

Transparency in the processing of temporal ambiguity: The case of embedded tense

Giuliano Armenante, Vera Hohaus & Britta Stolterfoht*

Abstract. We report the results of one acceptability rating study and two self-paced reading studies on the form-meaning mismatch in the interpretation of past-under-past in complement clauses in English. Across the three experiments, we find an off-line and on-line preference for the backward-shifted interpretation, in line with predictions of the structural approach to the ambiguity when assuming a processing preference for morphological transparent interpretation.

Keywords. tense semantics; consecutio temporum; sequence-of-tense parameter; form-meaning mismatches; ambiguity; processing strategies; comprehension

1. Introduction. In English, embedded tenses in certain configurations give rise to ambiguities, most prominently in the case of past tense in a stative complement clause embedded under a past-marked verb of reported speech. This past-under-past configuration is illustrated in (1). The sentence is compatible with two different direct utterances made by Oliver, indicated in (1-a) and (1-b), which correspond to a backward-shifted reading (under which Amber’s illness pre-dates Oliver’s utterance) and a simultaneous reading (under which her illness temporally overlaps with his utterance). Generalising broadly, structural approaches to this ambiguity treat SIM as derived from BACK by additional morpho-syntactic technology that allows the lower past tense not to be interpreted as such (prominently, Ogihara 1989, 1996, Stowell 1996, Kusumoto 1999, 2005).

- (1) Oliver said that Amber was sick.
- a. Oliver said: “Amber was sick.” (backward-shifted reading, BACK)
 - b. Oliver said: “Amber is sick.” (simultaneous reading, SIM)

We sketch one possible implementation of such a view in (3), where the lower PAST-operator that we see in the Logical Form for BACK is deleted to derive SIM, assuming that the use of past-tense morphology on the embedded verb can also be licensed by the PAST-operator in the matrix clause. In the absence of such an operator in the embedded clause and in interaction with the semantics of *say* in (2), Amber’s illness and Oliver’s utterance are interpreted to share an evaluation time.

- (2) For any possible world $w \in D_w$, time $t \in D_i$, individual $x \in D_e$
and tensed proposition $p \in D_{\langle s, \langle i, t \rangle \rangle}$, $\llbracket say_{\text{(simplified)}} \rrbracket(w)(t)(p) = 1$ if and only if
in all worlds w' compatible with x ’s utterance at t in w , $p(w')(t) = 1$.

*This research was funded by the Deutsche Forschungsgemeinschaft DFG (Project B1, Tübingen Collaborative Research Centre 833, ID #75650358). The research underwent ethical review by the Tübingen Ethics Committee (reference #2017_0328.53) and the Proportionate University Ethics Committee at the University of Manchester (reference #s 2018-5463-7466, 2019-5463-10129). We would like to thank the audience of the 2019 Tübingen “Processing Tense” workshop as well as Petra Augurzky, Núria Barrios-Jurado, Sigrid Beck, Maximilian Berthold, Chen-An Chang, Lilian Gonzalez Rodriguez, Robin Hörnig, Zahra Kolagar, Alina Mclellan, Anne Mucha, Paul Stott, Rolf Ulrich, and Malte Zimmermann. Authors: Giuliano Armenante, Universität Potsdam (armenante@uni-potsdam.de), Vera Hohaus, The University of Manchester (vera.hohaus@manchester.ac.uk) & Britta Stolterfoht, Eberhard Karls Universität Tübingen (britta.stolterfoht@uni-tuebingen.de).

- (3) a. Logical Form for BACK:
 [PAST_{t*} [λt Oliver say_{w@,t} [λw λt' PAST_{t'} [λt'' Amber sick_{w,t''}]]]],
 where t* is the utterance time and w@ the actual world
- b. Logical Form for SIM:
 [PAST_{t*} [λt Oliver say_{w@,t} [λw ————— [λt'' Amber sick_{w,t''}]]]]

In this paper, we investigate the processing predictions of this analysis. More generally, understanding how the processing system handles this particular ambiguity can also inform our understanding of the strategies the processing system adopts when there is no one-to-one mapping between form and meaning and multiple interpretations are available.

2. Previous research. While considered a “touchstone for the adequacy of the semantics for tense” (von Stechow 2009; 3), embedded tenses have received relatively little attention in the processing literature. While we find the experimental results in Dickey (2000, 2001)’s landmark study to be overall inconclusive, they have been taken to point towards a processing preference for SIM. The data from one of the adult-control groups in Hollebrandse (2000)’s research on first language acquisition also suggests a slight acceptability preference for SIM. Gennari (2004) employs a design that relies on additional manipulations, but also observes an advantage for overlapping temporal intervals in reading times. Mucha et al. (2022) in their comparison of English and Polish, however, observe higher acceptability for BACK compared to SIM for both languages.

3. Research questions and experimental hypotheses. When it comes to processing, the structural approach outlined in the introduction can plausibly be taken to predict a processing preference for BACK over SIM, given that an additional operation derives the former from the latter to allow the past tense morphology in the embedded clause not to be interpreted. More abstractly, this apparent mismatch between form and meaning is plausibly dispreferred. We translate these considerations here into the experimental hypothesis H1 in (4).

- (4) Hypothesis H1 “What you see is what you get” (WYSIWYG):
 Comprehension is driven by morphological transparency. A past tense should initially always be interpreted as such, favouring BACK over SIM.

Structurally, however, the resulting Logical Form underlying SIM can also be argued to be simpler than the structure that derives BACK, which may result in a processing preference for SIM. We summarise such a view in (5), as our alternative experimental hypothesis H2.

- (5) Hypothesis H2 “Representational Simplicity”:
 Comprehension is driven by structural simplicity at Logical Form. A past tense in the relevant configuration should initially not be interpreted, favouring SIM over BACK.

In the next section, we report two sets of experiments designed to test these two hypotheses using reading times and acceptability judgments. In general, we expect a processing preference to be reflected in shorter reading times and higher acceptability ratings.

4. Experimental data.

4.1. EXPERIMENTS 1 AND 2. The first two experiments investigated the two hypotheses for the type of configuration from the introduction, the temporal ambiguous interpretation of a clausal

complement of a verb of saying with a past-marked, stative predicate. The two experiments share design and materials, with Experiment 1 a reading-time study and Experiment 2 a rating study, where a preceding context disambiguated between readings.

Design. Both experiments adopted a 2x2 factorial design, with factors EVALUATION TIME (*past* vs *future*) and intended INTERPRETATION (*BACK* vs *SIM*). EVALUATION TIME was manipulated within the visually presented context; see Figure 1 for an example. The factor INTERPRETATION relates to the target items presented. A sample item from the two experiments is in (6), where the vertical bar indicates the presentation segments for Experiment 1.

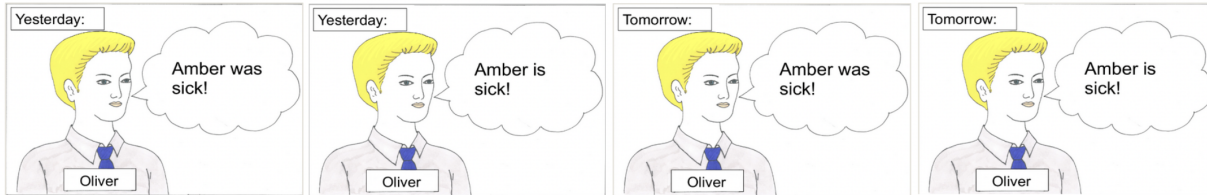


Figure 1: Sample visual contexts for Experiments 1 and 2

- (6) a. **Context:** Oliver, yesterday: “Amber was sick!” *past, BACK*
Target: Oliver | said | that | Amber | was | sick.
- b. **Context:** Oliver, yesterday: “Amber is sick!” *past, SIM*
Target: Oliver | said | that | Amber | was | sick.
- c. **Context:** Oliver, tomorrow: “Amber was sick!” *future, BACK*
Target: Oliver | will say | that | Amber | was | sick.
- d. **Context:** Oliver, tomorrow: “Amber is sick!” *future, #SIM*
Target: Oliver | will say | that | Amber | was | sick.

Of particular interest for our research questions are the two *past* conditions in (6-a) and (6-b), for which an ambiguous past-under-past sentence is evaluated in a context that establishes *BACK* or *SIM*. Note however also that simultaneous readings are not available when embedding a past-marked stative predicate under a future-marked verb of saying. As a result (and as indicated by # above), there is a mismatch between context and target for condition (6-d), which was designed to act as a control condition.

Predictions. Of the hypotheses formulated earlier, H1 “WYSIWYG” predicts a processing advantage for *BACK*, while H2 “Representational Simplicity” predicts an advantage for *SIM*. For the first experiment, we expect an effect to arise on the auxiliary or possibly at the spillover region (= the adjective), since it is at this point that the processor can establish the temporal relation between the embedded and the matrix clause and align it with the preceding context. (We assume that the the temporal information retrievable from the context is used for ambiguity resolution; see also, for instance, Trueswell & Tanenhaus 1991.) Under H1, we expect longer reading times for the non-transparent *past, SIM* compared to *past, BACK*. H2 predicts that the reading times in the critical region a shorter for the structurally simpler *past, SIM*, compared to *past, BACK*. For the second experiment, H1 predicts higher acceptability ratings for *past, BACK* compared to *past, SIM*, while we expect *past, SIM* to be rated better than *past, BACK* under H2.

4.1.1. METHODS.

Materials. We constructed 48 sets of experimental items like (6), embedded among 40 fillers, for a total of 88 trials, with each participant being exposed to 12 experimental items per condition. Each trial comprised a context picture paired with a target sentence. In order to minimize variation between items, all target sentences followed the rigid template in (7).¹

(7) NAME1 *said/will say that* NAME2 *was* ADJECTIVE.

Fillers contained reported speech constructions similar to the experimental items but targeting a different kind of aspectual ambiguity. We also designed 24 comprehension question relating to the names of the characters involved or the time of the story.

Participants. Participants ($N_{\text{Experiment 1}} = 43$, $N_{\text{Experiment 2}} = 40$) were all native speakers of English. Five participants from the initial 48 were excluded from Experiment 1 because their accuracy in answering comprehension questions was below the 75%-threshold. Participants received a cash reimbursement for their participation.

Procedure. Experiment 1 was a moving-window self-paced reading study, while Experiment 2 was an acceptability rating study. Across experiments, participants first saw the context picture, followed by the target sentence. This was read word-by-word in the first experiment, while in the second experiment, participants were asked to rate the sentence based on its fit with the context picture on a scale from 1 “bad fit” to 6 “good fit”. Of the trials, 3/11 ended with a comprehension question. Each experimental session started with written instructions, followed by two practice trials containing relative clauses, and then two blocks of the 88 trials with a break halfway through. Out of the 88 trials, four lists were formed following a Latin Square design, with stimuli presented in a randomised fashion. An average session took approximately 30 minutes, with participants tested individually at the University of Manchester Psycholinguistics Laboratory.

4.1.2. RESULTS AND DISCUSSION. The data analysis was conducted in the SPSS (Statistical Package for the Social Sciences) programming environment, performing a repeated-measures analysis of variance (ANOVA). In our exploratory analyses, we speak of a significant effect only if the probability of making a Type I error (α error) is below 5%.

Reading times. Reading times were corrected for outliers first, computing only values within the 90-1,300 ms region. In a second step, reading times with 2.5 standard deviations away from the mean were excluded. The resulting mean values across all sentence regions are in Figure 2. Mean reading times for the two regions of interest, the embedded auxiliary and the adjective that followed it, are listed in Table 1. The results show an effect of EVALUATION TIME for both regions of interest. At the auxiliary (that is, the embedded *was*), future attitude reports took significantly longer to read compared to past ones ($F_{1,42} = 19.5$, $p < .001$, $\eta^2 = .318$). The same effect was found at the adjective ($F_{1,42} = 9.6$, $p < .005$, $\eta^2 = .186$). By contrast, an effect of INTERPRETATION only reached significance at the adjective, with SIM evoking longer reading times than BACK ($F_{1,42} = 4.4$, $p < .05$, $\eta^2 = .096$), with no significant interaction between the two factors. Zooming in on the

¹If NAME1 was female, then NAME2 name was male, with the number of female and male names kept equal. Only negative polar adjectives like *furious*, *late*, and *drunk* featured in the stimuli in all three experiments, as they were deemed to not bias against one reading over the other in the relevant configuration.

two *past* conditions, we performed a *t*-test for EVALUATION TIME = *past*. While no difference was found at the auxiliary ($t(42) = 1.187, p = .242$), the effect approached significance at the adjective ($t(42) = 1.939, p = .059$), with BACK associated with faster reading times than SIM.

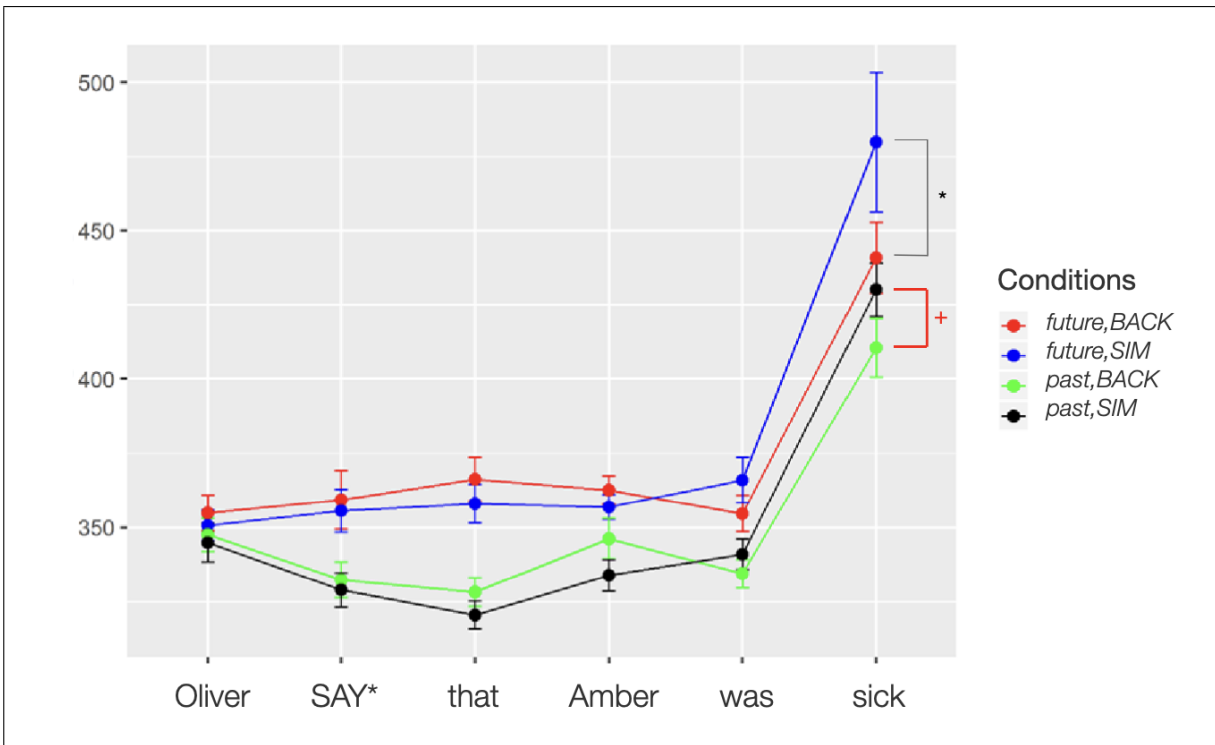


Figure 2: Mean reading times in ms for Experiment 1, with SAY* = said/will say

Region	Condition	Mean reading time in ms	Standard deviation in ms
Embedded auxiliary	<i>past,BACK</i>	334.4	95.4
	<i>past,SIM</i>	340.9	105.9
	<i>future,BACK</i>	354.6	99.1
	<i>future,SIM</i>	365.9	128.0
Adjective	<i>past,BACK</i>	410.5	111.1
	<i>past,SIM</i>	430.1	126.7
	<i>future,BACK</i>	440.9	112.8
	<i>future,SIM</i>	479.8	200.8

Table 1: Mean reading times for Experiment 1 with standard deviations for regions of interest

Acceptability ratings. Mean acceptability scores are listed in Table 2.²

	<i>past</i>	<i>future</i>
<i>BACK</i>	5.98	5.54
<i>SIM</i>	5.38	4.16

Table 2: Mean values of participants’ ratings for Experiment 2 (based on a scale from 1 “bad fit” to 6 “good fit”)

Overall, the *future, SIM* control condition in Experiment 2 received the lowest (but still relatively high) ratings, which show that the processor was sensitive to the temporal mismatch in this condition (despite the lack of a significant interaction in Experiment 1). By contrast, past-under-past was rated the highest when the intended reading was backward-shifted (in the *past, BACK* condition), but the other interpretation (in the *past, SIM* condition) still averaged a high rating. The statistical analysis revealed an effect of EVALUATION TIME ($F_{1,47} = 244.9, p < .001, \eta^2 = .839$), with *future* conditions rated significantly worse than *past*. Furthermore, we observe an effect of INTERPRETATION, in that participants rated *BACK* conditions significantly higher than *SIM* conditions ($F_{1,47} = 107.1, p < .001, \eta^2 = .695$). The interaction between EVALUATION TIME and INTERPRETATION is significant ($F_{1,47} = 65.7, p < .001, \eta^2 = .583$), with *BACK* rated significantly higher than *SIM* for future compared to past evaluation times. Upon closer inspection, the difference between the two interpretations was also significant in the *past* conditions ($t(39) = 4.921, p < .001$), pointing to a preference for the backward-shifted reading.

Discussion. We cautiously take these results to indicate a processing preference for the morphologically transparent, backward-shifted interpretation of past embedded under past in complement clauses. The results thus lend initial support not only to an analysis of embedded tenses that derives the simultaneous reading from the backward-shifted reading, but also in favour of a WYSIWYG processing strategy, as outlined in H1.

One potential challenge for this conclusion relates to a potential priming effect, however: For *past, BACK* and *future, BACK* conditions, both the visual context as well as the target item contained the same form of the auxiliary *to be* (that is, *was*), while *SIM* conditions involved the same lemma but different lexemes (that is, *is* in the visual context and *was* in the stimulus). We cannot exclude the possibility that this property of the design may have facilitated the processing of *BACK* conditions over *SIM* conditions. The design of the experiment we present in the next section sought to address this concern.

4.2. EXPERIMENT 3. We further tested H1 “WYSIWYG” and H2 “Representational Simplicity” in a third experiment, in a design where the embedded past was locally ambiguous and the intended interpretation was not already retrievable from the preceding context. The design was intended to establish whether any processing preferences exists when it comes to *BACK* and *SIM* and to address the methodological concerns raised in the discussion of Experiments 1 and 2. It was also intended to provide insight into some aspects of how the ambiguity affects the incremental composition of temporal meaning in these cases.

²Due to a coding error, five experimental items had to be removed from the analysis.

4.2.1. DESIGN. The experiment tested ambiguous and non-ambiguous past complement clauses with different continuations, which served as disambiguation in the case of ambiguous past. The experiment adopted a 2x2 factorial design, with factors AMBIGUITY (+*amb* vs -*amb*) and EVENT (*event1* vs *event2*), where the first event temporally preceded the second, and the second event coincided with the time at which the matrix verb was interpreted. We will explain the rationale behind this design in more detail for the sample item in (8), where *the rehearsal* and *the concert* make reference to the two events in question.

- (8) a. **Context:** After last week’s final rehearsal, last night, +*amb, event1*
 John’s band finally gave a concert, where I spoke to him about Mary... (= BACK)
Target: John | said | that | Mary | was sick, |
 so | that’s why | she | missed | **the rehearsal** | with | great | regret.
- b. **Context:** After last week’s final rehearsal, last night, +*amb, event2*
 John’s band finally gave a concert, where I spoke to him about Mary... (= SIM)
Target: John | said | that | Mary | was sick, |
 so | that’s why | she | missed | **the concert** | with | great | regret.
- c. **Context:** The other day John’s band taped their final rehearsal. -*amb, event1*
 And when this morning I spoke to John about Mary...
Target: John | thinks | that | Mary | was sick |
 and | that’s why | she | missed | **the rehearsal** | with | great | regret.
- d. **Context:** The other day John’s band finally gave a concert. -*amb, event2*
 And when this morning I spoke to John about Mary...
Target: John | thinks | that | Mary | was sick |
 and | that’s why | she | missed | **the concert** | with | great | regret.

In the +*amb* conditions, the context was set up in such a way as to introduce both the concert and the rehearsal that preceded it, and established that John and the speaker had a conversation at the concert. As a result, the target sentence without the *that’s why*-continuation in (8-a) and (8-b) is ambiguous, with the embedded clause plausibly allowing for both BACK and SIM. This ambiguity is resolved by *the rehearsal* or *the concert* in the continuation, with the former only compatible with BACK and the latter only with SIM. Such an ambiguity does not arise in the -*amb* conditions, where the matrix verb is in the present tense; past-under-present embeddings do not allow for multiple temporal interpretations. The -*amb* conditions served as baseline conditions that allowed us to explore the time course of the ambiguity resolution and to control for effects of length and frequency associated with the nouns used for disambiguation.

Predictions. The design of the experiment was intended to allow us to explore the preferred temporal interpretation of ambiguous past (as reflected in reading times), but also the time course of ambiguity resolution, that is, when the processing system decided on the temporal interpretation of the locally ambiguous complement clause. Does the processor delay committing to one of the interpretations as long as possible, in an attempt to avoid a potential revision later down the line? Or is there a default preference for one reading that will allow for an interpretation to be determined soon as possible? From these questions, there derive three regions of interest for this experiment, namely, the embedded predicate (that is, *was sick* in (8) above), the disambiguating region (*the*

rehearsal or the concert, for instance), and a spillover region, which was defined as the disambiguating region +1. Under H1 “WYSIWYG”, we expect an interaction at the disambiguating or the spillover region, with *event2* evoking longer reading times than *event1* in the +*amb* conditions but not in the -*amb* conditions. H2 “Representational Simplicity” predicts shorter reading times for the *event2*, +*amb* condition (= SIM), compared to the *event1*, +*amb* condition (= BACK). In relation to the incrementality of the composition of the temporal meaning in these sentences, any such difference could either be a result of the processing system revising a previously assigned interpretation or the result of a dispreferred temporal interpretation that has been delayed but is now forced by the disambiguation. As for the embedded predicate, if the ambiguity goes unnoticed and the processing system commits to a temporal interpretation right away (likely, the preferred interpretation), we may not expect an effect of AMBIGUITY. If both interpretations that are available at this point undergo more active consideration, we expect an increase in reading times for +*amb* conditions over -*amb* conditions.

4.2.2. METHODS.

Materials. 16 experimental items like (8) were constructed alongside 48 fillers, to yield a total of 64 trials. Trials included a context sentence followed by a target sentence. Of the trials, 3/8 ended with a comprehension question. Disambiguated nouns were carefully paired so that they would depict events with a stereotypical temporal order (as in the case of a rehearsal and a concert), to facilitate the retrieval of the intended temporal interpretation. As in the first two experiments, the target sentences followed a fairly rigid template, with *said* in all the +*amb* conditions and *thinks* in all the -*amb* conditions. Filler items exhibited a temporal or an aspectual local ambiguity, with a disambiguating continuation segment that was either a consecutive sentence similar to the one used for the experimental items or a relative clause.

Participants. 78 participants, native speakers of English, were recruited via the platform Prolific. Ten participants were excluded as they scored below 70% in the comprehension task, bringing the total number of participants to 68. Participants were reimbursed for their participation.

Procedure. The experiment adopted a self-paced reading task, with contexts and targets displayed entirely masked on the screen together (see (8) above for the segmentation adopted), but separated by line breaks. Three practice trials preceded the experimental session, which was split into four blocks. Participants had the opportunity to take a break at the end of each trial. Four lists were created following a Latin square design, with stimuli presented in a randomised fashion. An experimental session took approximately 25 minutes to complete online.

4.2.3. RESULTS AND DISCUSSION.

Reading times. Results were analysed as in Experiment 2.³ Reading times were corrected for outliers, trimming values below 1,200 ms for the context and above 1,500 ms for the disambiguating region in the continuation. Reading times were then log-transformed, with absolute values below a standard deviation of 3 from the mean being removed. The reading times computed for the two regions of interest to the temporal interpretation are reported in Table 3. No effect of AMBIGUITY or EVENT was found at the disambiguating region. Although not reaching significance

³Due to a programming error, two items had to be removed from the analysis.

($F_{1,67} = 3.127, p = .082, \eta^2 = .045$), we found a trend towards an interaction, with faster reading times for *event1* (= the backward-shifted interpretation) than *event2* (= the simultaneous interpretation) in the +*amb* condition compared to the -*amb* baseline; see also Figure 3. No effect was observed at the spillover region ($F_{1,67} = .729, p = .396, \eta^2 = .011$). Lastly, there was no effect of AMBIGUITY at the embedded predicate region (for the example in (8), *was sick*). There is thus no evidence that locally ambiguous sentences slowed down the processing system at the first point where it would have been possible to assign a temporal interpretation to the sentence and where the temporal ambiguity would have been detectable.

Region	AMBIGUITY	EVENT	Mean
Disambiguation	+ <i>amb</i>	<i>event1</i> (= BACK)	5.878
		<i>event2</i> (= SIM)	5.928
	- <i>amb</i>	<i>event1</i>	5.922
		<i>event2</i>	5.869
Spillover	+ <i>amb</i>	<i>event1</i> (= BACK)	5.991
		<i>event2</i> (= SIM)	5.990
	- <i>amb</i>	<i>event1</i>	5.973
		<i>event2</i>	5.929

Table 3: Log-transformed mean reading times for regions of interests

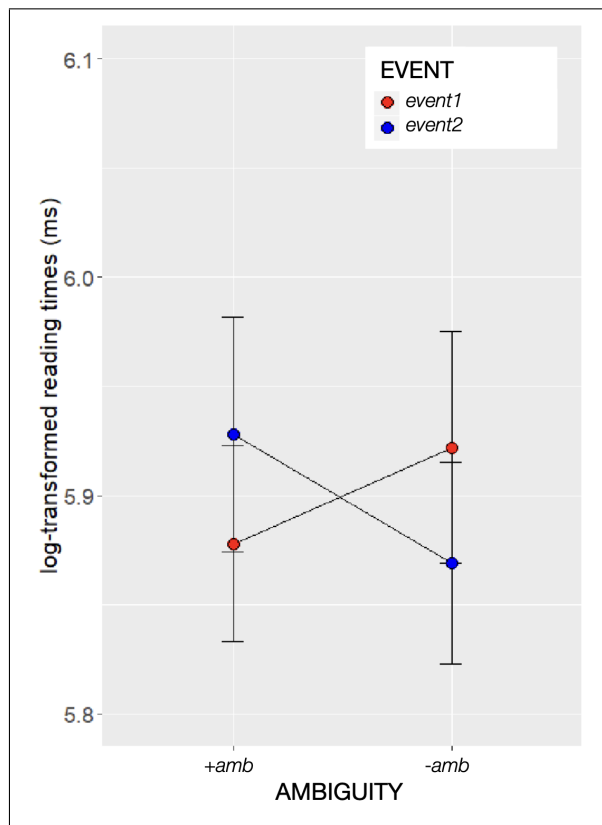


Figure 3: Reading times of participants for Experiment 3 at the disambiguating region

Discussion. Overall, these findings are compatible with the results from the previous two experiments and with a general processing preference for backward-shifted interpretations over simultaneous ones, in line with H1 “WYSIWYG”. In relation to the on-line composition of the temporal interpretation, in the absence of an effect at the embedded predicate, it appears that the ambiguity did not slow down the processing system. It is plausible that the preferred interpretation was assigned as soon as possible and later revised if incompatible with the continuation.

Before we turn to the overall discussion, let us briefly comment on the lack of a significant interaction between AMBIGUITY and EVENT in this experiment, especially given the trend observed in the data. One possible explanation for this may be design-related, the other may be related to more general cognitive principles: In relation to the design, the sample size was likely not large enough to produce sufficient statistical power, given the small effect size ($\eta^2 = .45$) and the challenge of constructing a higher number of good stimuli with the adopted design. From a cognition perspective, it has been observed that temporal remoteness is harder to process than temporal proximity, in that events that are temporally distant come with a processing cost compared to those closer to the speech time (see, for instance, Gennari 2004). BACK involves an event that is temporally further away from both the utterance time to which the sentence is anchored and the time at which the matrix predicate is interpreted than the event involved in SIM (that is, the rehearsal precedes the concert precedes the utterance time in the example item). Since the events features prominently in the items, we may want to speculate that this cognitive bias was a confounding effect and curbed the magnitude of the expected interaction. A second, related confound might also have played an inhibiting role: More recently mentioned referents are easier to recover, and thus to process. Within the design of Experiment 3, these were by choice always noun phrases that would force SIM. We opted against presenting the two events in a non-chronological order in the context for half of the trials to preserve naturalness and readability of the stimuli.

5. Overall discussion. Taken together, the three studies presented here provide initial evidence for an on- and off-line preference for BACK over SIM in the interpretation of ambiguous past-under-past. Taken together, we cautiously take these results to provide support for H1 “WYSIWYG” and a processing strategy that favours morphological transparency over structural simplicity to resolve form-meaning mismatches. These results challenge the often implicitly assumed preference for SIM, but are in line with more recent findings in Mucha et al. (2022). Based on our findings in Experiment 3, we have tentatively suggested that BACK may constitute a default interpretative strategy during on-line comprehension and that the ambiguous past in complement clauses does not cause the processing system to delay the composition of the temporal meaning of the sentence. These are preliminary findings, however, which should be investigated further in an adequately controlled experimental setting.

While a preference for BACK is not compatible with H2 “Representational Simplicity” as discussed in Section 3, the structurally simpler Logical Form for SIM may not necessarily translate to a processing advantage over BACK if we discuss the two structural representations as a type of filler-gap constellation, with the embedded tense operator a filler and the temporal variable associated with the embedded predicate the gap. Processing difficulty in filler-gap constellations is often measured in terms of the linear or structural distance between the filler and the gap, with

increased distance in a dependency increasing processing load.⁴ Assuming for the Logical Form underlying SIM a reduced structure as in (3), which lacks a tense operator (that is, filler) in the embedded clause, the temporal variable (that is, the gap) associated with the embedded predicate will be bound by the matrix predicate (and in turn the matrix PAST), leading to a non-local chain. By contrast, the Logical Form underlying BACK is characterised by a local chain, which in return should translate into a processing advantage for BACK over SIM. Given these considerations, no preference would be expected for SIM, even under a processing hypothesis based on the structural representations underlying the different readings.

6. Conclusion. This paper investigated the processing of a prominent form-meaning mismatch in English that relates to the temporal interpretation of sentence. More specifically, we investigated the processing of past-tensed complement clauses embedded under a past-marked verb of saying/ attitude verb, which allow for both simultaneous and backward-shifted interpretations. Overall, the results provide preliminary evidence in favour of a preference for BACK in both reading and rating tasks that prevents a processing delay in composing temporal meaning in these cases. We took these findings to provide empirical support for a processing strategy that hinges on a transparent mapping of morphological form to meaning where, by WYSIWYG, past morphology by default translates to past meaning. In the light of these results, we reasoned that competing hypotheses built on structural simplicity may have to be revised, as the reduced temporal representations underlying SIM come at the cost of a potential processing difficulty associated with their non-local filler-gap dependencies.

Beyond these findings, we hope that this paper will raise awareness for what we consider an under-researched area in semantic processing and highlight some of the methodological questions in researching the processing of embedded tense. Directions for further research not only include addressing some of these methodological challenges with different designs, but also broadening the empirical footing of the research programme to include the processing of progressive-marked past predicates, which also give rise to BACK and SIM (see, for instance, Kusumoto 1999), and a wider array of embedded environments, including relative clauses and adjunct clauses. Relative clauses in particular are compatible with a wider range of temporal configurations than complement clauses, without exhibiting structural ambiguity (see, for instance, Kusumoto 2005, Hohaus 2019). Processing data from a wider range of languages would also be welcome in future research, especially given the range of well-documented cross-linguistic variation in the interpretation of embedded tenses (see, for instance, Grønn & von Stechow 2010, Bochnak et al. 2019).

References

- Bochnak, M. Ryan, Vera Hohaus & Anne Mucha. 2019. Variation in tense and aspect, and the temporal interpretation of complement clauses. *Journal of Semantics* 36(3). 407–452. <https://doi.org/10.1093/jos/ffz008>.
- Collins, Chris. 1994. Economy of derivation and the generalised proper binding condition. *Linguistic Inquiry* 25(1). 45–61. <http://www.jstor.org/stable/4178848>.

⁴For relevant discussion and references, see Collins (1994), Hamilton (1995), O’Grady (1997), Hawkins (1999, 2004), and O’Grady et al. (2003), among many others.

- Dickey, Michael Walsh. 2000. *The processing of tense*. Amherst: University of Massachusetts dissertation.
- Dickey, Michael Walsh. 2001. *The processing of tense: Psycholinguistic studies on the interpretation of tense and temporal relations*. Dordrecht: Kluwer.
- Gennari, Silvia 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>.
- Grønn, Atle & Arnim von Stechow. 2010. Complement tense in contrast: The SOT parameter in Russian and English. In Atle Grønn & Irena Marijanović (eds.), *Russian in contrast: Grammar*, 109–153. Oslo: Oslo Studies in Language 2(1).
- Hamilton, Robert L. 1995. The noun-phrase accessibility hierarchy in SLA. In Fred R. Eckman, Diane Highland, Peter W. Lee, Jean Mileham & Rita Rutkowski Weber (eds.), *Second language acquisition: Theory and pedagogy*, 101–114. Mahwah: Erlbaum.
- Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–285. <https://doi.org/10.2307/417261>.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hohaus, Vera. 2019. The temporal interpretation of complement and relative clauses: Contrasting English and Samoan. *Journal of the Southeast Asian Linguistics Society* 12(3). 42–60.
- Hollebrandse, Bartjan. 2000. *The acquisition of sequence of tense*. Amherst: University of Massachusetts dissertation.
- Kusumoto, Kiyomi. 1999. *Tense in embedded context*. Amherst: University of Massachusetts dissertation.
- Kusumoto, Kiyomi. 2005. On the quantification over times in natural language. *Natural Language Semantics* 13(4). 317–357. <https://doi.org/10.1007/s11050-005-4537-6>.
- Mucha, Anne, Agata Renans & Jacopo Romoli. 2022. Sequence of tense and cessation implicatures: Evidence from Polish. *Natural Language and Linguistic Theory* online first. <https://doi.org/10.1007/s11049-022-09545-2>.
- Ogihara, Toshiyuki. 1989. *Temporal reference in English and Japanese*. Austin: University of Texas dissertation.
- Ogihara, Toshiyuki. 1996. *Tense, attitudes, and scope*. Dordrecht: Kluwer.
- O’Grady, William. 1997. *Syntactic development*. Chicago: The University of Chicago Press.
- O’Grady, William, Miseon Lee & Miho Choo. 2003. A subject-object asymmetry in the acquisition of relative clauses in Korean as a second languages. *Studies in Second Languages Acquisition* 25(3). 433–448. <https://doi.org/10.1017/S0272263103000172>.
- von Stechow, Arnim. 2009. Tenses in compositional semantics. In Wolfgang Klein & Ping Li (eds.), *The expression of time*, 129–166. Berlin: De Gruyter. <https://doi.org/10.1515/9783110199031.129>.
- Stowell, Tim. 1996. The phrase structure of tense. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 277–291. Berlin: Springer.
- Trueswell, John C. & Michael K. Tanenhaus. 1991. Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes* 6(4). 303–338. <https://doi.org/10.1080/01690969108406946>.