

Temporal proximity inferences in complex sentence comprehension: Evidence from English complement and relative clauses

Elena Marx (marx_elena@phd.ceu.edu)

Central European University, Department of Cognitive Science, Quellenstraße 51
Vienna, 1100 Austria

Eva Wittenberg (wittenberge@ceu.edu)

Central European University, Department of Cognitive Science, Quellenstraße 51
Vienna, 1100 Austria

Abstract

In language comprehension, mental representations of temporal relations between described situations are often construed by inference. While the basis for these inferences remains unclear, growing evidence suggests that abstract predicate properties – such as dynamicity and causal structure – play a crucial role in temporal event construal. Across two self-paced reading experiments, replicating and extending Gennari (2004), we find that temporal proximity inferences are shaped by these properties, but only for stative predicates that generally encode non-dynamic situations without causal structure: Participants consistently expected states to overlap in time, in both complement (Experiment 1) and relative clause constructions (Experiment 2). These findings indicate that temporal proximity inferences arise as a general feature of (non)-dynamicity, supporting models of language comprehension that prioritize abstract event structural properties in shaping temporal inferences.

Keywords: temporal construal, event cognition, dynamicity, complex sentences, temporal proximity inferences

Introduction

Just as we intuitively understand that the label *chair* does not refer to the same kind of thing as *anteater*, we also understand that *standing* is fundamentally different from *walking*. Whereas *standing* describes an object or agent maintaining an upright position, *walking* denotes an unhurried, self-propelled motion from one place to another. Crucially, these differences between predicates go beyond surface-level semantics: Much like the noun referents *chair* and *anteater* differ in animacy, verbal predicates such as *standing* and *walking* differ with respect to more abstract properties of the situations they describe (Dowty, 1977; Jackendoff, 1990; Mourelatos, 1978; Taylor, 1977; Vendler, 1967; Verkuyl, 1972).

From a cognitive perspective, a crucial question is: Which aspects of a mental representation are influenced by these abstract properties when people read sentences? Specifically, do differences between predicates systematically predict inferences about time, such as about the temporal proximity between situations?

Across linguistic theories, one fundamental distinction has been drawn between predicates that describe states and those that describe events (Bach, 1986; Dowty, 1979; Jackendoff,

1990; Rappaport Hovav & Levin, 1998; Vendler, 1967). A key difference between these types of predicates lies in whether they convey dynamic concepts (Dowty, 1979; Kearns, 2000; Olsen, 1994; Smith, 1999): While stative predicates describe situations that do not change over time (e.g., *to love*, *to belong*), eventive predicates describe situations that are dynamic and involve temporal progression (e.g., *to run*, *to build*). Importantly, this difference has cognitive ramifications: How temporal order between described situations is inferred is predicted by their dynamic properties (Marx & Wittenberg, 2022, 2024, 2025; Marx, Jardón, & Wittenberg, accepted).

Tied to the property of dynamicity, however, is that eventive and stative predicates also differ in their causal structure (Gennari, 2004; Gennari & Poeppel, 1998): Eventive predicates typically entail a cause that propels an action (e.g., *to jump*) or brings about a change from an initial to a resulting state (e.g. *to break*, *to destroy*). In both cases, causes involve some sort of agentive force: either another event (e.g., *The storm destroyed the city*), or a volitional agent performing the action (e.g., *The boy jumped up and down*). Stative verbs, by contrast, do not entail internal causal structure: They denote stable properties or relations between participants which exist independent of a causal agent or event (Dowty, 1979; Parsons, 1990; but see Dixon, 1982 for resultative vs. non-resultative states).

Naturally, causal structure is closely linked to temporal inference, as causes must logically precede their effects (Bechlivanidis & Lagnado, 2016; Briner, Virtue, & Kurby, 2011; Lascarides & Asher, 1993; Oversteegen, 2005). More importantly for our purposes, however, the presence or absence of an agentive force introduces different logical entailments which, in turn, may shape how comprehenders infer temporal proximity in language comprehension: On the one hand, situations described by eventive predicates typically end when their causes cease, giving rise to the inference that events – since they may causally trigger one another – should happen in close temporal proximity (Kamp & Reyle, 1993; Moens & Steedman, 1988). Situations described by stative predicates, on the other hand, can persist without a cause and so they should hold for indefinite periods of time and overlap with other situations in the world (Abusch, 1997; Dowty, 1986; Gennari, 2003; Hinrichs,

1986; Kamp & Reyle, 1993; Lascarides & Asher, 1991; Ogihara, 1996; ter Meulen, 1995).

Empirical evidence for these proximity inferences primarily comes from studies on on-line discourse comprehension (Ditman, Holcomb, & Kuperberg, 2008; Kelter, Kaup, & Claus, 2004; Rinck & Weber, 2003; Zwaan, 1996), showing that comprehenders incur greater processing costs when integrating temporal distance between two described events compared to temporal closeness. These findings, however, cannot rule out other factors that might drive proximity inferences: linearity, a strong principle in establishing narrative progression (Fleischman, 1990; Schmerling, 1975; Tai, 1983), or discourse coherence reasoning (Kehler, 2006). Thus, it remains unclear whether observed proximity effects are solely driven by event type or by broader principles that determine temporal relations in narratives.

At the same time, much less is known about how comprehenders process temporal relations involving states: In a limited number of studies, using verbal complement clauses as linguistic context, comprehenders interpreted states as overlapping with other events in time (Dickey, 2000; Mucha, Renans, & Romoli, 2023; but see Armenante, Hohaus, & Stolterfoht, 2023). However, those studies employed offline judgment tasks, leaving open the question of how proximity inferences emerge in online processing, and how such inferences about states relate to the above-mentioned empirical findings about events.

To the best of our knowledge, only one study has investigated temporal proximity inferences across event types during on-line sentence comprehension: In Gennari (2004), participants read complement clause constructions in a self-paced reading paradigm, where the subordinate clause described either a temporally close or distant event, or an overlapping or non-overlapping state. Reading times were slower when the embedded event that was temporally distant, rather than close from the main clause event, and when the embedded state did not overlap, rather than overlap with the main clause event. In these conditions, participants revised their previous inferred temporal relation between the two situations encoded in the sentence.

However, Gennari (2004) exclusively used verbal complement clause constructions – a sentence type in which, from a formal point of view, temporal structure is predicted by syntax (Abusch, 1997; Enç, 1987; Ogihara, 1996; Stowell, 2007; von Stechow, 1995): In complement clause constructions, the embedded clause functions as an argument of the main clause predicate, constraining interpretations such that the embedded situation must be evaluated relative to the time established in the main clause.

Furthermore, diachronically, complement clauses have been argued to have evolved from causal clauses (Deutscher, 2000; Givón, 1991; Zuckermann, 2006; but see Hernáiz, 2024), suggesting that causality underlies these constructions: This is particularly the case for factive predicates (e.g., *to be delighted*, *to know*) where the complement often represents the cause of the event

expressed by the matrix clause (e.g., *John was delighted that his paper got accepted* implies *John was delighted because his paper got accepted*). Against this backdrop, an important question is: Do Gennari's (2004) findings reflect general temporal proximity inferences driven by the dynamic and causal properties of events and states, or are they specific to the linking constraints for complement clause constructions? The present set of studies is designed to answer this question, using reading times as a proxy for processing effort elicited by temporal inferences. In Experiment 1, we replicate Gennari (2004), asking whether event type drives inferences of temporal proximity in complement clauses: Integrating temporally distant events or non-overlapping states in the subordinate clause should result in longer reading times than integrating temporally close events or overlapping states. In Experiment 2, we expand the empirical base by examining a construction type that is traditionally considered to lack systematic causal, temporal, or conditional relations (Thomson & Martinet, 1985): relative clauses. Predictions mirrored those in Experiment 1.

Experiment 1: Complement Clauses

Participants

240 native English speakers were recruited via Prolific.¹ The experiment was programmed and run online via IBEX (Zehr & Schwarz, 2018) and lasted approximately 15 minutes.

Materials

We took 16 past-under-past complement clause pairs from Gennari (2004): eight with embedded events like (1), and eight with embedded states like (2).

- | | | |
|--|--|-------------|
| (1) <i>The department chair complained today that the dean decided on a new professor earlier</i> | | |
| a. | <i>this week without his consent</i> | close |
| b. | <i>this year without his consent</i> | distant |
| (2) <i>Union representatives complained this week that the unemployment rate was alarming</i> | | |
| a. | <i>this month and requested government intervention</i> | overlap |
| b. | <i>last month and requested government intervention</i> | non-overlap |

Embedded events were encoded by telic verbs (e.g., *to decide*, *to reject*); embedded states either by *be* + adjective or by psychological verbs (e.g. *to be alarming*, *to hate*).

All sentences contained two temporal adverb phrases: one locating the main clause event in the past, and the other locating the embedded event or state before the main clause event. Embedded adverb phrases consisted of a determiner (i.e., *this*, *last*) and a temporal noun (i.e., *week*, *month*). Number of characters, frequency of the nouns and frequency of determiners and nouns combined were matched across sentences, according to Gennari (2004).

Sentence pairs (a-b) differed only in terms of embedded temporal adverbs: In the proximal conditions (1a, 2a), the

¹ We overshot the preregistered sample size (N=72), as online data collection resulted in high data exclusion and low power

(54.1% estimated power for observed $R^2=0.061$, using G*Power (Faul et al., 2007), with $\alpha=0.05$, two-tailed).

adverb established a close temporal relation to the main clause event for embedded events and a temporal overlap for embedded states. In the distal conditions (1b, 2b), temporal adverbs established a large temporal gap between main clause events and embedded event, and a non-overlap with embedded states. All materials can be found at <https://osf.io/q8msz/>.²

Procedure

Participants read sentences like (1-2) one word at a time in a self-paced moving window paradigm (Just, Carpenter, & Woolley, 1982): Each word was presented individually in its sentence position, indicated by a dashed line, with individual dashes representing individual words. Participants pressed the spacebar, to change the next dash into the corresponding word and to return the previously read word back into a dash. Times between spacebar presses indicated reading times for each word of the sentence.

After each sentence, participants answered a yes/no comprehension question targeting the embedded temporal inference, to ensure that people correctly understood the temporal information in the embedded clause.

There were 16 critical trials and 100 filler trials, all presented in a randomized order. Filler sentences consisted of complex, or one or two simplex sentences. Subsequent attention questions targeted the comprehension of temporal order or general content of the sentences.

Participants were randomly assigned to one of four lists, in a Latin square design that crossed embedded event type and temporal proximity across items.

Statistical Analysis

Trials were only analyzed when the attention check question was answered correctly. Individual data points were excluded if they exceeded 1.000ms or fell below 200ms.

We fitted a series of linear mixed-effects regression models (i.e., *lmer* from *lme4*, Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2014) and compared each full model to reduced models via likelihood ratio tests. Embedded event type (state vs. event) and temporal proximity (proximal: close/overlap vs. distal: distant/non-overlap) were coded as contrast-coded fixed effects and participant and item as random intercepts.

Like Gennari (2004), we individually analyzed reading times for (a) the word preceding the embedded temporal adverb, (b) the determiner, (c) the critical noun, and (d) the word following the temporal adverb. Additionally, we analyzed reading speed (words/s) across (e) the critical four-word region and (f) the remaining sentence after it.

The most complex random effect structure without convergence failure was fitted (Barr, Levy, Scheepers, & Tily, 2013), so that random slopes varied across models: participant and item variance on proximity for (a) and (d); item variance on proximity for (b), participant variance on

proximity for (c); participant variance on event type and proximity and item variance on proximity for (e); participant variance on event type for (f).

Planned pairwise comparisons served to assess interactions more thoroughly. Analyses were pre-registered at <https://osf.io/eg2fn>. Due to limited space, we depart slightly from our preregistration and focus on the critical four-word and the following region. Full data and analyses are, however, available in the project folder.

Results

All results are outlined in Table 1. We excluded 20.8% incorrect trials, 1.9% data points slower than 1.000ms and 7.5% data points faster than 200ms.

The previous word (Figure 1) was read marginally faster for embedded states ($M_{state}=375.0$, $SD=141.8$) than for embedded events ($M_{event}=388.4$, $SD=146.9$). This difference was significant between proximal conditions ($SD_{close}=151.4$ vs. $SD_{overlap}=140.6$, Table 1A). There were no reading time differences for the determiner (Figure 1, Table 1B).

For the critical noun (Figure 1), reading times were marginally slower for embedded states ($M_{state}=407.6$, $SD=163.8$) than for embedded events ($M_{event}=389.7$, $SD=144.1$), which was significant between distal conditions ($SD_{distant}=148.4$ vs. $SD_{non-overlap}=162.6$, Table 1C). The following word (Figure 1) was read faster for distal events than for distal states ($SD_{distant}=139.6$ vs. $SD_{non-overlap}=149.5$, see Table 1D).

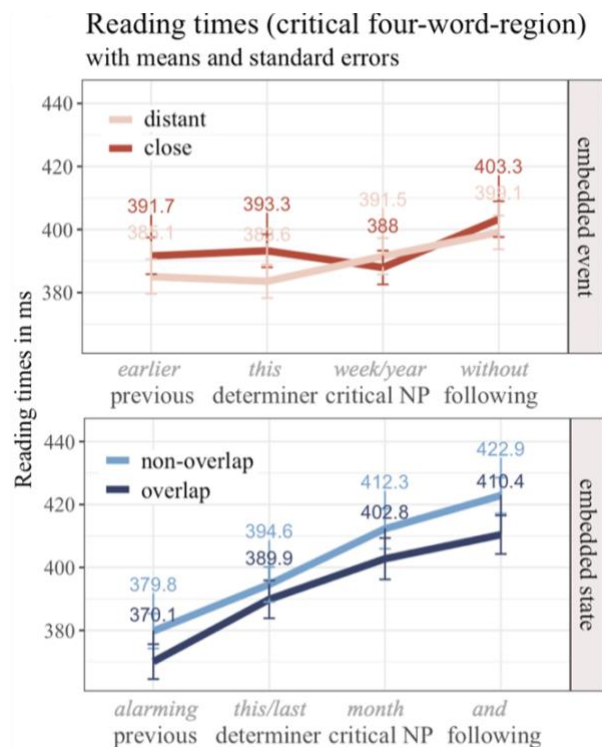


Figure 1 Reading times for each word of the critical region in Experiment 1; error bars represent standard errors.

² We preregistered an additional study in the original project; however, this study's focus and design could better be reported in a different format, leading us to only report two studies here.

Reading speed in the critical region was faster for states ($M_{state}=3.51$, $SD=2.75$) than for events ($M_{event}=3.33$, $SD=2.31$), particularly for proximal states ($M_{overlap}=3.67$, $SD=3.07$) compared to both distal states ($M_{non-overlap}=3.36$, $SD=2.40$; $\beta=0.16$, $t=2.17$, $p=0.03$) and proximal events ($M_{close}=3.26$, $SD=2.09$, see Table 1E).

The remaining sentence was read faster for proximal states than proximal events ($M_{overlap}=3.14$, $SD=2.03$ vs. $M_{close}=2.96$, $SD=1.55$, see Table 1F).

Table 1 Summary of results of the regression models and pairwise comparisons in Experiment 1 for each measure.

	regression model		pairwise comparisons	
A previous (reading time)	proximity: $Df=1$, $\chi^2=0$, $p=0.99$ event type: $Df=1$, $\chi^2=3.27$, $p=0.07$ interaction: $Df=1$, $\chi^2=2.31$, $p=0.12$	ns . ns	events: $\beta=3.28$, $t=0.82$, $p=0.41$ states: $\beta=-4.86$, $t=-1.25$, $p=0.21$ proximal: $\beta=10.80$, $t=2.69$, $p=0.007$ distal: $\beta=2.65$, $t=0.68$, $p=0.50$	ns ns ** ns
B determiner (reading time)	proximity: $Df=1$, $\chi^2=0.09$, $p=0.76$ event type: $Df=1$, $\chi^2=0.16$, $p=0.69$ interaction: $Df=1$, $\chi^2=1.61$, $p=0.20$	ns ns ns	events: $\beta=4.84$, $t=1.30$, $p=0.19$ states: $\beta=-2.37$, $t=-0.58$, $p=0.56$ proximal: $\beta=1.70$, $t=0.43$, $p=0.67$ distal: $\beta=-5.50$, $t=-1.44$, $p=0.15$	ns ns ns ns
C critical NP (reading time)	proximity: $Df=1$, $\chi^2=1.46$, $p=0.23$ event type: $Df=1$, $\chi^2=3.63$, $p=0.06$ interaction: $Df=1$, $\chi^2=0.20$, $p=0.65$	ns . ns	events: $\beta=-1.79$, $t=-0.46$, $p=0.65$ states: $\beta=-4.76$, $t=-1.04$, $p=0.30$ proximal: $\beta=-7.41$, $t=-1.76$, $p=0.08$ distal: $\beta=-10.37$, $t=-2.43$, $p=0.02$	ns ns . *
D following (reading time)	proximity: $Df=1$, $\chi^2=0.99$, $p=0.32$ event type: $Df=1$, $\chi^2=2.56$, $p=0.11$ interaction: $Df=1$, $\chi^2=2.48$, $p=0.12$	ns ns ns	events: $\beta=2.09$, $t=0.54$, $p=0.59$ states: $\beta=-6.23$, $t=-1.49$, $p=0.14$ proximal: $\beta=-3.56$, $t=-0.85$, $p=0.39$ distal: $\beta=-11.88$, $t=-3.03$, $p=0.002$	ns ns ns **
E four-word region (reading speed)	proximity: $Df=1$, $\chi^2=0.58$, $p=0.45$ event type: $Df=1$, $\chi^2=4.36$, $p=0.04$ interaction: $Df=1$, $\chi^2=6.57$, $p=0.01$	ns * *	events: $\beta=-0.07$, $t=-1.16$, $p=0.24$ states: $\beta=0.16$, $t=2.17$, $p=0.03$ proximal: $\beta=-0.21$, $t=-2.99$, $p=0.003$ distal: $\beta=0.02$, $t=-0.35$, $p=0.73$	ns * ** ns
F remaining sentence (reading speed)	proximity: $Df=1$, $\chi^2=0.04$, $p=0.83$ event type: $Df=1$, $\chi^2=0.08$, $p=0.78$ interaction: $Df=1$, $\chi^2=0.40$, $p=0.53$	ns ns ns	events: $\beta=-0.04$, $t=-0.89$, $p=0.37$ states: $\beta=0.03$, $t=0.59$, $p=0.56$ proximal: $\beta=-0.09$, $t=-1.93$, $p=0.05$ distal: $\beta=-0.02$, $t=-0.41$, $p=0.69$	ns ns . ns

Discussion of Experiment 1

We partially replicated Gennari's (2004) findings in Experiment 1, showing that temporal proximity inferences are influenced by the dynamic and causal properties of

stative predicates: All effects were driven by the fact that non-overlapping states were read slower or overlapping states were read faster than other conditions. At the same time, reading times for eventive predicates did not differ between temporally close or distant events.

In Experiment 2, we asked whether such temporal proximity inferences arise across complex sentences, using relative clauses as a less temporally constrained construction type.

Experiment 2: Relative Clauses

Participants

We recruited 240 native English speakers online through the Testable Minds platform (<https://www.testable.org/>).

Materials and Procedure

In Experiment 2, we used subject-extracted relative clauses that relativized the syntactic object of the main clause. Sentences were derived from the materials in Experiment 1, by changing main clause predicates into transitive predicates and topicalizing the temporal adverb in the main clause, to avoid processing difficulties due to an inserted constituent between antecedent head noun and relative clause (see 3-4).

- | | | |
|-----|--|-------------|
| (3) | <i>Today, the department chair complained to the dean who decided on a new professor earlier</i> | |
| | a. <i>this week</i> without his consultation or consent | close |
| | b. <i>this year</i> without his consultation or consent | distant |
| (4) | <i>This week, union representatives complained about the unemployment rate that was alarming</i> | |
| | a. <i>this month</i> and requested government intervention within the next days. | overlap |
| | b. <i>last month</i> and requested government intervention within the next days. | non-overlap |

Furthermore, we extended sentences after the critical region, to create a three-word spill-over region for all items, to assess the magnitude of the effect on sentence processing, while controlling for sentence wrap-up effects (Just et al., 1982; Mitchell & Green, 1978).

Experimental manipulations, procedure and filler trials were the same as in Experiment 1: embedded events (3) were contrasted with embedded states (4), and temporal proximity was either proximal (i.e., 3a, 4a) or distal (i.e., 3b, 4b), established via temporal adverbs in the main and relative clause. As in Experiment 1, the embedded clause was always temporally located before the main clause.

Statistical Analysis

In addition to the statistical analyses in Experiment 1 (i.e., mixed effect models with maximal random effect structure; planned pairwise comparisons), we analysed reading speed in a three-word spill-over region after the critical region (g). Random slopes were coded as follows: participant variance on proximity for (a); participant and item variance on proximity for (b); participant variance on proximity and event type for (e), while excluding item as a random effect; participant variance on event type for (f). No random slopes were included for the analysis of (c), (d) and (g).

Results

Results are shown in Table 2. 28.7% of all trials were excluded due to incorrect responses, while 1.8% and 9.2% of data points were excluded for having durations that were either too long or too short, respectively.

Reading times for the previous word, the determiner and the critical noun (Figure 2) did not differ across conditions (see Table 2A-C).

The word following the embedded temporal adverb (Figure 2) was read slower for embedded states ($M_{state}=400.4$, $SD=163.7$) than for embedded events ($M_{event}=383.2$, $SD=146.9$). Reading times were furthermore slower for distal states both compared to proximal states ($SD_{non-overlap}=170.0$ vs. $SD_{overlap}=156.3$) and distal events ($SD_{distant}=141.5$, see Table 2D).

Reading speed across the critical region was slower for distal states ($M_{non-overlap}=3.54$, $SD=2.41$) compared to both proximal states ($M_{overlap}=3.99$, $SD=3.29$), and distal events ($M_{distant}=3.92$, $SD=3.19$), as reflected by a significant interaction (see Table 2E).

In the following sentence, states were read marginally slower than events ($M_{state}=3.74$, $SD=3.01$, $M_{event}=4.08$, $SD=4.01$). Reading speed was faster for proximal events ($M_{close}=4.32$, $SD=4.52$) than for both distal events ($M_{distant}=3.83$, $SD=3.40$) and proximal states ($M_{overlap}=3.81$, $SD=3.15$, see Table 2F). However, in the spillover region, the opposite was the case: Reading speed for states was significantly faster than for events in the proximal conditions ($M_{overlap}=3.78$, $SD=2.50$ vs. $M_{close}=3.48$, $SD=2.44$) and marginally faster in the distal conditions ($M_{non-overlap}=3.72$, $SD=2.47$ vs. $M_{distant}=3.48$, $SD=2.37$, see Table 2G).

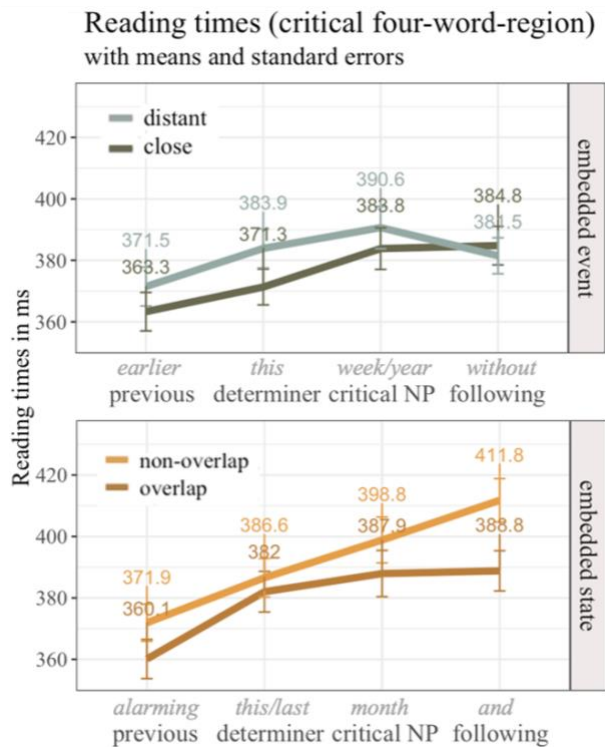


Figure 2 Reading times for each word of the critical region in Exp. 2; error bars represent standard errors.

Table 2 Summary of results of the regression models and pairwise comparisons in Experiment 2 for each measure.

	regression model		pairwise comparisons	
A	proximity: $Df=1, \chi^2=2.48, p=0.12$ event type: $Df=1, \chi^2=0.04, p=0.84$ interaction: $Df=1, \chi^2=0.62, p=0.43$	ns	events: $\beta=-4.08, t=-0.92, p=0.36$ states: $\beta=-5.90, t=-1.34, p=0.18$ proximal: $\beta=1.60, t=0.36, p=0.72$ distal: $\beta=-0.23, t=-0.05, p=0.96$	ns ns ns ns
B	proximity: $Df=1, \chi^2=0.85, p=0.36$ event type: $Df=1, \chi^2=1.14, p=0.29$ interaction: $Df=1, \chi^2=0.20, p=0.66$	ns ns ns	events: $\beta=-6.29, t=-1.46, p=0.15$ states: $\beta=-2.28, t=-0.50, p=0.62$ proximal: $\beta=-5.35, t=-1.22, p=0.22$ distal: $\beta=-1.34, t=-0.36, p=0.76$	ns ns ns ns
C	proximity: $Df=1, \chi^2=1.17, p=0.28$ event type: $Df=1, \chi^2=1.11, p=0.29$ interaction: $Df=1, \chi^2=0.12, p=0.73$	ns ns ns	events: $\beta=-3.41, t=-0.71, p=0.48$ states: $\beta=-5.46, t=-1.03, p=0.30$ proximal: $\beta=-2.05, t=-0.41, p=0.69$ distal: $\beta=-4.10, t=-0.81, p=0.42$	ns ns ns ns
D	proximity: $Df=1, \chi^2=1.55, p=0.21$ event type: $Df=1, \chi^2=4.48, p=0.03$ interaction: $Df=1, \chi^2=3.26, p=0.07$	ns * .	events: $\beta=1.65, t=0.38, p=0.70$ states: $\beta=-11.49, t=-2.39, p=0.02$ proximal: $\beta=0.83, t=0.20, p=0.84$ distal: $\beta=-15.14, t=-3.30, p<0.001$	ns * ns ***
E	proximity: $Df=1, \chi^2=1.85, p=0.17$ event type: $Df=1, \chi^2=1.44, p=0.23$ interaction: $Df=1, \chi^2=6.32, p=0.01$	ns ns *	events: $\beta=-0.01, t=-0.08, p=0.94$ states: $\beta=0.23, t=2.80, p=0.005$ proximal: $\beta=-0.04, t=-0.48, p=0.63$ distal: $\beta=0.19, t=2.42, p=0.02$	ns ** ns *
F	proximity: $Df=1, \chi^2=2.07, p=0.15$ event type: $Df=1, \chi^2=3.42, p=0.06$ interaction: $Df=1, \chi^2=0.95, p=0.33$	ns . ns	events: $\beta=0.24, t=2.22, p=0.03$ states: $\beta=0.07, t=0.83, p=0.41$ proximal: $\beta=0.26, t=2.41, p=0.02$ distal: $\beta=0.08, t=0.95, p=0.34$	* ns * ns
G	proximity: $Df=1, \chi^2=0.04, p=0.83$ event type: $Df=1, \chi^2=0.92, p=0.34$ interaction: $Df=1, \chi^2=0.01, p=0.93$	ns ns ns	events: $\beta=-0.003, t=-0.05, p=0.96$ states: $\beta=0.03, t=0.43, p=0.67$ proximal: $\beta=-0.15, t=-2.17, p=0.03$ distal: $\beta=-0.12, t=-1.73, p=0.08$	ns ns * .

Discussion of Experiment 2

Experiment 2 largely replicated Experiment 1 with relative clause constructions: Participants expected embedded states to overlap and read temporal adverb phrases slower when they specified a non-overlap. Overall, there was no evidence for temporal proximity inferences regarding events: Reading

times were equally fast for temporally close or distant embedded events, apart from reading speed across the remaining sentence, which did not control for sentence wrap-up effects. The effect was reverse in the spill-over region, pointing towards a general difference between processing of stative and eventive predicates.

General Discussion

The present study investigated whether different event types systematically predict temporal inferences between situations during on-line sentence comprehension. Following previous work (Gennari, 2004), we used reading times as a proxy for processing effort triggered by temporal inferences. We asked: Do state and event predicates lead to different temporal proximity inferences, based on their dynamic and causal properties?

In two self-paced reading experiments, we partly replicated Gennari (2004) using complement clauses (Experiment 1) and extended the study to relative clauses (Experiment 2) – a construction type that lacks systematic semantic relations between two described situations (Thomson & Martinet, 1985). In both experiments, participants read past events in the main clauses and either past events or states in the subordinate clauses, with temporal proximity manipulated through temporal adverbs: For embedded events, proximity was either close or distant relative to the main clause event; for embedded states, proximity was either overlapping or non-overlapping with the main clause event.

Our results show that temporal proximity inferences are reliably predicted by event type: Proximal states elicited faster reading times than distal states, consistent with the expectations that states should overlap in time, and in line with Gennari's (2004) previous findings. Crucially, this pattern of results held both for complement and relative clause constructions, highlighting the robustness of the effect and its independence from specific clause-linking relations.

However, in contrast to the original study, we found no evidence for temporal proximity inferences for event descriptions: Reading times did not differ between temporally close or distant embedded events, suggesting that integrating temporal distance for events did not impose additional processing costs for comprehenders.

In principle, our study confirms that people are sensitive to a well-established linguistic distinction between predicates that encode states and those that encode events (Bach, 1986; Dowty, 1979; Olsen, 1994; Vendler, 1967); and use this distinction as a basis for temporal proximity inferences during sentence comprehension. However, our pattern of results challenges assumptions about the precise nature of this distinction as the basis for such inferences: While prior studies have attributed these inferences to the presence or absence of causal structure in these event types (Gennari, 2004; Gennari & Poeppel, 1998; see also Kamp & Reyle, 1993; Moens & Steedman, 1988), the complete absence of temporal proximity inferences for events in our data suggests that dynamicity alone might be the critical factor guiding these inferences.

We believe there are good reasons for this conclusion: First, causal inference is not straightforward when it comes to temporal proximity (Buehner & McGregor, 2006; Gong & Bramley, 2024; Henne, Kulesza, Perez, & Houcek, 2021). While causes must logically precede their effects, this does not necessitate that effects occur immediately after their causes: For example, the decision to take a painkiller might occur hours before the actual act of taking it, and the effect (i.e., pain relief) might not manifest until later. Similarly, causal chains in climate systems can involve significant time lags, with carbon emissions leading to observable environmental consequences only decades later.

Second, situations described by stative predicates might elicit expectations of overlaps not only because they can persist without a cause, but also because they are temporally unbounded and homogeneous (Bennett & Partee, 1972; Dowty, 1986; Lascarides & Asher, 1991): Lacking inherent temporal limits and dynamic change, states may simply be easier to align and co-exist with other situations. Crucially, while these temporal properties are compatible with causality, they are not inherently tied to it. In fact, we propose an alternative mechanism for temporal proximity inferences that explains our results without invoking additional causal reasoning: a general Figure-Ground bias.

On such an account, language encodes spatial, temporal, and conceptual relationships in systematic ways (Gleitman, Gleitman, Miller, & Ostrin, 1996; Levinson, 1996; Talmy, 1975; Zlatev, 2010): Moving or movable entities are usually conceptualized as the "Figure" against a more stable "Ground" entity (Gleitman et al., 1996, p. 358; see also Sun, Firestone, & Hafri, 2023 for temporal relations between Figure and Ground entities). Given their non-dynamic properties, static states should thus align more closely with the concept of Ground, while dynamic events should be more likely conceived as Figures (Hopper, 1979; Reinhart, 1984; Wärvik, 2004). The inference of temporal overlap emerges as a natural by-product of this distinction: Grounds provide a stable and continuous background against which Figures unfold.

A Figure-Ground explanation is also in line with previous empirical findings concerning temporal inferences, using different methods, and several languages (Marx, Iwan, & Wittenberg, 2025; Marx & Wittenberg, 2022, 2024, 2025). This is particularly important since we know from previous studies that reading times are always to be interpreted with caution: Faster reading times may reflect processing effort, but they may also indicate that people abandon a parse (Morgan, von der Malsburg, Ferreira, & Wittenberg, 2020; Nicenboim, Logačev, Gattei, & Vasishth, 2016). However, as we have triangulated our reading times results with previous behavioral evidence showing that states and events indeed result in different temporal inferences, we are confident that our explanation aligns both with the on-line processing pattern obtained here, and the fact that event type differences affect the actual temporal inferences people draw during comprehension.

References

- Abusch, D. (1997). Sequence of Tense and Temporal De Re. *Linguistics and Philosophy*, 20, 1–50.
- Armenante, G., Hohaus, V., & Stolterfoht, B. (2023). Transparency in the processing of temporal ambiguity: The case of embedded tense. *Proceedings of Experiments in Linguistic Meaning*, 2, 1–12. <https://doi.org/10.3765/elm.2.5390>
- Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, 9(1), 5–16.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition*, 146, 58–66. <https://doi.org/10.1016/j.cognition.2015.09.001>
- Bennett, M., & Partee, B. H. (1972). Toward the Logic of Tense and Aspect in English. *Compositionality in Formal Semantics*, 59–109. <https://doi.org/10.1002/9780470751305.ch4>
- Briner, S. W., Virtue, S., & Kurby, C. A. (2011). Processing causality in narrative events: Temporal order matters. *Discourse Processes*, 49(1), 61–77. <https://doi.org/10.1080/0163853X.2011.607952>
- Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking and Reasoning*, 12(4), 353–378. <https://doi.org/10.1080/13546780500368965>
- Deutscher, G. (2000). *Syntactic change in Akkadian: the evolution of sentential complementation*. OUP Oxford.
- Dickey, M. W. (2000). *The processing of tense*. University of Massachusetts Amherst.
- Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin and Review*, 15(4), 750–756. <https://doi.org/10.3758/PBR.15.4.750>
- Dixon, R. M. W. (1982). Where have All the Adjectives Gone? In *Where have All the Adjectives Gone?* <https://doi.org/10.1515/9783110822939>
- Dowty, D. R. (1977). Toward a semantic analysis of verb aspect and the English “imperfective” progressive. *Linguistics and Philosophy*, 1(1), 45–77. <https://doi.org/10.1007/BF00351936>
- Dowty, D. R. (1979). *Word meaning and Montague Grammar: The semantics of verbs and times in Generative Semantics and in Montague’s PTQ*. Dordrecht: Springer Dordrecht.
- Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy*, 9(1), 37–61.
- Enç, M. (1987). Anchoring Conditions for Tense. *Linguistic Inquiry*, 18(4), 633–657.
- Fleischman, S. (1990). *Tense and narrativity: From medieval performance to modern fiction*. <https://doi.org/10.7560/780903>
- Gennari, S. P. (2003). Tense meanings and temporal interpretation. *Journal of Semantics*, 20(1), 35–71. <https://doi.org/10.1093/jos/20.1.35>
- Gennari, S. P. (2004). Temporal references and temporal relations in sentence comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(4), 877–890. <https://doi.org/10.1037/0278-7393.30.4.877>
- Gennari, S. P., & Poeppel, D. (1998). *Events versus States: Empirical Correlates of Lexical Classes*.
- Givón, T. (1991). The evolution of dependent clause morpho-syntax in Biblical Hebrew. *Approaches to Grammaticalization*, 2, 257–310.
- Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition*, 58(3), 321–376. [https://doi.org/10.1016/0010-0277\(95\)00686-9](https://doi.org/10.1016/0010-0277(95)00686-9)
- Gong, T., & Bramley, N. (2024). Evidence from the future. *Journal of Experimental Psychology: General*, 153(3), 864–872.
- Henne, P., Kulesza, A., Perez, K., & Houcek, A. (2021). Counterfactual thinking and recency effects in causal judgment. *Cognition*, 212(July 2020), 104708. <https://doi.org/10.1016/j.cognition.2021.104708>
- Hernáiz, R. (2024). Causal clauses as source of sentential complementation: cross-linguistic evidence and methodological issues. *Linguistics Vanguard*, 1–11. <https://doi.org/10.1515/lingvan-2022-0010>
- Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and Philosophy*, 9(1), 63–82.
- Hopper, P. J. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Discourse and Syntax* (pp. 211–241). https://doi.org/10.1163/9789004368897_010
- Jackendoff, R. (1990). *Semantic structures*. The MIT Press.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–238. <https://doi.org/10.1037/0096-3445.111.2.228>
- Kamp, H., & Reyle, U. (1993). Tense and aspect. In H. Kamp & U. Reyle (Eds.), *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory* (pp. 483–689).
- Kearns, K. (2000). *Semantics*. New York: St. Martin’s Press.
- Kehler, A. (2006). Discourse coherence. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 241–265). <https://doi.org/10.1002/9780470756959.ch11>
- Kelter, S., Kaup, B., & Claus, B. (2004). Representing a Described Sequence of Events: A Dynamic View of Narrative Comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(2),

- 451–464. <https://doi.org/10.1037/0278-7393.30.2.451>
- Lascarides, A., & Asher, N. (1991). Discourse relations and defeasible knowledge. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, (1), 55–62. <https://doi.org/10.3115/981344.981352>
- Lascarides, A., & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16, 437–493. <https://doi.org/10.1007/BF00986208>
- Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology*, 25(1), 353–382. Retrieved from <http://arjournals.annualreviews.org/loi/anthro>
- Marx, E., Iwan, O., & Wittenberg, E. (2025). Dynamicity predicts inferred temporal order in complex sentences: Evidence from English, German, and Polish. *Cognitive Science*, 49, e70038. <https://doi.org/10.1111/cogs.70038>
- Marx, E., Jardón, N., & Wittenberg, E. (n.d.). The state-before-event inference emerges across tenses. *Open Mind*.
- Marx, E., & Wittenberg, E. (2022). Event structure predicts temporal interpretation of English and German past-under-past relative clauses. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 439–445. Retrieved from <https://escholarship.org/uc/item/5z61175d>
- Marx, E., & Wittenberg, E. (2024). Eventuality type predicts temporal order inferences in discourse comprehension. *Glossa Psycholinguistics*, 3, 1–24. <https://doi.org/10.5070/G60116579>
- Marx, E., & Wittenberg, E. (2025). Temporal construal in sentence comprehension depends on linguistically encoded event structure. *Cognition*, 254(October 2024), 105975. <https://doi.org/10.1016/j.cognition.2024.105975>
- Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, 30(4), 609–636. <https://doi.org/10.1080/14640747808400689>
- Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14(2), 15–28.
- Morgan, A. M., von der Malsburg, T., Ferreira, F., & Wittenberg, E. (2020). Shared syntax between comprehension and production: Multi-paradigm evidence that resumptive pronouns hinder comprehension. *Cognition*, 205, 104417. <https://doi.org/10.1016/j.cognition.2020.104417>
- Mourelatos, A. P. D. (1978). Events, processes, and states. *Linguistics and Philosophy*, 2, 415–434.
- Mucha, A., Renans, A., & Romoli, J. (2023). Sequence of tense and cessation implicatures: Evidence from Polish. *Natural Language and Linguistic Theory*, 41(1), 267–346. <https://doi.org/10.1007/s11049-022-09545-2>
- Nicenboim, B., Logačev, P., Gattei, C., & Vasishth, S. (2016). When High-Capacity Readers Slow Down and Low-Capacity Readers Speed Up: Working Memory and Locality Effects. *Frontiers in Psychology*, 7. Retrieved from <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2016.00280>
- Ogihara, T. (1996). *Tense, Attitudes and Scope*. Springer Dordrecht.
- Olsen, M. B. (1994). *The Semantics and Pragmatics of Lexical and Grammatical Aspect*. <https://doi.org/10.4324/9781315052267>
- Oversteegen, L. (2005). Causality and tense - Two temporal structure builders. *Journal of Semantics*, 22(3), 307–337. <https://doi.org/10.1093/jos/ffh021>
- Parsons, T. (1990). *Events in the Semantics of English: A Study in Subatomic Semantics*. MIT Press.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 97–134). Stanford, CA: CSLI Publications.
- Reinhart, T. (1984). Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics*, 22(6), 779–810. <https://doi.org/10.1515/ling.1984.22.6.779>
- Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event-indexing model. *Memory & Cognition*, 31(8), 1284–1292. <https://doi.org/10.3758/bf03195811>
- Schmerling, S. F. (1975). Asymmetric conjunction and rules of conversation. In P. Cole & J. L. Morgan (Eds.), *Speech Acts* (pp. 211–231). https://doi.org/10.1163/9789004368811_009
- Smith, C. S. (1999). Activities: States or Events? *Linguistics and Philosophy*, 22(5), 479–508.
- Stowell, T. (2007). The syntactic expression of tense. *Lingua*, 117(2), 437–463. <https://doi.org/10.1016/j.lingua.2005.08.003>
- Sun, Z., Firestone, C., & Hafri, A. (2023). How to build a scene: Relational representations are constructed in a canonical order. *Journal of Vision*, 23(9), 4750. <https://doi.org/10.1167/jov.23.9.4750>
- Tai, J. H. Y. (1983). Temporal sequence and Chinese word order. In J. Haiman (Ed.), *Iconicity in Syntax: Proceedings of a Symposium on Iconicity in Syntax* (pp. 49–72). Stanford: John Benjamins Publishing Company.
- Talmy, L. (1975). Figure and Ground in Complex Sentences. *Annual Meeting of the Berkeley Linguistics Society*, 1, 419–430. <https://doi.org/10.3765/bls.v1i0.2322>
- Taylor, B. (1977). Tense and Continuity. *Linguistics and Philosophy*, 1(2), 199–220. Retrieved from <http://www.jstor.org/stable/25000953>
- ter Meulen, A. (1995). *Representing time in natural language: The dynamic interpretation of tense and aspect*. MIT Press.

- Thomson, A. J., & Martinet, A. V. (1985). *A Practical English Grammar*. Oxford, UK: Oxford University Press.
- Vendler, Z. (1967). Verbs and times. In Z. Vendler (Ed.), *Linguistics and Philosophy* (pp. 97–121). <https://doi.org/10.7591/9781501743726-005>
- Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects*. <https://doi.org/10.1007/978-94-017-2478-4>
- von Stechow, A. (1995). On the Proper Treatment of Tense. *Semantics and Linguistic Theory*, 5, 362. <https://doi.org/10.3765/salt.v5i0.2700>
- Wårvik, B. (2004). What is foregrounded in narratives? Hypotheses for the cognitive basis of foregrounding. In T. Virtanen (Ed.), *Approaches to cognition through text and discourse* (pp. 99–148). <https://doi.org/10.1515/9783110892895.99>
- Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX)*. <https://doi.org/https://doi.org/10.17605>
- Zlatev, J. (2010). Spatial Semantics. In *The Oxford Handbook of Cognitive Linguistics*. <https://doi.org/10.1093/oxfordhb/9780199738632.013.0013>
- Zuckermann, G. (2006). *Complement clause types in Israeli*. Oxford University Press.
- Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning Memory and Cognition*, 22(5), 1196–1207. <https://doi.org/10.1037/0278-7393.22.5.1196>