UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Two Negatives Make a Positive: Reducing Referential Uncertainty through Negation and Order Reversal Eases Processing in Counterfactuals

Permalink
https://escholarship.org/uc/item/1rx1b25r

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Evcen, Ebru
Wittenberg, Eva

Publication Date
2024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Two Negatives Make a Positive: Reducing Referential Uncertainty through Negation and Order Reversal Eases Processing in Counterfactuals

Ebru Evcen (eevcen@ucsd.edu)
Department of Linguistics, 9500 Gilman Drive 0108
San Diego, CA 92093 USA

Eva Wittenberg (wittenberge@ceu.edu)
Department of Cognitive Science, 1025 Quellenstraße 51, A-1100
Vienna, Austria

Abstract

Counterfactual statements are famously difficult to process, and so are negated sentences and infrequent clause orders. Here, we argue that their combination can ease much of the processing cost when these difficult constructions align to clarify what is being referred to, thereby reducing referential uncertainty. In Experiment 1, we tested how affirmative and negative counterfactual statements (e.g., If there had been (no) zebras, there would have been (no) lions) are interpreted using a web-based eye-tracking paradigm. We found that negation facilitates processing, particularly when a Question under Discussion is about the actual state of affairs. In Experiment 2, reversing the clause order resulted in easier comprehension. These results provide support for a model of incremental language processing that puts the construal of semantic representations front and center.

Keywords: counterfactual interpretation; implicit negation; QuD accommodation; referential uncertainty

Introduction

Language enables us to articulate not only the realities of the world but also hypothetical alternatives. Counterfactuals, i.e., mental representations of scenarios that diverge from past events or states, are a prime example of this capability. We investigate how listeners navigate the dynamics of counterfactuality and arrive at an interpretation during sentence processing.

Imagine hearing: "If there had been zebras, there would have been lions in the zoo" This counterfactual statement invites the listener to envisage a hypothetical scenario with the existence of zebras and lions, but simultaneously infer that, in reality, neither zebras nor lions were present. The understanding of such utterances involves the mental juxtaposition of the hypothetical world, where these animals were present, with the implied actual one, where these animals were absent (e.g., Byrne, 2024; Byrne & Tasso 1999; Johnson-Laird & Byrne, 2002). Listeners draw upon both the hypothetical and implied actual state to arrive at the implied actual state interpretation for successful communication. However, identifying the exact animals in the implied actual state remains a challenge since the referential set of the actually present animals is very broad: Apart from zebras and lions, there is no information about the animals in the zoo. Here we ask how this referential uncertainty interacts with the way counterfactuals are processed in real time.

Empirical work indicates that dual meanings in counterfactual utterances, both the actual and hypothetical content, are accessible to comprehenders (e.g., Fillenbaum, 1974; Thompson & Byrne, 2002; Quelhas et al., 2018). The interesting question regarding online sentence processing is therefore not whether counterfactuals can convey a dual meaning, but whether they *always* do and how exactly such dual meaning relates to an incremental build-up of the sentence meaning (see for a review, Kulakova & Nieuwland, 2016).

Some studies demonstrate the rapid representation of the implied actual state (Santamaria et al., 2005; Ferguson & Sanford, 2008; Ferguson, 2012; de Vega et al., 2007; de Vega & Urrutia, 2012) whereas others indicate that listeners only represent the hypothetical state (Ferguson et al., 2009; Nieuwland & Martin, 2012; Nieuwland, 2013). This divergence in findings may partly be explained by the 'Question under Discussion' (QuD) (Roberts, 1996; 2004) inherent in the stimuli. Decoding what others imply often involves inferring their intentions (Grice, 1975; Sperber & Wilson, 2002). Consider the statement from before: 'If there had been zebras, there would have been lions', which implies the absence of both animals. However, this inference is not fixed and varies depending on the QuD: When the QuD revolves around describing an ideal zoo visit, the interpretation of the statement may shift. In such contexts, the focus might lean more toward the hypothetical state where zebras and lions were present. That is, when the QuD leans towards hypothetical states, these scenarios are likely to be represented more vividly, and vice versa (Evcen & Wittenberg, 2022).

The idea that there is a tight connection between the interpretation of an utterance and the communicative goals has been well established (Clifton & Frazier, 2012; Ronai & Xiang, 2021; Degen, 2013; Degen & Tanenhaus, 2015; Ippolito, 2013). The QuD can guide listeners' expectations for upcoming content and aid in extracting meaning from an utterance. Here we test two other cases that facilitate the incremental assembly of mental representations during

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

counterfactual processing: explicit negation and clause order. The central idea is that the complexity involved in understanding counterfactuals is largely due to the ambiguity in reference during their processing, i.e., how easily QuD is accommodated. That is, the QuD accommodation happens incrementally and automatically, which either triggers the representation of the hypothetical state and contributes to the extra processing cost of arriving at the implied actual state interpretation or facilitates implied actual state interpretation without additional cost (see for a similar discussion in negation processing, Dale & Duran, 2011; Tian & Breheny, 2016; Tian et al., 2016).

We first ask whether negated counterfactuals are easier or more difficult to understand than affirmative ones. The two-step views of negation suggest that negated counterfactuals should have longer processing times than affirmative ones. Under this view, understanding a negative statement typically involves mentally constructing the affirmative scenario first, which is an effortful process (Carpenter & Just, 1975; Kaup et al., 2007; Lüdtke et al., 2008). However, the application of this view to counterfactuals is not straightforward, as counterfactuals carry implicit negation with them. An alternative hypothesis is that explicit negation in counterfactuals eases the processing cost as it might facilitate the interpretation of the implied actual state. Consider the negated counterfactual "If there had been no zebras, there would have been no lions." Here, the negation creates an imaginary scenario where both zebras and lions are absent, yet simultaneously implying their presence. The negated counterfactual narrows down the interpretative scope for the listener, reducing the ambiguity inherent in the affirmative counterpart. We call this alternative hypothesis the QuD-accommodation hypothesis.

The second case that might aid or hinder comprehension of counterfactuals is the structural ordering of clauses. On the one hand, one hypothesis here is that the canonical order, where the *if* clause precedes the main clause, should be easier to comprehend because *if* in the antecedent clause is what implies the conditional meaning and perfect modal structure in the consequent clause is more ambiguous, such that it may imply past possibilities and necessities as well as counterfactuality. On the other hand, the QuD-accommodation hypothesis predicts that reversed order, where the main clause precedes *if* clause as in "There would have been lions if there had been zebras", aids in focusing the implied actual interpretation. This reversed order allows for immediate actual state inference triggered by modal auxiliary (would) as to what was present or what was not in this case. Conversely, the canonical order initially triggers a hypothetical world, and then the consequent is evaluated within the bounds of that hypothetical world (Evans & Over, 2004; Haigh & Stewart, 2011). This not only increases the cognitive load but also potentially deepens the engagement with the hypothetical state.

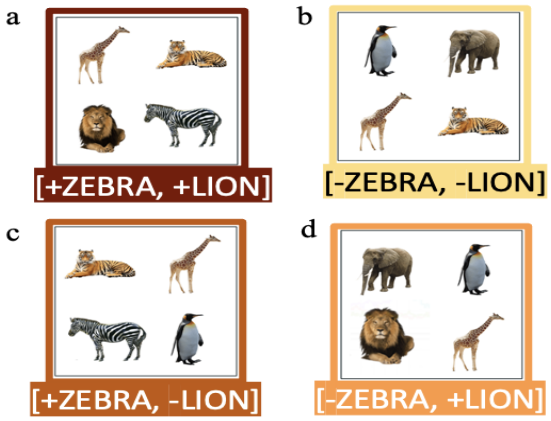

Figure 1: Example visual display in Experiment 1 & Experiment 2.

The Current Study

Our overall hypothesis is that some of the cost of interpretation of counterfactuals can be explained by referential uncertainty during processing, and when this uncertainty is reduced (through explicit negation in Experiment 1 and order reversal in Experiment 2) processing costs shrink**.** We compare the time course of eye movements during the processing of counterfactuals, using a visual world paradigm built on Evcen and Wittenberg (2022). Extending Orenes et al. (2019), they provided evidence that making the implied actual state interpretation a contextually relevant alternative via an explicit QuD shifted the interpretation from the hypothetical to implied actual interpretation. Keeping the QuD 'What was actually here?', focusing on the actual state of affairs, the same, we build on this work in three steps:

First, we created a more comprehensive visual setting than previous work (Orenes et al., 2019, Evcen & Wittenberg, 2022) and replaced the crossed-out images with a set of other related objects to represent absence and/or negation. This gave us images depicting a) the presence of both referents [+ZEBRA, +LION] and b) the absence of both referents [-ZEBRA, -LION]. Additionally, we introduced two images depicting c) the first-referent-only state [+ZEBRA, -LION], d) the second-referent-only state [-ZEBRA, +LION] (Fig. 1). This design enabled us to track participants' gaze patterns as the utterance unfolded in time, allowing for more fine-grained observations of online event construal.

Second, we tested the hypothesis that explicit negation in counterfactual reduces representational uncertainty, and eases processing cost. The QuD accommodation hypothesis predicts that participants will update their attention to the implied actual state interpretation more rapidly in negated counterfactuals than they do in affirmative counterfactuals whereas two-step views predict overall delays in the presence of explicit negation. Third, we tested whether the order in which the counterfactual is presented facilitates processing. The QuD accommodation hypothesis predicts faster looks to the implied actual state in the reversed order.

All materials, data, and analysis code for the experiments are available at https://osf.io/ueq5z.

Experiment 1: Canonical Order

Participants: 102 (N=82 post exclusions) native speakers of English participated in the study via Amazon Mechanical Turk. Using CloudResearch (Litman et al., 2017), we limited participants to those in the United States, with a task acceptance rate of 80% or higher, and a minimum of 100 tasks completed. Following Morgan et al. (2020), participants with more than 25% track loss during the approximately 8000 ms eye-tracking duration were excluded, as well as those who did not provide data for at least one trial in each of the experiment's four cells.

Design and Materials: The experiment used a 2x2 factorial design which manipulated Sentence Type (declarative, counterfactual) and Polarity (affirmative, negative) within subjects. Declarative sentences acted as our control condition. Participants were presented with 6 vignettes per condition, giving a total of 24 critical trials. Each vignette contained three sentences: a context sentence, e.g. 'While Jack was at the zoo visiting the animals with his parents, he said to his friend', followed by the critical utterance (1)-(2) and a concluding sentence (e.g., 'Finally, Jack and his family went to a restaurant to eat'). Trials featured pre-recorded computerized auditory and visual input, pairing each sentence with a visual scene of four images. Post-trial, participants answered the QuD 'What was actually there?' by clicking one of the pictures. This question was included in the instructions and repeated after each trial. Additionally, 12 filler items similar to the experimental sentences in implicit negation and inference were included (e.g., Jack loved all the animals except for zebras and lions). The order of items and image positions on-screen were randomized for each participant.

(1) If there had been zebras, then there would have been lions.
(Declarative: There were no zebras and there were no lions)

(2) If there had been no zebras, then there would have been no lions.
(Declarative: There were zebras and there were lions)

The experiment was conducted online using PennController IBEX (Zehr & Schwarz, 2018) with Webgazer.js for eye-tracking (Papautsaki et al., 2017). Participants joined remotely, using their webcams for eye-tracking. While webcam-based tracking shows higher variance and lower sampling rates, it maintains comparable accuracy to in-lab studies (Degen et al., 2020; Vos et al., 2022).

Data treatment and analyses: Fixations were analyzed in two key time windows centered on the first (zebras) and the second referent (lions), chosen for their disambiguating roles. To account for saccade planning, we offset the onset of each time window by 200ms (Hallett, 1986). To analyze behavioral responses (i.e., image selection at the end of the sentence), we ran Poisson regression models to examine variations in the count of selection on each image across different conditions, with a random intercept of participant to account for individual variances. For the eye tracking data, we used Growth Curve Analysis (GCA; Mirman et al., 2008) to examine dynamic changes in fixation proportions to the target image (implied actual state) over the competitor (hypothetical state), employing the empirical logit transformation of fixation probabilities every 50 ms (Barr, 2008), which served as our dependent variable. We conducted separate GCAs for each time window, including contrast-coded fixed effects of sentence type, polarity, and their time interactions (linear, quadratic, cubic). We fitted linear mixed effects models including random slopes for sentence type and polarity by participant and item. We report the output of the best-fitting model.

Results and Discussion: We excluded trials where participants looked at none of the images on the screen (track loss) and clicked on neither the hypothetical state nor the implied actual state image (loss of attention), leading to 25% of trials being excluded from the analyses.

For the behavioral responses, we report the proportion of selections on the target and competitor in counterfactuals in Table 1.[1] There were two groups of people: those who clicked on the hypothetical state and those who clicked on the implied actual state in both polarity conditions. Notably, the data revealed a tendency for a greater number of clicks on the hypothetical state in the affirmative condition (β=0.12, SE=0.06, z=1.86, p=.06) whereas there was no such trend in the negative condition. For the inferential analyses, we focus on the analyses where participants click on the implied actual state and discuss the data with clicks on the hypothetical state in the General Discussion.

Figure 3 shows participants' visual search patterns in the counterfactual sentences by polarity: For counterfactual affirmative sentences, participants' attention remained on images with [+ZEBRA] for around 1500 milliseconds, later transitioning to images without zebras or lions. In contrast, with counterfactual negative sentences, participants quickly directed their attention to actual state images, disregarding hypothetical ones throughout the sentence.

First referent window: We found a significant interaction between sentence type and polarity ($X^2(1)$ =164.52, p=.001), which suggests participants looked more to the critical image in counterfactual negative sentences than in counterfactual affirmative sentences (β=0.3, SE=0.12, t=2.42, p<.05). In declarative sentences, participants showed the opposite pattern: They looked more to the critical image in the affirmative condition than in the negative condition (β=-0.26, SE=0.07, t=-3.46, p<.001). The fixation patterns also varied across time. We found a three-way interaction between time terms, sentence type, and polarity ($X^2(3)$ =69.07, p=.001). To resolve this interaction, we ran separate models for each sentence type. In the counterfactual condition, participants'

[1] Due to space constraints, we report clicks only for the implied actual and hypothetical states in the counterfactual condition. Clicks on other images were minimal (<3%) across scenarios, and the control condition showed high accuracy.

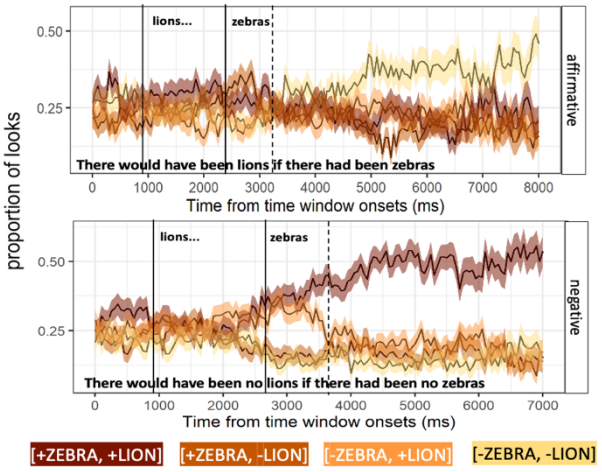

Figure 3: Fixation probabilities for counterfactual affirmative and negative sentences in Experiment 1, with standard errors shown as ribbons. Sentence end is indicated by a dotted line.

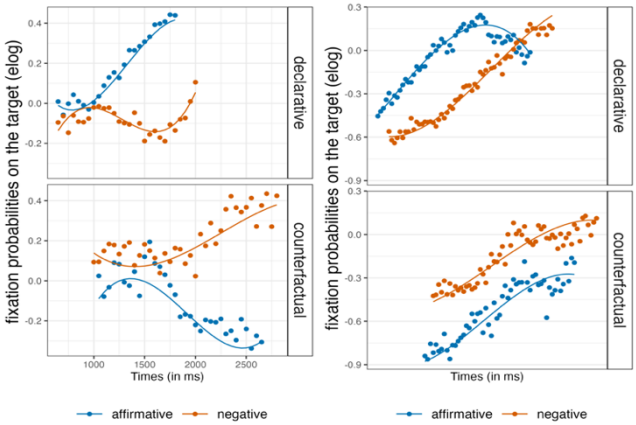

Figure 4: Fitted lines for the looks to the implied actual state by sentence type and polarity in the first referent (left) and second referent window (right) in Experiment 1.

fixations on the target image increased linearly for negative sentences (polarity*linear term, β=1.2, SE=0.16, t=7.47, p<.001). In contrast, fixations on the affirmative counterfactual sentences fluctuated more with an initial rise followed by a decrease (polarity*cubic term, β=-0.48, SE=0.16, t=-2.94, p<.01). In declarative condition, participants showed the opposite pattern: They fixated on the target image quickly in the affirmative condition (polarity*linear term, β=-0.91, SE=0.14, t=-6.34, p<.001) whereas their fixations to the target image showed a greater curvature with an initial decrease followed by an increase in the declarative negative condition (polarity*cubic term, β=0.45, SE=0.12, t=3.62, p<.001) (Fig 4).

Second referent window: We found an effect of sentence type (X^2(1) =16.75, p=.001), such that participants looked at the target image significantly more in the declarative condition compared to counterfactual condition (β=-0.23, SE=0.07, t=-3.08, p<.01). There was a significant main effect of polarity (X^2(1) =8.97, p=.01) on participants' gaze towards the target image. Specifically, participants looked at the target image more when they were listening to negative sentences compared to affirmative sentences (β=0.13, SE=0.06, t=2.32, p<.05). Additionally, there was also a significant interaction between sentence type and polarity (X^2(3) =444.66, p=.001) such that participants looked at the target image while listening to counterfactual negative sentences more than they did for counterfactual affirmative sentences (β=0.32, SE=0.09, t=3.38, p<.01). However, there were more fixations to the target image in declarative affirmative sentences than declarative negative sentences (β=-0.22, SE=0.06, t=-3.41, p<.01). Additionally, we found a three-way interaction between time terms, sentence type, and polarity (X^2(3) =111.01, p=.001). For counterfactual sentences, fixations on the target image showed a linear increase in both affirmative and negative counterfactuals (β=1.55, SE=0.09, t=15.98, p<.001). As for the declarative sentences, there was a linear increase when participants were listening to declarative negative sentences (polarity*linear term, β=1.48, SE=0.15, t=9.7, p<.001) and there was a reversed U-shaped pattern for declarative affirmative sentences (polarity*quadratic term, β=1.57, SE=0.15, t=10.4, p<.001) (Fig 4).

Exp. 2: Reversed Order

Participants: 78 (N=68 post exclusions) native speakers of English were recruited in the same way as in Experiment 1.

Design and Materials: The procedure and tasks were the same as in Exp. 1, this time reversing the clause order. The conditions again involved declarative sentences as control conditions and counterfactual sentences in affirmative and negative structure, as in (3)-(4).

(3) There would have been lions if there had been zebras.
(4) There would have been no lions if there had been no zebras.

Results and Discussion: Similar to Experiment 1, we excluded trials where participants failed to look at any images or clicked on neither the competitor nor the target image, removing about 27% of the trials due to track loss.

For the behavioral responses, there were again two groups of responders: those who clicked on the hypothetical state and those who clicked on the implied actual state in both polarity conditions (see Table 1). There was a significantly greater number of clicks on the implied actual state than on the hypothetical state in both the affirmative condition (β=0.82, SE=0.08, z=10.07, p<.001) and negative condition (β=-0.96, SE=0.08, z=-11.45, p<.001). For the inferential analyses, like in Exp. 1, we focused on the analyses where participants click on the implied actual state.

Figure 5 presents the pattern of visual searches among participants for four images in the counterfactual condition by polarity.

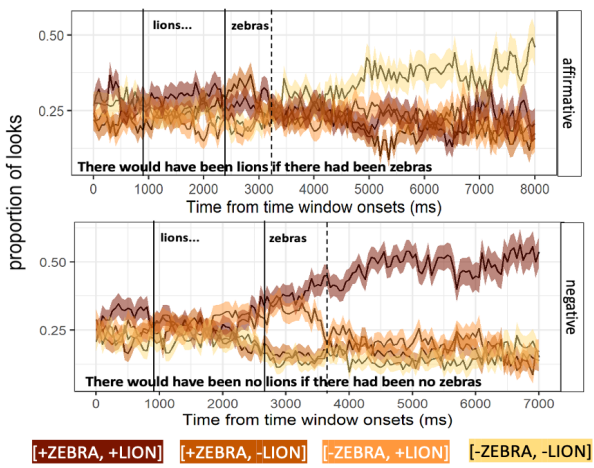

Figure 5: Fixation probabilities for counterfactual affirmative and negative sentences in Experiment 2, with standard errors shown as ribbons. Sentence end is indicated by a dotted line.

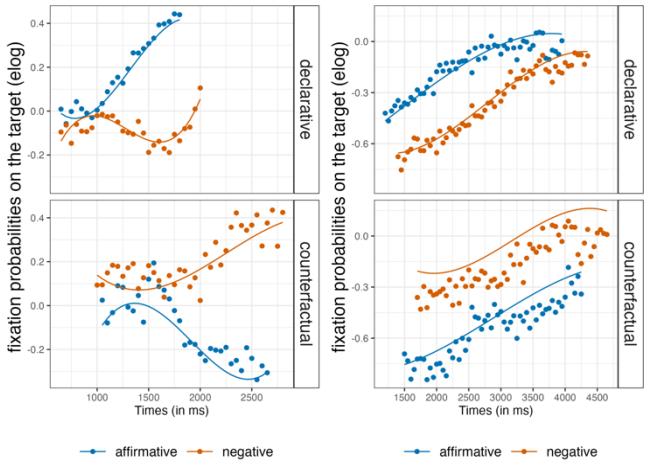

Figure 6: Fitted lines for the looks to the implied actual state image by sentence type and polarity in the first referent (left) and second referent window (right) in Experiment 2.

For the affirmative sentences, participants focused on hypothetical alternatives throughout the utterance transitioning to the actual state image only post-utterance. In contrast, in negative sentences, participants immediately focused on the actual state, disregarding the hypothetical alternatives during and after the utterance.

First referent window: We observed a significant interaction between sentence type and polarity ($X^2(1) = 100.3$, $p = .001$), indicating greater fixation on the critical image in counterfactual negative versus affirmative sentences. Conversely, in declarative sentences, this pattern reversed, with more fixation in affirmative contexts. Regarding eye fixation patterns, there was a significant three-way interaction among time terms, sentence type, and polarity ($X^2(3) = 34.03$, $p = .001$). Further analysis by sentence type revealed that in counterfactual negatives, fixation on the target image linearly increased (polarity*linear term, $\beta = 0.93$, SE=0.18, $t = 5.24$, $p < .001$) while in counterfactual affirmative sentences, fixation patterns were stable. Additionally, counterfactual negative sentences showed a U-shaped fixation pattern, with a slight decrease followed by an increase (polarity*quadratic term, $\beta = 0.38$, SE=0.18, $t = 2.14$, $p < .05$). In the declarative control condition, affirmative sentences led to quicker fixation on the target (polarity*linear term, $\beta = -0.48$, SE=0.16, $t = -2.94$, $p < .01$) while negative sentences showed an initial and subsequent increase in fixation (polarity*cubic term, $\beta = 0.52$, SE=0.14, $t = 3.67$, $p < .001$). This pattern suggests a delay in identifying the correct referent in counterfactual affirmatives and declarative negatives (Fig 6).

Second referent window: There was a significant interaction between sentence type and polarity ($X^2(1) = 485.46$, $p = .001$), showing increased target image fixation in declarative affirmative over counterfactual affirmative sentences, and more in counterfactual negative than in declarative negative sentences. Time-based differences were explored through growth curve analyses, revealing a three-way interaction ($X^2(3) = 20.467$, $p = .001$). Key findings are as follows: We found higher fixation probabilities on the target image in counterfactual negative versus affirmative sentences (main effect of polarity, $\beta = 0.30$, SE=0.08, $t = 3.45$, $p < .01$). Furthermore, in counterfactual affirmatives, there was a linear increase in target fixation over time (polarity*linear term, $\beta = -0.44$, SE=0.19, $t = -2.21$, $p < .05$) whereas fixation pattern in negatives remained constant during this window. In declarative sentences, there was more frequent fixation on the target in affirmative than in negative condition (main effect of polarity, $\beta = -0.23$, SE=0.07, $t = -3.05$, $p < .01$). Also, for affirmative declarative sentences, initial linear increase in eye gaze was followed by steadiness (polarity*quadratic term, $\beta = 0.46$, SE=0.16, $t = 2.75$, $p < .01$), while in negative sentences, fixation increased linearly (polarity*linear term, $\beta = 0.51$, SE=0.17, $t = 3.04$, $p < .01$). These patterns suggest delayed fixation on the target image in counterfactual affirmative sentences and a quicker target setting in affirmative declarative sentences (Fig 6).

Experiment 1 vs 2: The effect of clause order

We compared Experiment 1 vs 2 to assess the impact of clause order on processing, excluding declarative control sentences from the analyses as their pattern did not differ across experiments due to no structural differences.

Behavioral Responses: For the affirmative counterfactuals, the model revealed a significant effect of Clause Order on the number of clicks on the target image ($X^2(1) = 4.25$, $p < .05$) such that there were significantly more clicks on the target image in the reversed order compared to the canonical order condition ($\beta = 0.45$, SE=0.22, $t = 2.02$, $p < .05$) (see Table 1). For the negative counterfactuals, the model did not reveal a main effect of Clause Order, indicating that clicks on the target image did not differ significantly depending on the clause order.

Table 1: The proportion of clicks on target and competitor in counterfactual sentences by polarity and clause order

polarity & clause order	implied actual state [target]	hypothetical state [competitor]
canonical, affirmative	**36.99 %**	58.17 %
canonical, negative	**46.19 %**	49.41 %
reversed, affirmative	**54.30 %**	39.32 %
reversed, negative	**55.42 %**	38.55 %

First referent window: We found that participants looked at the critical image significantly more when listening to counterfactual negative sentences than counterfactual affirmative sentences in both linear and reversed order (main effect of polarity, $X^2(1) = 7.46$, $p = .01$). There was also an interaction between polarity and linear time term ($\beta = 1.15$ SE=0.16, $t = 6.94$, $p < .001$), suggesting that probability of fixations on the target image increased linearly in the negative sentences. Finally, the looking patterns in affirmative sentences showed greater curvature with an initial increase in the looks to the target image followed by a slight decrease (polarity*cubic term, $\beta = -0.40$, SE=0.13, $t = -2.91$, $p < .01$). Other effects and interactions were not significant.

Second referent window: We found that participants looked at the target image significantly more when the sentence was negative than when it was affirmative (main effect of polarity, $X^2(1) = 16.49$, $p = .001$). Regarding time-wise differences, there was an interaction between Clause Order and linear time term ($\beta = -0.41$, SE=0.2, $t = -2.01$, $p < .05$), indicating that the probability of fixations on the target image increased linearly in negative sentences in the canonical order whereas the increase was less steep in the reversed order.

General Discussion

Our goal was to investigate the incremental nature of counterfactual conditionals to contribute to understanding the complexity inherent in processing counterfactual language, focusing on the interplay of negation and clause order. We found a close alignment in processing between counterfactual negative and affirmative declarative sentences, and similarly between counterfactual affirmative and declarative negative sentences, albeit with a notable delay in the counterfactual affirmative scenario. In the reversed clause order, this pattern was repeated, with significantly more looks at the implied actual state (target) over the hypothetical one than in the canonical order.

We argue that part of the complexity in counterfactual statements might be attributed to how easily the QuD is accommodated: The clearer and more straightforward it is to represent the actual state of affairs in line with the QuD, the quicker and more precise the understanding of the counterfactual. In our experiments, we kept the QuD unchanged but made it easier for participants to attend on the implied actual state. We achieved this by reducing uncertainty in references using explicit negation and by changing the order of the clauses. These changes overall led to notably faster and more accurate comprehension.

The findings align with theories that closely integrate event representation with language comprehension. Unlike processing models that center around syntactic parsing of grammatical structures (e.g., Frazier & Clifton, 1996), various theories emphasize the extraction of pragmatically plausible event construals from the linguistic signals. These include approaches from "good-enough comprehension" (Ferreira et al., 2002; Ferreira & Patson, 2007), to dynamic pragmatic accounts of negation (Tian et al., 2010; Tian & Breheny, 2016), and proposals suggesting QuD effects on restricting hypothesis space (Skordos & Barner, 2019; Skordos & Papafragou, 2016).

These accounts are particularly relevant when considering negated counterfactuals. Two-step views propose that negation adds complexity to sentence processing. However, this might not hold in the context of counterfactuals. Here, negation may not add extra processing burdens but instead help restrict the referential set more quickly through negation, thus facilitating easier access to the underlying semantic content (also see Espino & Byrne, 2020; Orenes & Santamaría, 2014; Orenes et al., 2021).

One issue that stands is that similarly to Evcen and Wittenberg (2022), nearly half of our participants in the standard order and a third in the reversed order did not focus on the implied actual state at all. Despite extending prior studies (Orenes et al., 2019; Evcen & Wittenberg, 2022), and addressing factors like QuD and visual design, this issue persisted. One reason could be that the hypothetical content in affirmative counterfactuals might be represented due to the challenges in QuD accommodation. Another potential cause is the absence of a causal link between the antecedent and the consequent. While omitting causal links can reduce the influence of causal reasoning, which may otherwise overshadow the effect of grammatical cues, it also poses a problem. Causal reasoning is integral to counterfactual thinking, so without the causal link, listeners might be doing a heuristic match to what is mentioned in the conditional (Evans, 1996). This might contribute to the observed split between participants responding based on hypothetical or actual states, suggesting that some might adopt a consistent strategy influenced by the non-causal nature of the materials. These findings might also explain the 'systematic' hypothetical-actual state responder split: It might be the case that some participants adopt a working strategy based on the form of the material and apply the same strategy throughout the absence of causal content. In conclusion, the present data supports models of language processing that move beyond decoding syntactic structures and focus on constructing coherent event representations (see for a review, Ünal et al., 2019).

Acknowledgments

We thank Lea Zaric and Miguel Mejia for helping with stimuli creation, and anonymous reviewers for helpful discussion and advice.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439.

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

Byrne, R. M. (2024). How people think about the truth of hypothetical impossibilities. *Memory & Cognition*, *52*(1), 182-196.

Byrne, R. M., & Tasso, A. (1999). Deductive reasoning with factual, possible, and counterfactual conditionals. *Memory & Cognition*, *27*, 726-740.

Christiansen, M. H., & MacDonald, M. C. (2009). A usage-based approach to recursion in sentence processing. *Language Learning*, *59*, 126-161.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*(3), 472-517.

Clifton Jr, C., & Frazier, L. (2012). Discourse integration guided by the 'question under discussion'. *Cognitive Psychology*, *65*(2), 352-379.

Dale, R., & Duran, N. D. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, *35*(5), 983-996.

de Vega, M., & Urrutia, M. (2012). Discourse updating after reading a counterfactual event. *Psicologica: International Journal of Methodology and Experimental Psychology*, *33*(2), 157–173.

de Vega, M., Urrutia, M., & Riffo, B. (2007). Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition*, *35*(6), 1410–1421.

Degen, J. (2013). *Alternatives in pragmatic reasoning*. [Doctoral dissertation, University of Rochester].

Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive Science*, *39*(4), 667-710.

Degen, J., Kürşat, L., & Leigh, D. D. (2021). Seeing is believing: Testing an explicit linking assumption for visual world eye-tracking in psycholinguistics. In T. Fitch, C. Lamm, H. Leder, and K. Tessmar-Raible (Eds.), *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*, 1500-1506.

Diessel, H. (2001). The ordering distribution of main and adverbial clauses: A typological study. *Language*, *77*(2), 433-455.

Espino, O., & Byrne, R. M. (2020). The suppression of inferences from counterfactual conditionals. *Cognitive Science*, *44*(4), e12827.

Evans, J. S. B. (1996). On the mental representation of conditional sentences. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(4), 1086-1114.

Evans, J. S. B., Over, D. E., (2004). *If: Supposition, pragmatics, and dual processes*. Oxford University Press.

Evcen, E., & Wittenberg, E. (2022). Making the Question under Discussion explicit shifts counterfactual interpretation. In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 1855-1862.

Ferguson, H. J. (2012). Eye movements reveal rapid concurrent access to factual and counterfactual interpretations of the world. *Quarterly Journal of Experimental Psychology*, *65*(5), 939–961.

Ferguson, H. J., & Sanford, A. J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language*, *58*(3), 609–626.

Ferguson, H. J., Scheepers, C., & Sanford, A. J. (2009). Expectations in counterfactual and theory of mind reasoning. *Language and Cognitive Processes*, *25*(3), 297–346.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71-83.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15.

Fillenbaum, S. (1974). Information amplified: Memory for counterfactual conditionals. *Journal of Experimental Psychology*, *102*(1), 44-49.

Frazier, L., & Clifton, C. (1996). *Construal*. MIT Press.

Grice, H. P. (1975). Logic and Conversation. In P. Cole, & J. L. Morgan. (Eds.), *Syntax and Semantics* Speech Acts (pp. 41-58). New York: Academic Press.

Haigh, M., & Stewart, A. J. (2011). The influence of clause order, congruency, and probability on the processing of conditionals. *Thinking & Reasoning*, *17*(4), 402-423.

Ippolito, M. (2003). Presuppositions and implicatures in counterfactuals. *Natural Language Semantics*, *11*(2), 145-186.

Johnson-Laird, P. N., & Byrne, R. M. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*(4), 646.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology*, *60*(7), 976-990.

Kulakova, E., & Nieuwland, M. S. (2016). Understanding counterfactuality: A review of experimental evidence for the dual meaning of counterfactuals. *Language and Linguistics Compass*, *10*(2), 49–65.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime. Com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442.

Lüdtke, J., Friedrich, C. K., De Filippis, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence–picture verification paradigm. *Journal of Cognitive Neuroscience*, *20*(8), 1355-1370.

Mirman, D. (2017). *Growth curve analysis and visualization using R*. CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494.

Morgan, A. M., von der Malsburg, T., Ferreira, V. S., & Wittenberg, E. (2020). Shared syntax between comprehension and production: Multi-paradigm evidence that resumptive pronouns hinder comprehension. *Cognition*, *205*, 104417.

Nieuwland, M. S. (2013). "If a lion could speak...": Online sensitivity to propositional truth-value of unrealistic counterfactual sentences. *Journal of Memory and Language*, *68*(1), 54–67.

Nieuwland, M. S., & Martin, A. E. (2012). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition*, *122*(1), 102–109.

Orenes, I., & Santamaría, C. (2014). Visual content of words delays negation. *Acta Psychologica*, *153*, 107-112.

Orenes, I., Espino, O., & Byrne, R. M. (2022). Similarities and differences in understanding negative and affirmative counterfactuals and causal assertions: Evidence from eye-tracking. *Quarterly Journal of Experimental Psychology*, *75*(4), 633-651.

Orenes, I., Garcia-Madruga, J. A., Gomez-Veiga, I., Espino, O., & Byrne, R. M. (2019). The comprehension of counterfactual conditionals: Evidence from eye-tracking in the visual world paradigm. *Frontiers in Psychology*, *10*, 1172.

Papoutsaki, A., Laskey, J., & Huang, J. (2017). Searchgazer: Webcam eye tracking for remote studies of web search. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, 17-26.

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2018). The relation between factual and counterfactual conditionals. *Cognitive Science*, *42*(7), 2205–2228.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*(1), 1-23.

Roberts, C. (1996). Information structure in discourse: Toward a unified theory of formal pragmatics. *Ohio State University Working Papers in Linguistics*, *49*, 91–136.

Roberts, C. (2004). Context in dynamic interpretation. *The Handbook of Pragmatics*, *197*, 220.

Ronai, E., & Xiang, M. (2021). Exploring the connection between Question Under Discussion and scalar diversity. *Proceedings of the Linguistic Society of America*, *6*(1), 649-662.

Santamaría, C., Espino, O., & Byrne, R. M. (2005). Counterfactual and semifactual conditionals prime alternative possibilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 1149-1154.

Skordos, D., & Barner, D. (2019). Language comprehension, inference, and alternatives. In Cummins, C., (Ed). *The Oxford Handbook of Experimental Semantics and Pragmatics*. Oxford University Press

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, *153*, 6–18.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language*, *17*(1–2), 3–23.

Thompson, V. A., & Byrne, R. M. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1154-1170.

Tian, Y., Breheny, R. (2016). Dynamic pragmatic view of negation processing. In Larrivée, P., Lee, C. (Eds.), Negation and polarity: Experimental perspectives (pp. 21-43). Cham, Switzerland: Springer International.

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *The Quarterly Journal of Experimental Psychology, 63*(12), 2305–2312.

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context–why and when we represent the positive argument. *Language, Cognition and Neuroscience*, *31*(5), 683-698.

Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic meaning. *Topics in Cognitive Science*, *13*(1), 224-242.

Vos M.T., Minor S., & Ramchand G.C. (2022). Comparing infrared and webcam eye tracking in the Visual World Paradigm. *Glossa Psycholinguistics, 1*(1), 1-37.

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX).