an explanation of the *in X time* test with examples and instructions to interpret the *in X time* phrase as referring to how long it took to complete the event. The instructions explained the differences between events with and without a natural end. For each test pair, the sentence that was preferred more often was classified as ''telic''; the sentence that was preferred less often was classified as ''atelic.'' Over the 24 pairs, the set of telic predicates classified in this manner was preferred an average of 79% in the *in X time* test, $p < .0001$ by binomial test.

Word frequency. The verbs for telic and atelic predicates did not differ in mean frequency of occurrence (50,565 and 45,687 respectively) (Davies, 2009), $F < 1$.

Structure and length. The test sentences consisted of syntactic categories in the following order: a subordinating conjunction, a subject phrase, a verb, a direct object (in some cases with a prepositional phrase), an adverbial modifier, and a main clause. The average number of words in the sentences was 17.0. The object of the initial verb was always singular. The predicate consisted of the verb, the object noun phrase and, if present, a prepositional modifier of the object. Excluding an initial space, the average number of characters was 33.1 in telic and atelic predicates, $F < 1$. Excluding an initial space and the comma, the average number of characters was 17.2 in durative modifiers (SD = 1.14), 10.2 in frequency modifiers (SD = 1.14) and 9.5 in frame modifiers (SD = 0.93).

Table 1
Examples of test sentences in Experiment 1.

Telic, frame: *Though the hiker climbed a high mountain in the Rockies last year, she still preferred the Adirondacks.*
Telic, durative: *Though the hiker climbed a high mountain in the Rockies for several years, she still preferred the Adirondacks.*
Telic, frequency: *Though the hiker climbed a high mountain in the Rockies every year, she still preferred the Adirondacks.*
Atelic, frame: *Though the hiker admired a high mountain in the Rockies last year, she still preferred the Adirondacks.*
Atelic, durative: *Though the hiker admired a high mountain in the Rockies for several years, she still preferred the Adirondacks.*
Atelic, frequency: *Though the hiker admired a high mountain in the Rockies every year, she still preferred the Adirondacks.*

Interpretations. An online survey assessed differences in the interpretation of the test items. Two lists of 48 target sentences consisted of the initial clauses of 24 test items each in Experiments 1 and 2. Half of the items from each experiment were classified as telic and half as atelic. One atelic target sentence from Experiment 2 was discarded because of an error in wording. The two lists differed in whether the target sentence contained the telic or atelic member of the predicate. A sample item was *The hiker climbed a high mountain in the Rockies for several years*. Following each target sentence, participants saw four additional sentences that expressed (a) iterative, (b) imperfective, (c) result state and (d) durative interpretations of the target as in

a. The hiker climbed a mountain more than once over several years.
b. The hiker began climbing a mountain and several years later had not completed climbing it.
c. The hiker completed climbing a mountain and it remained climbed for several years.
d. The hiker began climbing a mountain, continued climbing it and stopped after several years.

A third set of thirty-two native English-speaking participants who received course credit at Montclair State University made judgments about the interpretation of the sentences. Their task was to indicate which of the four additional sentences best describes their interpretation of the target sentence. Table 2 shows the mean number of choices by participant for each of the four interpretations. For items from Experiment 1, iterative interpretations were more common by participants for telic predicates than for atelic predicates (6.57 vs. 4.81), $F1$ (1, 31) = 15.7, MSE = 3.46, $p < .001$, $\eta_p^2 = .34$, $F2$ (1, 23) = 3.13, MSE = 11.2, $p = .09$, $\eta_p^2 = .12$. Durative interpretations were more common by items for atelic predicates than for telic predicates (3.28 vs. 2.38), $F1$ (1, 31) = 3.90, MSE = 3.37, $p = .06$, $\eta_p^2 = .11$, $F2$ (1, 23) = 5.98, MSE = 5.30, $p = .02$, $\eta_p^2 = .21$. The mean number of imperfective and result state interpretations did not differ for telic vs. atelic predicates, all ps > .10. For items from Experiment 2, result state interpretations were more common for telic predicates than for atelic predicates (6.22 vs. 3.03), $F1$ (1, 31) = 35.9, MSE = 4.53, $p < .001$, $\eta_p^2 = .54$, $F2$ (1, 23) = 16.2, MSE = 10.2, $p = .001$, $\eta_p^2 = .41$. Durative interpretations were more common for atelic predicates than for telic predicates (3.90 vs. 1.94), $F1$ (1, 31) = 29.7, MSE = 2.96, $p < .001$, $\eta_p^2 = .49$, $F2$ (1, 23) = 29.8, MSE = 5.54, $p < .001$, $\eta_p^2 = .56$. The number of iterative and imperfective interpretations did not differ for telic vs. atelic predicates, all ps > .10.

Acceptability. An online survey assessed differences in the acceptability of the test items. A fourth set of thirty native English-speaking participants who received course credit at Montclair State University made judgments about the acceptability of the sentences. The survey consisted of six lists containing the 24 test items without the conjunction. A sample item was *The hiker admired a high mountain in the Rockies for several years*. Following each sentence, participants saw five options: unacceptable, marginally unacceptable, neutral, marginally acceptable, and acceptable. Their task was to indicate which of the five options best describes the sentence. The instructions defined acceptable as easy to read with a clear interpretation and unacceptable as awkward or difficult to interpret. Scores were assigned to ratings on a scale from 1 to 5 with 1 = unacceptable and 5 = acceptable. Sentences with telic predicates were rated more acceptable than those with atelic predicates for frequency modifiers (M = 3.84 vs. 3.08), $F1$ (1, 29) = 15.3, MSE = 0.577, $p = .001$, $\eta_p^2 = .35$, $F2$ (1, 23) = 7.97, MSE = 0.588, $p = .01$, $\eta_p^2 = .26$, and for frame modifiers (M = 4.30 vs. 3.78), $F1$ (1, 29) = 9.97, MSE = 0.415, $p = .004$, $\eta_p^2 = .26$, $F2$ (1, 23) = 5.60, MSE = 0.591, $p = .03$, $\eta_p^2 = .20$. Predicate telicity had no effect on acceptability ratings for durative modifiers (telic: M = 3.12, atelic: M = 3.40), $F1$ (1, 29) = 1.98, MSE = 0.609, $p > .05$, $\eta_p^2 = .06$, $F2$ (1, 23) = 1.55, MSE = 0.6222, $p > .05$, $\eta_p^2 = .06$.

Situational likelihood. An online survey assessed differences in the situational likelihood of the test items. A fifth set of thirty native English-speaking participants who received course credit at Montclair State University made judgments about the likelihood of events that sentences report. The survey consisted of six lists containing 24 test items without the conjunction. A sample item was *The hiker admired a high mountain in the Rockies for several years*. Following each

Table 2
Mean number of choices of interpretations for test items in Experiments 1 and 2 depending on predicate telicity.

	Interpretation			
	Iterative	Imperfective	Result state	Durative
Experiment 1				
Telic	6.57	1.88	4.03	2.38
Atelic	4.81	1.50	3.41	3.28
Experiment 2				
Telic	2.34	0.94	6.22	1.94
Atelic	2.69	0.94	3.03	3.90

sentence, participants saw five options: very unlikely, unlikely, neutral, likely, and very likely. Their task was to indicate which of the five options best describes the likelihood of the event that the sentence reports. Scores were assigned to ratings on a scale from 1 to 5 with 1 = very unlikely and 5 = very likely. Mean situational likelihood ratings did not differ overall for telic and atelic predicates (3.31 vs. 3.39), both *p*s > .05. Mean situational likelihood ratings were lower for durative modifiers than for frame modifiers (3.04 vs. 3.84), $F1$ (1, 29) = 28.0, MSE = .688, $p < .001$, $\eta_p^2 = .49$, $F2$ (1, 23) = 43.2, MSE = .362, $p < .001$, $\eta_p^2 = .65$, and lower for frequency modifiers than for frame modifiers (3.18 vs. 3.84), $F1$ (1, 29) = 22.2, MSE = .582, $p < .001$, $\eta_p^2 = .43$, $F2$ (1, 23) = 25.2, MSE = .410, $p < .001$, $\eta_p^2 = .52$. Mean situational likelihood ratings did not differ for durative modifiers and frequency modifiers (3.04 vs. 3.18), both *p*s > .05. For durative modifiers, mean situational likelihood ratings were lower by participants for telic predicates than for atelic predicates (2.82 vs. 3.26 respectively), $F1$ (1, 29) = 11.0, MSE = .267, p = .002, $\eta_p^2 = .28$, $F2$ (1, 23) = 3.34, MSE = .701, p = .08, $\eta_p^2 = .13$. For frequency and frame modifiers, situational likelihood ratings did not differ for telic and atelic predicates (for telic and atelic predicates respectively, frequency: 3.13 vs. 3.24; frame: 3.80 vs. 3.88), all *p*s > .05.

The effects of predicate telicity differed for situational likelihood and acceptability. For frame and frequency modifiers, sentences with telic predicates were more acceptable than those with atelic predicates (as noted in the preceding section, frame: 4.30 vs. 3.78; frequency: 3.84 vs. 3.08) but sentences with telic and atelic predicates did not differ in situational likelihood (frame: 3.80 vs. 3.88; frequency: 3.18 vs. 3.2). For durative modifiers, sentences with telic and atelic predicates did not differ in acceptability (3.12 vs, 3.40) but sentences with telic predicates were rated as less situationally likely than those atelic predicates by participants (2.83 vs. 3.86). These patterns indicate that situational likelihood and acceptability are different properties.

1.1.3. Procedure

The experiment was conducted with an SR Research Eye Link 1000 desktop eye-tracking system and a View Sonic monitor. The font was fixed-width Monaco 16. The screen width was 160 characters with a resolution of 1280 by 1024. The monitor was 28 in from the participant producing a visual angle of 0.8°. Each participant rested his or her chin and forehead on bars. The system was calibrated with a 9-point grid for right eye tracking with corneal reflection. Maximum drift error was set at 0.4 degrees at the beginning of an experimental session and checked several times during each session. A trial began when the participant focused on a circle in the center of the screen. When the participant's gaze was stable, the Eye Track software presented a square near the left edge of the screen. (The source of all referenced software is http://www.umass.edu/eyelab/software/.) When the participant fixated on the square, the software presented the sentence. Each test sentence had a line break after the comma at the end of the modifier. The second line was one inch below the first. Participants were instructed to read each sentence normally. When they reached the end of the sentence, participants shifted their gaze to a sequence of XXX one inch below and 3–5 spaces to the right of the period. They then pressed a button on the left side of a game controller. This button press presented a comprehension question (see below). Participants answered the question by pressing the top (‘‘yes’’) or bottom (‘‘no’’) button on the game controller. Each participant read 24 test sentences interspersed among 80 filler sentences. The test sentences had the form of those in Table 1. Each of six lists had four test sentences in each of six conditions: 2 (Predicate Type: telic, atelic) × 3 (Modifier Type: durative, frequency, frame). Eight participants were assigned to each list.

Comprehension questions. A comprehension question appeared after each sentence to encourage participants to attend to the meaning of the sentence. A thematic role question followed half of the items and an aspectual question followed the other half of the items. Thematic role questions asked about the patient (e.g., *What did the hiker climb*? A Rocky Mountain vs. An Adirondack Mountain). Aspectual questions asked whether the sentence reported one vs. multiple events (e.g., *How many times did Howard definitely send something*? Just one vs. More than one). Question type varied between items so that a particular item always received the same question.

1.1.4. Data analysis

The eye movement software enabled manual adjustment of vertical displacement. The software combined fixations shorter than 80 ms. It excluded trials in which any single gaze duration exceeded 2000 ms (2.0% of the trials).

The regions of analysis were the predicate and the adverbial modifier. The predicate region consisted of the verb, the object and any prepositional phrase that preceded the adverbial modifier as in *climbed a high mountain in the Rockies*. The modifier region consisted of the adverbial phrase as in *last year*. The regions of analysis included the spaces before words and the comma after the adverbial modifier.

The data consisted of *first pass time* and *total time*. The first pass time in a region is the sum of fixation times from the first fixation when entering the region from the left until leaving the region. The total time in a region is the sum of all fixation times in the region including first pass fixations and fixations that return to the region. To control for variability in the length of regions, statistical analysis of fixation time differences used residual reading time (Ferreira and Clifton, 1986; Trueswell et al., 1994). Residual reading time is the difference between the actual fixation time on a region and the fixation time that

linear regression predicts from the number of characters in the region. Linear mixed effects analysis of the relation between residual reading time and telicity and situational likelihood was performed with R (Foundation, 2014) and *lme4* (Bates et al., 2012; Kuznetsova et al., 2015; Pinheiro and Bates, 2006).

1.2. Results and discussion

The modifier region was fixated on at least 95% of the trials for each condition.

Fixation times more than 2.5 standard deviations from the mean were trimmed. Telicity and situational likelihood were entered as fixed effects with subjects and items as random intercepts. For each combination of Modifier Type and Region, models with and without telicity or situational likelihood were compared using likelihood ratio tests. The best fitting random intercepts models were compared to random slopes models by-subjects and by-items. Because these comparisons showed no advantage for models with random slopes, only models using random intercepts are reported. Random intercept models that included a term for the interaction between telicity and situational likelihood were compared with models that had telicity and situational likelihood as fixed main effects. Because these comparisons showed no advantage for models with an interaction term, models with an interaction term are not reported.

Table 3 shows slopes, χ^2, degrees of freedom and p values for likelihood ratio tests of the relation between telicity and residual reading time and between situational likelihood and residual reading time. Values of $p < .05$ are reported as significant. Mean raw times (ms) and residual reading times on frame, durative and frequency modifiers appear in Appendices B.1–B.3.

Frame modifiers. Frame modifiers serve as a control for the effects of predicate–modifier mismatches in forming situation models with durative and frequency modifiers. Table 3 shows no effect of telicity on either measure of fixation time for sentences with frame modifiers. The effect of situational likelihood on residual first pass time on the modifier was significant, $\chi^2(1) = 4.575$, $p = .03$, with greater situational likelihood associated with faster residual first pass time (slope = -55.35 ± 26.30 standard errors), indicating an early relation between situational likelihood and reading time on the modifier. Fig. 1 shows the relation between situational likelihood and first pass time on frame modifiers. The effect of situational likelihood on residual total time on the modifier was also significant, $\chi^2(1) = 6.5215$, $p = .01$, with greater situational likelihood associated with faster residual total time (slope = -115.04 ± 42.54 standard errors). Fig. 2 shows

Table 3
Summary statistics for likelihood ratio tests of the effects of telicity and situational likelihood in Experiment 1.

	Predicate				Modifier			
	Slope	χ^2	*df*	*p*	Slope	χ^2	*df*	*p*
Frame modifiers								
Telicity								
First pass time	−10.91	0.0758	1	.78	13.907	0.5879	1	.44
Total time	−110.264	0.5516	1	.46	−22.95	0.0135	1	.91
Situational likelihood								
First pass time	65.88	1.1954	1	.27	−55.35	4.575	1	.03*
Total time	162.41	1.9959	1	.16	−115.04	6.5215	1	.01*
Durative modifiers								
Telicity								
First pass time	−17.35	0.1619	1	.69	65.48	6.3451	1	.01*
Total time	−91.81	0.2499	1	.62	−4.057	0.007	1	.93
Situational likelihood								
First pass time	111.68	3.4442	1	.06	−15.81	1.3202	1	.25
Total time	−96.07	0.6242	1	.43	3.404	0.0076	1	.93
Frequency modifiers								
Telicity								
First pass time	126.11	2.7224	1	.10	−13.88	0	1	1.0
Total time	−391.32	9.0887	1	.003**	−139.53	8.6599	1	.003**
Situational likelihood								
First pass time	−4.44	0.2746	1	.60	−45.79	3.6183	1	.06
Total time	17.63	0.1306	1	.72	−40.89	0.9015	1	.34

* $p < .05$.
** $p < .01$

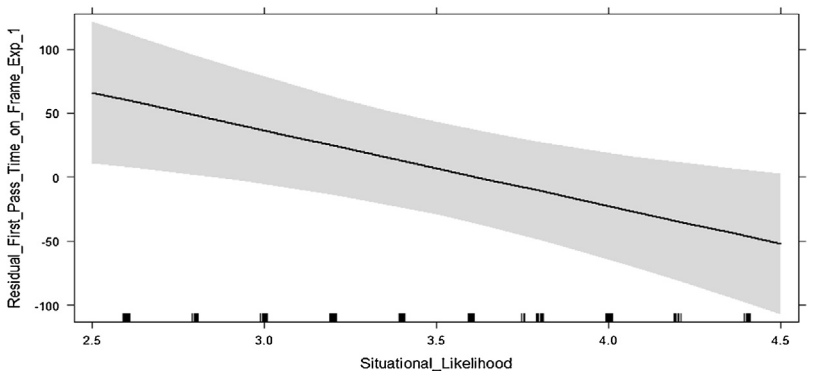

Fig. 1. Mean residual first pass time on frame modifiers depending on situational likelihood in Experiment 1.

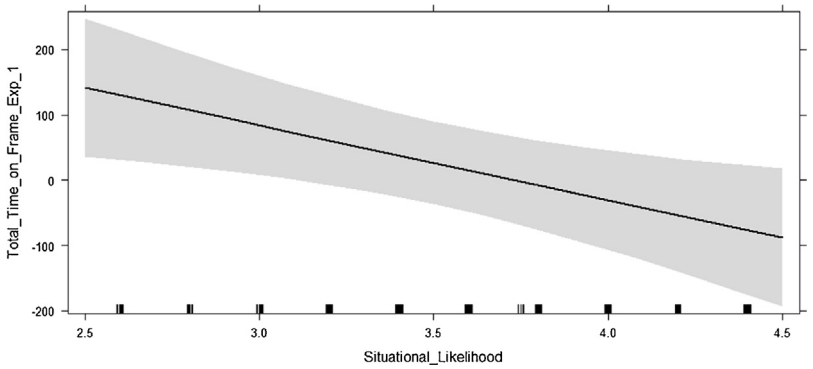

Fig. 2. Mean residual total time on frame modifiers depending on situational likelihood in Experiment 1.

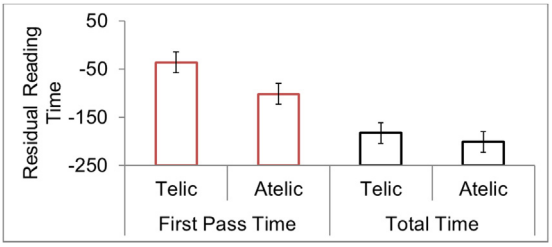

Fig. 3. Mean residual first pass time and total time on durative modifiers depending on predicate telicity in Experiment 1.

the relation between situational likelihood and total time on frame modifiers. No other effects approached significance. (See Appendix B.1 for raw and residual reading times.)

Durative modifiers. Table 3 shows that the only significant effect was the effect of telicity on residual first pass time on the modifier, with telic predicates increasing residual first pass time by 65.48 $\pm$ 30.74 standard errors, $\chi^2(1) = 6.3451$, $p = .01$. Fig. 3 shows mean residual first pass time and total time on durative modifiers. In a linear mixed model fit by a maximum likelihood test (Satterthwaite approximation to degrees of freedom) and situational likelihood and telicity as fixed factors, the effect of telicity on residual first pass time remained significant, $t(46.68) = 2.578$, $p = .01$, $\eta_p^2 = .72$, indicating that the reader does not use situational likelihood to resolve a semantic mismatch. Moreover, the interaction

between situational likelihood and telicity on residual first pass time was not significant, $\chi^2(1) = 0.8404$, $p = .36$. (See Appendix B.2 for raw and residual reading times.)

Frequency modifiers. Table 3 shows that the only significant effects of telicity for frequency modifiers appeared in total time. Atelic predicates increased residual total time on the predicate by 391.32 ± 122.85 standard errors, $\chi^2(1) = 9.0887$, $p = .003$, and on the modifier by 139.53 ± 45.01 standard errors, $\chi^2(1) = 8.6599$, $p = .003$. Fig. 4 shows mean first pass time and total time for frequency modifiers. The telicity effects on total time and not on first pass time indicate a late increase in reading time. In a linear mixed model fit by a maximum likelihood *t*-test (Satterthwaite approximation to degrees of freedom) with situational likelihood and telicity as fixed factors, the effect of telicity on residual total time on the modifier remained significant, $t(49.03) = 3.056$, $p = .004$, $\eta_p^2 = .75$. The only other effect that approached significance was a marginal effect of situational likelihood for first pass time on the modifier, $\chi^2(1) = 3.6183$, $p = .06$, with greater situational likelihood associated with faster residual first pass time (slope = -45.79 ± 22.90 standard errors). Fig. 5 shows the relation between situational likelihood and first pass time on frequency modifiers. This result suggests an early attempt to integrate an initial situation model with situational likelihood. (See Appendix B.3 for raw and residual reading times.)

The results of Experiment 1 demonstrate that when telic predicates have a salient preparatory process, first pass time on a durative modifier is longer for telic predicates than for atelic predicates, and total time on a frequency modifier is longer for atelic predicates than for telic predicates. Both results support type coercion theory. Neither situational likelihood nor sentence acceptability can explain these differences in reading time.

Situational likelihood was strongly related to reading time for frame modifiers but only marginally for frequency modifiers and not at all for durative modifiers. These results indicate that readers use the situational likelihood of a sentence primarily in the absence of a predicate–modifier mismatch. The lack of a relation between situational likelihood and reading time on durative modifiers and on frequency modifiers suggests that readers do not use situational likelihood to resolve type mismatches.

The telicity effects on first pass time on durative modifiers appear unrelated to ratings of sentence acceptability. For durative modifiers, predicate telicity was unrelated to acceptability, yet first pass time was longer for telic predicates, demonstrating differences in reading time that we cannot attribute to acceptability. For frequency and frame modifiers, sentences were more acceptable with telic predicates, yet first pass time was unrelated to predicate telicity, demonstrating

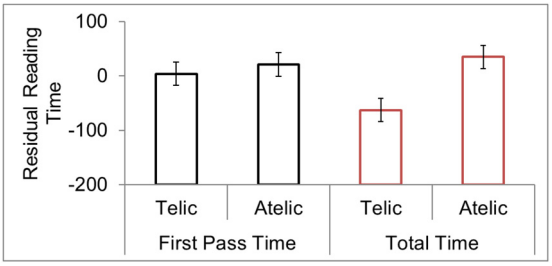

Fig. 4. Mean residual first pass time and total time on frequency modifiers depending on predicate telicity in Experiment 1.

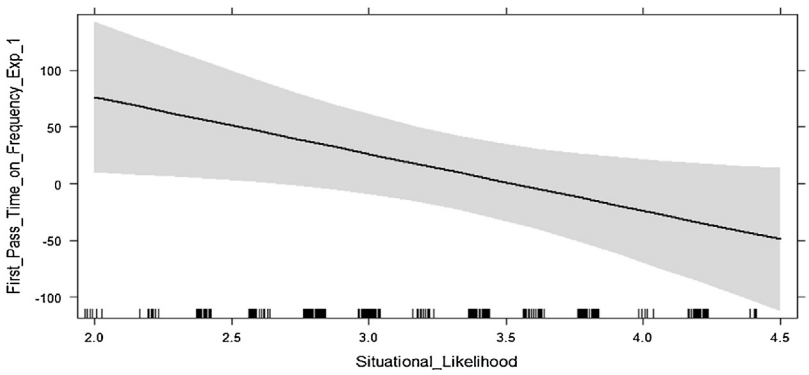

Fig. 5. Mean residual first pass time on frequency modifiers depending on situational likelihood in Experiment 1.

that differences in acceptability do not produce differences in reading time. On the other hand, differences in sentence acceptability were related to total time on frequency modifiers: sentences with telic predicates were rated more acceptable and they produced shorter total time. One interpretation of this relation between acceptability, predicate telicity and total time is that participants judge sentences with atelic predicates and frequency modifiers less acceptable because of a slow, effortful search for potential boundaries of the durative activity that the atelic predicate denotes.

Experiment 1 therefore demonstrates that violations of the telicity requirements of predicate modifiers affect reading time. In the case of frequency modifiers, the cognitive cost of searching for potential boundaries of durative situations may increase reading time more slowly.

2. Experiment 2

The aim of Experiment 2 was to determine whether predicates that make salient the state that follows from a culmination increase reading time for durative modifiers. Type mismatches between predicate and modifier appear in *stepped into a pool for an hour* and *watched a pool every hour.* If type mismatches increase cognitive effort, fixation times on durative modifiers (*for an hour*) will be longer for telic predicates (*stepped into a pool*) than for atelic predicates (*watched a pool*). Fixation times will also be longer on frequency modifiers (*every hour*) when the predicate is atelic than when it is telic. Moreover, if setting boundaries to a durative event requires an inference based on slow and effortful search, the fixation time effects on frequency modifiers will appear in total time rather than in first pass time.

As in Experiment 1, a second aim of Experiment 2 was to determine whether variation in situational likelihood explains differences in fixation time. If situations that are more likely are easier to read, fixation time will decrease as situational likelihood increases. If reading time is related to situational likelihood rather than the cognitive effort of shifting the aspectual type of the predicate, interactions between predicate type and modifier type will not appear when situational likelihood is included as a variable.

2.1. Method

2.1.1. Participants

Forty undergraduate students (8 male) from the Psychology subject pool at Montclair State University received course credit for participating. All participants had normal or corrected to normal vision, all were native speakers of English and all were naïve about the purpose of the experiment. The experiment lasted about 40 min.

2.1.2. Materials

Each of 24 sets of two-clause sentences had four conditions depending on Predicate Type and Modifier Type in the initial clause. The sentences for one set appear in Table 4. The complete set of sentences appears in Appendix A.2. The initial clause contained a predicate that was either telic (e.g., *halted a class*) or atelic (e.g., *attended a class*). Telic and atelic predicates used result state and activity verbs respectively. Telic result state verbs were drawn from Caudal (2005), Levin (1993), and Pinón (1999). The initial clause contained a durative modifier (e.g., *for several years*) or a frequency modifier (e.g., *every year*). The combination of Predicate Type and Modifier Type produced four versions of each item: Telic predicate with a durative modifier or a frequency modifier, and atelic predicate with a durative modifier or a frequency modifier.

Word frequency. The verbs for telic and atelic predicates did not differ in mean frequency of occurrence (21,025 and 22,749 respectively) (Davies, 2009), $F < 1$.

Structure and length. The test sentences consisted of syntactic categories in the following order: a subject phrase, a verb, a direct object, a prepositional phrase, an adverbial modifier, and a subordinate clause (see Table 4). The average number of words in the sentences was 17.1. The object of the initial verb was always singular. The predicate consisted of the verb, the object noun phrase and a prepositional modifier of the object. Excluding an initial space, the average number

Table 4
Examples of test sentences in Experiment 2.

Telic, durative: *The dean halted a class in Dickson Hall for an hour if there was a fire drill.*
Telic, frequency: *The dean halted a class in Dickson Hall every hour if there was a fire drill.*
Atelic, durative: *The dean attended a class in Dickson Hall for an hour if there was a fire drill.*
Atelic, frequency: *The dean attended a class in Dickson Hall every hour if there was a fire drill.*

of characters was 32.5 in telic predicates (SD = 4.4), 33.8 in atelic predicates (SD = 4.6), 12.6 in durative modifiers (SD = 3.1), and 11.2 in frequency modifiers (SD = 1.9).

Interpretations. The survey reported in the method section of Experiment 1 showed that for the materials in Experiment 2, result state interpretations are more common for telic (result state) predicates than for atelic predicates and that durative interpretations are more common for atelic predicates than for telic predicates. (See Experiment 1, Method, Interpretations.)

Acceptability. An online survey assessed differences in the acceptability of the test items. A second set of 24 native English-speaking participants who completed a course requirement at Montclair State University made judgments about the acceptability of the sentences. The survey consisted of four lists containing the initial clause of 24 test items. A sample item was *The dean halted a class in Dickson Hall for an hour.* The procedure was identical to that in Experiment 1 (see Experiment 1, Method, Acceptability). Participants rated sentences more acceptable with activity verbs rather than result state verbs for durative modifiers (M = 3.68 vs. 2.90), $F1$ (1, 23) = 26.5, MSE = 0.274, $p < .001$, $\eta^2_{partial} = .54$, $F2$ (1, 23) = 8.38, MSE = 0.868, p = .008, $\eta^2_{partial} = .27$, and frequency modifiers (3.74 vs. 3.48), $F1$ (1, 23) = 4.19, MSE = 0.189, p = .05, $\eta^2_{partial} = .15$, $F2$ (1, 23) = 5.60, MSE = 0.191, p = .03, $\eta^2_{partial} = .20$.

Situational likelihood. An online survey assessed differences in the situational likelihood of the test items. A third set of 28 native English-speaking participants who completed a course requirement at Montclair State University made judgments about the likelihood of events that sentences report. The survey consisted of four lists containing 24 test items without the conjunction. A sample item was *The dean halted a class in Dickson Hall for an hour.* The procedure was identical to that in Experiment 1 (see Experiment 1, Method, Situational Likelihood). Mean situational likelihood ratings did not differ for telic and atelic predicates (3.29 vs. 3.32) or for durative modifiers and frequency modifiers (3.31 vs. 3.30), and the interaction was not significant, all Fs < 1. The effects of predicate telicity differed for situational likelihood and acceptability. For both frequency modifiers and durative modifiers, sentences with telic and atelic predicates did not differ in situational likelihood but sentences with atelic predicates were more acceptable than those with telic predicates (see preceding section).

2.1.3. Procedure

Each participant read 24 test sentences and 101 filler sentences. Each of four lists had six test sentences in each of four conditions: 2 (Predicate Type: telic, atelic) × 2 (Modifier Type: durative, frequency). Twelve participants were assigned to each list. All other details of the procedure were identical to those in Experiment 1.

2.1.4. Data analysis

The eye movement software enabled manual adjustment of vertical displacement (http://www.umass.edu/eyelab/software/). The software combined fixations shorter than 80 ms. It excluded trials in which any single gaze duration exceeded 2000 ms (0.4% of the trials). The regions of analysis were the predicate and the adverbial modifier. The predicate region consisted of the verb, the object and the prepositional phrase that preceded the adverbial modifier as in *halted a class in Dickson Hall for an hour*. The modifier region consisted of the adverbial phrase as in *for an hour*. The regions of analysis included the spaces before words and the comma after the adverbial modifier. Experiment 2 collected first pass time and total time as in Experiment 1.

2.2. Results and discussion

The modifier region was fixated on at least 98% of the trials for all conditions.

Statistical analysis followed the same procedures as in Experiment 1. Table 5 shows slopes, χ^2, degrees of freedom and p values for likelihood ratio tests of the relation between telicity and residual fixation time and between situational likelihood and residual fixation time. Values of $p < .05$ are reported as significant. Mean raw and residual reading times on durative and frequency modifiers appear in Appendices B.4 and B.5.

Durative modifiers. Table 5 shows no effect of telicity for durative modifiers. Fig. 6 shows mean first pass time and total time on durative modifiers. The only significant effect for sentences with durative modifiers was the relation between situational likelihood and residual total time on the modifier, $\chi^2(1) = 4.5703$, p = .03, with greater situational likelihood associated with faster residual total time (slope = -129.2 ± 58.51 standard errors). Fig. 7 shows the relation between situational likelihood and total time on durative modifiers. (See Appendix B.4 for raw and residual reading times.)

Frequency modifiers. Table 5 shows a significant effect of telicity on total time on the modifier, $\chi^2(1) = 5.5142$, p = .02, where atelic predicates increased residual total time by 118.17 ± 48.29 standard errors. Fig. 8 shows mean first pass time and total time on frequency modifiers. No other effects for frequency modifiers approached significance. The significant effect of telicity for total time and the lack of an effect of telicity for first pass time suggests that processing difficulty occurs relatively late in reading. (See Appendix B.5 for raw and residual reading times.)

The results of Experiment 2 demonstrate that when a telic predicate has a salient result state, predicate telicity has no effect on first pass time on a durative modifier. This result does not support the prediction from type coercion theory that a

Table 5
Summary statistics for likelihood ratio tests of the effects of telicity and situational likelihood in Experiment 2.

	Predicate				Modifier			
	Slope	χ^2	*df*	*p*	Slope	χ^2	*df*	*p*
Durative modifiers								
Telicity								
First pass time	3.433	0.0049	1	.94	−8.79	0.1375	1	.71
Total time	75.26	0.417	1	.52	40.54	0.2729	1	.60
Situational likelihood								
First pass time	21.25	0.1428	1	.71	15.90	0.2333	1	.63
Total time	−174.55	3.464	1	.06	−129.2	4.5703	1	.03*
Frequency modifiers								
Telicity								
First pass time	−11.06	0.021	1	.88	−2.417	0.0071	1	.93
Total time	0.7295	0	1	.99	−118.17	5.5142	1	.019*
Situational likelihood								
First pass time	30.16	0.4575	1	.50	13.49	0.8282	1	.36
Total time	−0.4266	0	1	1.0	51.02	1.6898	1	.19

* $p < .05$.

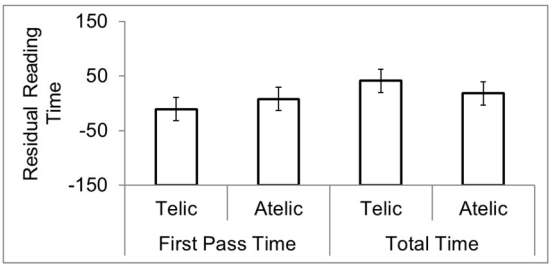

Fig. 6. Mean residual first pass time and total time on durative modifiers depending on predicate telicity in Experiment 2.

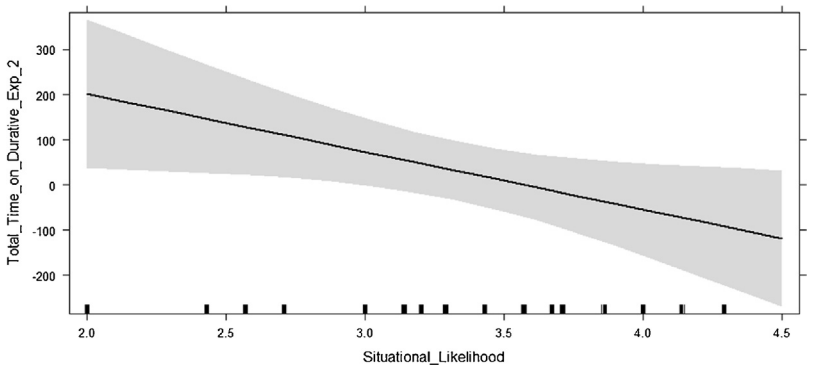

Fig. 7. Mean residual total time on durative modifiers depending on situational likelihood in Experiment 2.

durative modifier of a telic predicate increases reading time. As in Experiment 1, total time on a frequency modifier was longer for atelic predicates than for telic predicates. This result is consistent with a slow, effortful search for potential boundaries of the durative activity that the atelic predicate denotes.

Neither situational likelihood nor sentence acceptability can explain the patterns of reading time. Situational likelihood was related to reading time only in total time on durative modifiers, suggesting that readers primarily use situational likelihood after forming an initial model of a durative result state. Differences in sentence acceptability cannot explain the effect of predicate telicity on total time on frequency modifiers: sentences with atelic predicates were more acceptable than those with telic predicates, yet total time on frequency modifiers was longer for atelic predicates, suggesting the surprising conclusion that more acceptable sentences take longer to read.

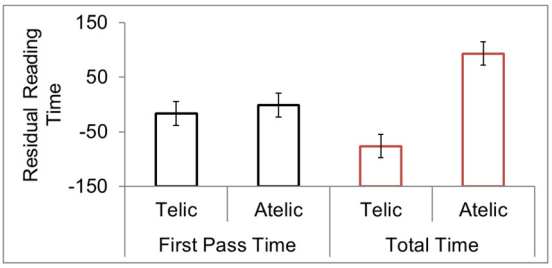

Fig. 8. Mean residual first pass time and total time on frequency modifiers depending on predicate telicity in Experiment 2.

The conclusion from Experiment 2 is that readers are flexible in using different senses of telic predicates depending on the modifier. With durative modifiers of result state predicates, the reader readily adopts a durative model. With frequency modifiers of result state predicates, the reader readily adopts a telic model.

3. General discussion

The results demonstrate that reading time for predicate modifiers is related to the meanings of predicates and modifiers and to the likelihood of the situation that the sentence describes. Mismatches between the meanings of predicates and their modifiers increased reading time depending on the salience of different stages of the event that the predicate denotes. Situational likelihood was related to reading time primarily in the absence of predicate–modifier mismatches. These results demonstrate that readers form and revise situation models of sentences based on the meanings of predicates and modifiers. This section summarizes and interprets the results regarding predicate–modifier mismatches and situational likelihood, and the relation between processing situation models in narratives and in sentences.

3.1. Type coercion theory and stage salience

The event structure of predicates can affect reading time on durative modifiers, confirming previous studies (Brennan and Pylkkänen, 2008; Paczynski et al., 2014; Pinango et al., 1999, 2006; Proctor et al., 2004; Todorova et al., 2000; Townsend, 2013). The present results also show that the distinction between telic and atelic predicates is not sufficient to explain reading time: reading time depends on the salience of different event stages of a telic predicate, suggesting a modification of type coercion theory. Telic predicates increased first pass time on durative modifiers when the preparatory process was salient but not when the result state was salient. This result suggests that on durative modifiers of preparatory process predicates, readers shift the interpretation of the predicate from telic to atelic by interpreting it as a series of distinct events or possibly as an incomplete event in which the culmination did not occur. This shift requires cognitive effort and it increases reading time. The absence of a telicity effect on durative modifiers of result state predicates suggests that when the result state is salient, readers quickly adopt a result state model, making the interpretation of the predicate compatible with the durative modifier and avoiding a costly shift to an iterative or imperfective interpretation.

The meanings of predicates affect reading time more slowly for frequency modifiers than for durative modifiers. Whereas telic predicates with a salient preparatory process increased first pass time on durative modifiers, predicate telicity had no effect on first pass time on frequency modifiers and atelic predicates increased total time on frequency modifiers. These patterns clarify the processes that occur in shifting from an atelic to a telic model. Because frequency modifiers count occasions of an event, longer total time for atelic predicates suggests that the combination of an atelic predicate and a frequency modifier induces a search for potential boundaries that could mark distinct and countable occasions of the activity. The effects of type shifts may appear later on frequency modifiers than on durative modifiers because a frequency modifier of an atelic predicate requires this slow effortful search for potential boundaries. Because interpretation with a durative modifier does not require a search for potential boundaries, total time on durative modifiers does not differ for telic and atelic predicates. These results suggest a further modification of type coercion theory: aspectual type shifts are not always an immediate response to a semantic incompatibility between predicate and modifier, but may instead involve slower use of knowledge from outside the language system.

The results suggest an explanation for previous inconsistencies in the evidence for type shifts. Previous studies have examined processing of particular predicate types such as semelfactives (e.g., *sneezed*), achievements (e.g., *sent*), or accomplishments (e.g., *consumed*) without systematic investigation of the interactions between the various possible meanings of predicates and modifiers (Brennan and Pylkkänen, 2008; Pickering et al., 2006; Proctor et al., 2004;

Todorova et al., 2000; Townsend, 2013). The present research demonstrates that the effects of predicates depend on an interaction between the salience of different stages of an event and the meanings of modifiers.

3.2. Situational likelihood and expectedness

The likelihood of the situation in an activated script cannot explain completely the variations in reading time. As situational likelihood increased, first pass time and total time on frame modifiers decreased, showing that the ratings of situational likelihood did assess knowledge that people use when reading frame modifiers. On the other hand, reading time for mismatching combinations of predicate and modifier was unrelated to situational likelihood. Even though telic predicates with a salient preparatory process increased first pass time on a durative modifier, first pass time in these cases was unrelated to situational likelihood. Similarly, atelic predicates increased total time on frequency modifiers, but total time in these cases was unrelated to situational likelihood. Variation in the likelihood of the event that the sentence reports cannot explain the effects of predicate–modifier mismatches.

A surprising result was a relation between situational likelihood and total time on durative modifiers of result state predicates and atelic predicates. Combined with the absence of a relation between situational likelihood and first pass time on durative modifiers of result state predicates, the relation between situational likelihood and total time suggests that readers use situational likelihood after adopting a result state model. A mechanism that explains this pattern is that readers initially adopt a durative/result state model for both atelic predicates and result state predicates. When re-reading the modifier, readers evaluate their durative/result state model according to its likelihood in the world. Evaluating the durative/result state model is easier if it is more typical in the activated script. Thus, word meanings appear to drive the formation of models. Readers evaluate these models in terms of their likelihood in the world.

At first glance, the absence of a relation between situational likelihood and reading time when the sentence contains a predicate–modifier mismatch seems to contradict demonstrations that expectedness in a narrative can explain reading time for event shifts (Pettijohn and Radvansky, 2016). The difference in results may be due to the difference in situational likelihood and expectedness. In the present study situational likelihood refers to the typicality of the reported event in the script that the sentence itself activates. Readers cannot judge situational likelihood until they have a model for the sentence. Because they base this model on the meanings of predicates and modifiers, the relations to reading time appear earlier for predicate–modifier meanings than for situational likelihood. On the other hand, expectedness refers to the typicality of the reported event in the script that preceding sentences in the narrative activate (Pettijohn and Radvansky, 2016). Readers have access to expectedness prior to reading the target sentence. Thus, one reason why the relation between situational likelihood/expectedness and reading time is smaller in isolated sentences than in narratives may be that narratives have more context that can activate a script, producing expectation of a particular event prior to reading a sentence that reports that event. A related explanation is that readers direct attention differently in isolated sentences and narratives (Townsend and Bever, 1988). When reading a sentence that appears in isolation or that is unexpected in a narrative, people focus attention more on words, phrases and other sentence-level properties. Reading a sentence that is highly expected in a narrative draws people's attention away from sentence-level properties. Nevertheless, sentence-level properties still affect processing in narratives (Townsend and Bever, 1991, 2001).

3.3. Situation models in narratives and sentences

The processes that occur in narratives and isolated sentences overlap. When reading narratives, people form situation models. These narrative-level models represent what the narrative is about. Readers form narrative-level models using information from the sentences in the narrative and from the script that the narrative activates. They use gaps in temporal relations, causal relations, protagonists, goals and locations to identify boundaries of major events and to revise narrative-level models. Forming a new narrative-level model can have a cognitive cost.

When reading isolated sentences, people form situation models. These sentence-level models represent what the sentence is about. Readers form sentence-level models using information from the words and phrases in the sentence and from the script that the sentence activates. They use aspectual information in predicates and modifiers to identify boundaries of events and to revise sentence-level models. Forming a new sentence-level model can have a cognitive cost.

The benefit of this analogy between situation models in narratives and sentences is that identifying mechanisms at one level of processing can provide insights into mechanisms at the other. Although script-based knowledge may be more salient at the narrative level than at the sentence level, the overlap in processes can inform the mechanisms that occur at these levels. For example, research on mental models in narratives draws a distinction between implicit and explicit inferences (Johnson-Laird, 1983). This distinction can help to explain reading time on frequency modifiers of atelic predicates in isolated sentences. Research on isolated sentences draws distinctions between telic and atelic predicates and the salience of different stages of events. These distinctions can help to explain event shifts that occur in narratives.

Acknowledgements

This work was supported by the National Institute of Child Health and Human Development under Grant 1R15HD055680-01A1. I am grateful to a reviewer for comments on an earlier draft of this manuscript, to Susan Rothstein for discussions, to Ryan King for statistical assistance, to Froogh Aziz, Kerry McDermott and Kendra McKim for testing participants, and to Chuck Clifton for his support.

Appendix A

A.1. Materials for Experiment 1

Within each set, predicates appear in the order telic/atelic. Adverbial modifiers appear in the order frame/durative/frequency. The comprehension question follows each sentence.

1. Even though Howard [sent/owed] a large check to his daughter [last year/for several years/every year], she still ran out of money.
   How many times did Howard definitely [send/owe] something?
2. Though the administration [responded to/considered] a document request [last year/for several years/every year], it still withheld crucial information.
   How many document requests did the administration definitely [respond to/consider]?
3. Although Janet [wrote/displayed] a long letter to her congressman [yesterday/for several days/every day], she didn't vote for him.
   How many letters did Janet definitely [write to/display from] her congressman?
4. Even though Brian [scored/guaranteed] a goal in the third period [last night/for several nights/every night], the team dropped into last place.
   How many goals did Brian definitely [score/guarantee] in the third period?
5. Although Joanne [made/owned] a duplicate house key [last year/for several years/every year], she still got locked out.
   How many duplicate house keys did Joanne definitely [make/own]?
6. Although the Senate [passed/debated] a new tax cut [last year/for several years/every year], the surplus continued to grow quite rapidly.
   How many new tax cuts did the Senate definitely [pass/debate]?
7. Though the hospital [moved/cared for] a newborn baby from the nursery [yesterday/for several days/every day], it was still too crowded.
   How many newborn babies did the hospital definitely [move/care for]?
8. Though the secretary [scheduled/anticipated] a conference with the president [yesterday/for several days/every day], there was no good news.
   How many conferences with the president did the secretary definitely [schedule/anticipate]?
9. Although the teacher [gave/planned] a lecture to the class [yesterday/for several days/every day], some students did poorly.
   How many lectures did the teacher definitely [give to/plan for] the class?
10. Even though Elizabeth [dug/planned] an irrigation ditch [last summer/for several summers/every summer], she abhors physical exertion.
    How many irrigation ditches did Elizabeth definitely [dig/plan]?
11. Even though the judge [announced/wanted] a new court procedure [last week/for several weeks/every week], the courtroom was still chaotic.
    How many new court procedures did the judge definitely [announce/want]?
12. Although Barb [hammered/left] a small nail in the plaster wall [last year/for several years/every year], the landlord warned her not to.
    How many small nails did Barb definitely [hammer/leave] in the wall?
13. Although the judge [made/allowed] a large donation to the Democratic Party [last year/for several years/every year], they still didn't win many elections.
    What did the judge [make/allow]?
14. Though the artist [sold/leased] an oil painting to a wealthy benefactor [last year/for several years/every year], she was still quite impoverished.
    What did the artist [sell/lease]?

15. Even though the police [investigated/reported] a traffic accident [yesterday/for several days/every day], the Highway Department refused to lower the speed limit.
    What did the police [investigate/report]?
16. Though the department chair [presented/described] a status report [every meeting/last meeting/for several meetings], the faculty still felt uninformed.
    What did the department chair [present/describe]?
17. Though the hiker [climbed/admired] a high mountain in the Rockies [last year/for several years/every year], she still preferred the Adirondacks.
    What did the hiker [climb/admire]?
18. Though the builder [converted/used] a vacant lot [into/for] a condominium [last year/for several years/every year], there was still not enough residential housing.
    What did the builder [convert/use]?
19. Even though the vandals [sprayed/kept] a can of paint on the cabinet [last night/for several nights/every night], the police did not add more guards.
    What did the vandals [spray/keep]?
20. Although their mother [marinated/served] a large steak [for/to] the kids [every day/yesterday/for several days], they really wanted hot dogs.
    What did their mother [marinate/serve]?
21. Though Cheryl [packed/stored] a cardboard box with books [last weekend/for several weekends/every weekend], she still had a lot to do for her move. What did Cheryl [pack/store] in the cardboard box?
22. Even though he [bought/kept] a rose for Tanya [last week/for several weeks/every week], she still felt unhappy.
    What did he [buy/keep] for Tanya?
23. Although the new contract [guaranteed/delivered] a bonus to the CEO [last year/for several years/every year], he still complained.
    What did the new contract [guarantee/deliver to] the CEO?
24. Although the boss [promised/paid] a large salary to Fred [last month/for several months/every month], he was still frugal.
    What did the boss [promise/pay] to Fred?

A.2. Materials for Experiment 2

Within each set, predicates appear in the order telic/atelic. Adverbial modifiers appear in the order durative/frequency. The comprehension question follows each sentence.

1. The man [ignited/looked for] a candle on the table [for eight nights/every night] after he lit a match.
   How many times did the man definitely [ignite/look for] the candle?
2. The camper [ignited/looked for] a log in the fireplace [for the night/every night] after she gathered the kindling.
   Who [ignited/looked for] the log?
3. The court [overturned/considered] a law passed by Congress [for a week/every week] after it heard the case.
   How many times did the court definitely [overturn/consider] a law?
4. The judge [overturned/considered] a decision from a lower court [for a week/every week] after she reviewed the arguments.
   What did the judge [overturn/consider]?
5. The policeman [awakened/kept an eye on] a suspect in the jail cell [for an afternoon/every afternoon] before the prosecutor arrived.
   How many times did the policeman definitely [awaken/keep an eye on] the suspect?
6. The teacher [awakened/kept an eye on] a student in the classroom [for an afternoon/every afternoon] before the school day ended.
   Where did the teacher [awaken/keep an eye on] the student?
7. The explorer [folded/sought] a map in her backpack [for a few minutes/every morning] after she decided on a destination.
   How many times did the explorer definitely [fold/seek] a map?
8. The secretary [folded/sought] a letter on her desk [for a few minutes/every morning] after she found an envelope.
   Who [folded/sought] a letter?
9. The professor [unfolded/held] a newspaper in the kitchen [for a few minutes/every morning] before his wife brewed some coffee.
   How many times did the professor definitely [unfold/hold] a newspaper?

10. The woman [unfolded/held] a magazine in the waiting room [for a few minutes/every morning] before she saw the doctor.
    What did the woman [unfold/hold] in the waiting room?
11. The chemist [sanitized/monitored] a test tube in her laboratory [for a month/every month] before she mixed the chemicals.
    How many times did the chemist definitely [sanitize/monitor] a test tube?
12. The archaeologist [sanitized/monitored] a fossil in her workshop [for a month/every month] before she loaded the truck.
    Where did the archaeologist [sanitize/monitor] a fossil?
13. The navy [sank/guarded] a battle ship in the harbor [for a year/every year] to prevent an attack.
    How many times did the navy definitely [sink/guard] a battle ship?
14. The pirates [sank/guarded] a boat in the lagoon [for a year/every year] so their prisoners could not escape.
    Who [sank/guarded] a boat?
15. The cop [halted/attended] a parade on Fifth Avenue [for an afternoon/every afternoon] if the weather was bad.
    How many times did the cop definitely [halt/attend] a parade?
16. The dean [halted/attended] a class in Dickson Hall [for an hour/every hour] if there was a fire drill.
    What did the dean [halt/attend]?
17. The tourist [stepped into/watched] a pool behind the hotel [for an hour/every hour] in order to cool off.
    How many times did the tourist definitely [step into/watch] a pool?
18. The bear [stepped into/watched] a river behind the mountain [for an hour/every hour] in order to catch a fish.
    Where did the bear [step into/watch] a river?
19. The kids [capsized/examined] a canoe in the lake [for a day/every day] if a storm approached.
    How many times did the kids definitely [capsize/examine] the canoe?
20. The trapper [capsized/examined] a raft in the swamp [for a day/every day] if it was windy.
    Who [capsized/examined] the raft?
21. The clerk [entered/managed] a store in the mall [for a day/every day] when it was open.
    How many times did the clerk definitely [enter/manage] a store?
22. The owner [entered/managed] a restaurant in the city [for a day/every day] when she was in the neighborhood.
    What did the owner [enter/manage]?
23. The detective [left/waited in] a police station on Main Street [for three nights/every night] when she suspected criminal activity.
    How many times did the detective definitely [leave/wait in] the police station?
24. The student [left/waited in] an office in College Hall [for the night/every night] when she had an appointment.
    Where did the student [leave/wait in] an office?

Appendix B

B.1. Mean reading time (ms) for frame modifiers in Experiment 1 depending on region and predicate type. Residual reading time appears in parentheses

Predicate type	Region			
	Predicate		Modifier	
	Mean	SE	Mean	SE
First pass time				
Telic	1051	58.9	379	22.9
	(−92.4)	(46.4)	(−1.74)	(20.7)
Atelic	1103	59.5	386	19.7
	(−111)	(60.0)	(−0.59)	(20.3)
Total time				
Telic	2016	113	610	43.7
	(−109)	(80.3)	(7.73)	(33.8)
Atelic	2206	79.1	654	40.6
	(−68.9)	(95.3)	(38.7)	(35.2)

B.2. Mean reading time (ms) for durative modifiers in Experiment 1 depending on region and predicate type. Residual reading time appears in parentheses

Predicate type	Region			
	Predicate		Modifier	
	Mean	SE	Mean	SE
First pass time				
Telic	1126	60.6	630	35.7
	(−78.3)	(66.8)	(−36.1)	(25.9)
Atelic	1118	58.9	586	26.9
	(−105)	(76.2)	(−101)	(26.4)
Total time				
Telic	2128	107	1034	70.0
	(−66.2)	(89.1)	(−182)	(53.9)
Atelic	2254	85.2	1011	59.8
	(−59.1)	(92.4)	(−201)	(42.2)

B.3. Mean reading time (ms) for frequency modifiers in Experiment 1 depending on region and predicate type. Residual reading time appears in parentheses

Predicate type	Region			
	Predicate		Modifier	
	Mean	SE	Mean	SE
First pass time				
Telic	1187	38.8	417	21.9
	(−43.6)	(44.3)	(4.08)	(15.9)
Atelic	1061	59.4	432	21.9
	(−105)	(59.6)	(21.1)	(20.4)
Total time				
Telic	1954	78.7	608	40.1
	(−373)	(94.8)	(−62.6)	(31.1)
Atelic	2222	97.4	709	39.9
	(100)	(71.3)	(35.0)	(28.8)

B.4. Mean reading time (ms) for durative modifiers in Experiment 2 depending on region and predicate type. Residual reading time appears in parentheses

Predicate type	Region			
	Predicate		Modifier	
	Mean	SE	Mean	SE
First pass time				
Telic	1044	43.1	511	26.0
	(−190.2)	(46.3)	(−10.7)	(16.6)
Atelic	1077	50.5	526	25.4
	(−235)	(59.3)	(8.03)	(16.3)
Total time				

Telic	2260	81.4	921	58.2
	(−155.7)	(73.8)	(41.0)	(37.6)
Atelic	2250	81.6	847	46.1
	(−62.8)	(67.9)	(18.1)	(26.9)

B.5. Mean reading time (ms) for frequency modifiers in Experiment 2 depending on region and predicate type. Residual reading time appears in parentheses

Predicate type	Region			
	Predicate		Modifier	
	Mean	SE	Mean	SE
First pass time				
Telic	1042	60.3	432	21.0
	(−183)	(50.8)	(−16.1)	(16.1)
Atelic	1051	49.2	443	22.3
	(−221.6)	(65.4)	(−0.96)	(18.5)
Total time				
Telic	2015	97.5	633	42.9
	(−157)	(82.4)	(−76.1)	(24.0)
Atelic	2026	76.2	758	64.9
	(−8.26)	(88.1)	(93.3)	(36.8)

References

Baggio, G., van Lambalgen, M., Hagoort, P., 2008. Computing and recomputing discourse models: an ERP study. J. Mem. Lang. 59 (1), 36–53.

Bates, D., Maechler, M., Bolker, B., 2012. lme4: Linear Mixed-Effects Models Using S4 Classes.

Bott, O., 2010. The Processing of Events. John Benjamins, Amsterdam.

Brennan, J., Pylkkänen, L., 2008. Processing events: behavioral and neuromagnetic correlates of aspectual coercion. Brain Lang. 106 (2), 132–143. http://dx.doi.org/10.1016/j.bandl.2008.04.003

Caudal, P., 1999. Result stages and the lexicon: the proper treatment of event structure. In: Paper presented at the Proceedings of the Ninth Conference on European Chapter of the Association for Computational Linguistics.

Caudal, P., 2005. Stage structure and stage salience for event semantics. In: Aspectual Inquiries. Springer, pp. 239–264.

Clark, E.V., 1971. On the acquisition of the meaning of before and after. J. Verbal Learn. Verbal Behav. 10 (3), 266–275.

Davies, M., 2009. The 385+ million word Corpus of Contemporary American English (1990–2008+): design, architecture, and linguistic insights. Int. J. Corpus Linguist. 14 (2), 159–190.

De Swart, H., 1998. Aspect shift and coercion. Nat. Lang. Linguist. Theory 16 (2), 347–385.

Deo, A., Pinango, M., 2011. Quantification and context in measure adverbs. In: Paper presented at the Proceedings of SALT.

Deo, A., Pinango, M., Lai, Y.Y., Foster-Hanson, E., 2012. Building multiple events: the cost of context retrieval. In: Paper presented at the Architectures and Mechanisms of Language Processing, .

Dowty, D., 1979. Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ, vol. 7. Springer, New York.

Dowty, D.R., 1986. The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? Linguist. Philos. 9 (1), 37–61.

Ferreira, F., Clifton, C., 1986. The independence of syntactic processing. J. Mem. Lang. 25 (3), 348–368.

Ferretti, T., McRae, K., Hatherell, A., 2001. Integrating verbs, situation schemas, and thematic role concepts. J. Mem. Lang. 44 (4), 516–547.

Ferretti, T., Kutas, M., McRae, K., 2007. Verb aspect and the activation of event knowledge. J. Exp. Psychol.: Learn. Mem. Cogn. 33 (1), 182.

Ferretti, T.R., Rohde, H., Kehler, A., Crutchley, M., 2009. Verb aspect, event structure, and coreferential processing. J. Mem. Lang. 61 (2), 191–205.

Filik, R., Paterson, K., Liversedge, S., 2009. The influence of only and even on online semantic interpretation. Psychon. Bull. Rev. 16 (4), 678–683.

Foundation, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Hinrichs, E., 1986. Temporal anaphora in discourses of English. Linguist. Philos. 9 (1), 63–82.

Husband, E., Stockall, L., 2015. Building aspectual interpretations online. In: Cognitive Science Perspectives on Verb Representation and Processing. Springer, pp. 157–186.

Jackendoff, R., 1997. Architecture of the Language Faculty. MIT Press, Cambridge, MA.

Johnson-Laird, P.N., 1983. Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Harvard University Press.

Kuperberg, G.R., Choi, A., Cohn, N., Paczynski, M., Jackendoff, R., 2010. Electrophysiological correlates of complement coercion. J. Cogn. Neurosci. 22 (12), 2685–2701.

Kuznetsova, A., Brockhoff, P., Christensen, R., 2015. Package 'lmerTest'. In: R Package Version, 2.0-29. .

Landman, F., 2011. Boolean Pragmatics. (unpublished manuscript).

Levin, B., 1993. English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press.
Madden, C.J., Ferretti, T.R., 2009. Verb aspect and the mental representation of situations. The Expression of Time, vol. 3. pp. 217–231.
Magliano, J.P., Schleich, M.C., 2000. Verb aspect and situation models. Discourse Process. 29 (2), 83–112.
Moens, M., Steedman, M., 1988. Temporal ontology and temporal reference. Comput. Linguist. 14 (2), 15–28.
Münte, T.F., Schiltz, K., Kutas, M., 1998. When temporal terms belie conceptual order. Nature 395 (6697), 71–73.
Paczynski, M., Jackendoff, R., Kuperberg, G., 2014. When events change their nature: the neurocognitive mechanisms underlying aspectual coercion. J. Cogn. Neurosci. 26 (9), 1905–1917. http://dx.doi.org/10.1162/jocn_a_00638
Pettijohn, K.A., Radvansky, G.A., 2016. Narrative event boundaries, reading times, and expectation. Mem. Cogn. 44 (7), 1064–1075.
Pickering, M., McElree, B., Frisson, S., Chen, L., Traxler, M., 2006. Underspecification and aspectual coercion. Discourse Process. 42 (2), 131–155.
Pinango, M., Zurif, E., Jackendoff, R., 1999. Real-time processing implications of enriched composition at the syntax–semantics interface. J. Psycholinguist. Res. 28 (4), 395–414.
Pinango, M., Winnick, A., Ullah, R., Zurif, E., 2006. Time-course of semantic composition: the case of aspectual coercion. J. Psycholinguist. Res. 35 (3), 233–244.
Pinheiro, J., Bates, D., 2006. Mixed-Effects Models in S and S-PLUS. Springer Science & Business Media.
Pinón, C., 1999. A Semantics for Durative Adverbials. In: Sinn und Bedeutung. Heinrich-Heine-Universität, Düsseldorf.
Proctor, A., Dickey, M., Rips, L., 2004. The time-course and cost of telicity inferences. In: Paper presented at the Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society.
Pustejovsky, J., 1991. The syntax of event structure. Cognition 41 (1–3), 47–81. http://dx.doi.org/10.1016/0010-0277(91)90032-Y
Pustejovsky, J., 1993. Type coercion and lexical selection. In: Semantics and the Lexicon. Springer, pp. 73–94.
Pylkkänen, L., McElree, B., 2006. The syntax-semantics interface: on-line composition of sentence meaning. Handbook of Psycholinguistics, vol. 2. pp. 537–577.
Radvansky, G., 2012. Across the event horizon. Curr. Dir. Psychol. Sci. 21 (4), 269–272.
Radvansky, G.A., Copeland, D.E., 2010. Reading times and the detection of event shift processing. J. Exp. Psychol.: Learn. Mem. Cogn. 36 (1), 210.
Rinck, M., Weber, U., 2003. Who when where: an experimental test of the event-indexing model. Mem. Cogn. 31 (8), 1284–1292.
Rothstein, S., 1995. Adverbial quantification over events. Nat. Lang. Semant. 3 (1), 1–31.
Smith, C.S., 1997. The Parameter of Aspect, vol. 43. Springer.
Stockall, L., Husband, E., 2014. Processing (the) events: lexical and structural ingredients of inner aspect. In: Schutze, C.T., Stockall, L. (Eds.), Connectedness: Papers by and for Sarah VanWagenen. UCLA Working Papers in Linguistics, vol. 18. pp. 255–271.
Todorova, M., Straub, K., Badecker, W., Frank, R., 2000. Aspectual coercion and the online computation of sentential aspect. In: Paper presented at the Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society.
Townsend, D.J., 1983. Thematic processing in sentences and texts. Cognition 13 (2), 223–261.
Townsend, D.J., 2013. Aspectual coercion in eye movements. J. Psycholinguist. Res. 42 (3), 281–306. http://dx.doi.org/10.1007/s10936-012-9216-4
Townsend, D.J., Bever, T., 1988. Knowledge representations during reading depend on reading strategy and reading skill. Practical Aspects of Memory: Current Research and Issues. Clinical and Educational Implications, vol. 2. pp. 309–314.
Townsend, D.J., Bever, T.G., 1991. The use of higher-level constraints in monitoring for a change in speaker demonstrates functionally distinct levels of representation in discourse comprehension. Lang. Cogn. Process. 6 (1), 49–77.
Townsend, D., Bever, T., 2001. Sentence Comprehension: The Integration of Habits and Rules. MIT Press, Cambridge, MA.
Townsend, D.J., Ravelo, N., 1980. The development of complex sentence processing strategies. J. Exp. Child Psychol. 29 (1), 60–73.
Trueswell, J., Tanenhaus, M., Garnsey, S., 1994. Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. J. Mem. Lang. 33 (3), 285–318.
Van Dijk, T.A., Kintsch, W., Van Dijk, T.A., 1983. Strategies of Discourse Comprehension. Academic Press, New York.
Vendler, Z., 1957. Verbs and times. Philos. Rev. 66 (2), 143–160.
Yap, F.H., Chu, P.C.K., Yiu, E.S.M., Wong, S.F., Kwan, S.W.M., Matthews, S., Tan, L.H., Li, P., Shirai, Y., 2009. Aspectual asymmetries in the mental representation of events: role of lexical and grammatical aspect. Mem. Cogn. 37 (5), 587–595.
Zwaan, R.A., 1999. Situation models: the mental leap into imagined worlds. Curr. Dir. Psychol. Sci. 8 (1), 15–18.
Zwaan, R., Radvansky, G., 1998. Situation models in language comprehension and memory. Psychol. Bull. 123 (2), 162.