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

Title
Abstract Sentences elicit more Uncertainty and Curiosity than Concrete Sentences

Permalink
https://escholarship.org/uc/item/7cj2g289

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

Authors
Mazzuca, Claudia
Villani, Caterina
Lamarra, Tommaso
et al.

Publication Date
2024

Peer reviewed

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

Abstract Sentences elicit more Uncertainty and Curiosity than Concrete Sentences

Claudia Mazzuca (claudia.mazzuca@uniroma1.it)
Sapienza University of Rome, Department of Dynamic and Clinical Psychology, and Health Studies,
via degli Apuli 1, Rome, 00185 Italy

Caterina Villani (caterina.villani6@unibo.it)
University of Bologna, Department of Modern Languages, Literatures, and Cultures,
via Cartoleria 5, Bologna, 40124 Italy

Tommaso Lamarra (tommaso.lamarra2@unibo.it)
University of Bologna, Department of Modern Languages, Literatures, and Cultures, via Cartoleria 5
Bologna, 40124 Italy

Marianna Marcella Bolognesi (m.bolognesi@unibo.it)
University of Bologna, Department of Modern Languages, Literatures, and Cultures, via Cartoleria 5
Bologna, 40124 Italy

Anna M. Borghi (anna.borghi@uniroma1.it)
Sapienza University, Department of Dynamic and Clinical Psychology, and Health Studies, via degli Apuli 1, Rome, 00185;
Institute of Cognitive Sciences and Technologies, National Research Council, Via Romagnosi 18A, Rome, 00196, Italy

Abstract

Are abstract sentences associated with specific constructs in dialogue, i.e., higher uncertainty, more curiosity and willingness to continue a conversation, and more causal questions? In three preregistered experiments we address these questions asking participants to evaluate the plausibility of linguistic exchanges referred to concrete and abstract concepts. Results support theories proposing that abstract concepts involve more inner monitoring and social dynamics compared to concrete concepts, and suggest that reaching alignment in dialogue is more effortful with abstract than with concrete concepts.

Keywords: Abstract concepts; concrete concepts; dialogue; alignment; social interaction; uncertainty; curiosity; metacognition

Introduction

Humans are remarkably good at talking to each other. Indeed, it has been proposed that we are hard-wired for dialogue (Pickering & Garrod, 2021). To be successful, a conversation rests upon the shared understanding of the interlocutors, that reach a common ground through what has been called *interactive alignment* (Pickering & Garrod, 2021). This mechanism—which is intrinsically a form of joint action—broadly consists of an unconscious process by which interlocutors linguistically align their representations at the same time at multiple levels.

However, conversational topics differ and some of them might require more effort. For instance—at least in English—abstract language is ubiquitous (Lupyan & Winter, 2018), and a consistent body of research shows that abstract concepts (e.g., "democracy") are processed and represented differently from concrete concepts (e.g., "hammer"; Banks et al., 2023; Binder et al., 2005; Bolognesi & Steen, 2018; Borghi, 2023; Conca et al., 2021; Dove, 2022; Henningsen-Schomers & Pulvermüller, 2022; Mazzuca et al., 2021). It has been suggested that whereas in the case of concrete concepts speakers can reach an alignment more easily, and simply by attending to a perceptual entity or by recalling it, with abstract concepts they need more "mutual monitoring" (Gandolfi et al., 2023). In the same vein, theories of conceptual representation have suggested that for abstract concepts social interaction is crucial. Because the meaning of abstract concepts is more open compared to concrete concepts, it can be co-built and negotiated with others (Borghi, 2022; Mazzuca & Santarelli, 2022). Consistent with this, in a study where participants were asked to imagine being involved in a real conversation and to respond to target sentences composed of different kinds of abstract and concrete concepts, responses to abstract and concrete sentences qualitatively differed (Villani et al., 2022). Specifically, abstract sentences elicited more expressions of uncertainty (e.g., "mmmh", "how is that?"), questions aimed at knowing more about the topic (e.g., "tell me more"), and more *why* and *who* questions compared to concrete sentences. Concrete sentences, instead, elicited less uncertain expressions, less questions aimed at knowing more, and more *where* and *when* questions. This pattern of results has been explained suggesting that abstract concepts, in general, evoke more social validation than concrete concepts. For instance, we might feel the need for other people's aid to understand—or simply to frame—better an abstract sentence compared to a concrete sentence (e.g., Social Metacognition, Borghi et al., 2018). Rating studies support this intuition, showing that on average more abstract words score higher in Social

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).

Metacognition, i.e., the need to rely on others to understand word meaning, than more concrete words (Villani et al., 2019). Along the same lines, people typically feel less confident about the meaning of abstract than concrete words (Fini et al., 2023; Mazzuca et al., 2022). In addition, a recent norming study broadening the notion of *socialness* (Diveica et al., 2023) and collecting ratings for more than 8,000 English words found that socialness is negatively correlated with concreteness, Body-Object-Interaction, and Imageability—hence suggesting that more concrete words are considered as less relevant on the social dimension. Finally, abstract concepts are consistently associated with internal, mental, and interoceptive states (Barsalou & Weimer-Hastings, 2005; Connell et al., 2018; Villani et al., 2019), whereas concrete concepts are typically associated with perceptual and sensorimotor states (Lynott et al., 2020).

However, the extent to which this pattern of results can be generalized remains an open question. Indeed, norming and linguistic production studies might only reflect participants' metacognitive awareness of the appropriateness of linguistic exchanges (linguistic production studies), or might conceal important conceptual information that presenting a word in isolation fails to convey (norming studies). If abstract sentences evoke more uncertainty and clarification expressions, and to some extent more social interaction than concrete sentences, then this should also be observed in behavioral-related measurements.

This Study

This study tackles these questions with three different experiments, preregistered at https://osf.io/erxd6. Each experiment targets a specific aspect of the relation between abstractness and social interactions leveraging on excerpts of simulated conversations. Participants are presented with sentences varying in abstractness, followed by three different types of possible follow-up expressions or questions. Sentences are matched for their morphological complexity and include only subject, verb, and noun (e.g., *"I made a cake", "I made a judgment"*). Follow-up items are selected from a previous Italian production study (Villani et al., 2022) and validated with two pilot studies. Follow-ups represent expressions of uncertainty vs. expression of certainty (Experiment 1); expressions of curiosity/signaling the willingness to know more vs. the willingness to end the conversation (Experiment 2); *why* and *who/for whom* questions vs. *where* and *when* questions (Experiment 3). Participants are asked to judge whether the sentence–follow-up combinations are plausible linguistic exchanges, as if they were embedded in a real conversation. We collected response times and frequencies. While this study has its own relevance for the literature on the link between abstractness, conversations, and social interaction, it also constitutes a conceptual replication of a previous production study (Villani et al., 2022).

Experiment 1: Abstractness and Uncertainty

Experiment 1 tests whether people feel more uncertain with abstract compared to concrete concepts. We hypothesized that abstract sentences elicit more uncertainty about their possible meaning and evoke a longer monitoring process compared to concrete sentences. So, we expected participants to judge more plausible expressions of uncertainty as follow-ups for abstract sentences compared to expressions signaling that the sentence has been understood, and the opposite for concrete sentences. We also expected a difference in RTs to abstract and concrete sentences as a function of the type of follow-up.

Method

Twenty-eight Italian speakers were recruited at the University of Bologna (M_{age} = 19.32; *SD* = 1.22, age range = 18 – 24). To generate the stimuli, we first selected 60 sentences from Villani et al. (2022), and asked a separate sample of 38 participants to rate them on 7-point Likert scales in terms of concreteness~abstractness, how much they thought sentences needed a context to be understood, naturalness, and familiarity. Participants' ratings served as basis for the selection of 42 sentences (*n* = 21 abstract; *n* = 21 concrete) with naturalness ratings > 4, that did not significantly differ in terms of naturalness and familiarity, but that differed in terms of abstractness. To generate the follow-ups, we first selected 24 among the most frequently produced follow-ups to abstract and concrete sentences in Villani et al. (2022). Then we asked a further sample of 22 participants to rate on 7-point Likert scales linguistic exchanges composed of abstract and concrete sentences randomly paired with the follow-ups in terms of (a) whether the answer signals uncertainty vs. certainty; (b) whether the answer signals curiosity vs. willingness of ending the conversation. For the present experiment, we selected follow-ups corresponding to the highest two and lowest two values on (a), i.e., "I did not understand; "What do you mean" vs. "Good job"; "Well done". Sentences–follow-up pairings were presented on a computer screen in two blocks separated by a short break. Participants were asked to decide whether the combinations could plausibly occur in a real conversation (yes vs. no). The order of sentences was randomized across participants, ensuring that each sentence was repeated only twice within a single block. Each trial started with a central black fixation cross of 500 *ms*, followed by a sentence lasting for 1500 *ms*, then the follow-up appeared, remaining on the screen until the response was given by participants up to a maximum of 3000 *ms*. RTs were recorded from follow-up onset. Participants provided their response by pressing two keys on the keyboard (i.e., "x" and "m"), the mapping of which was counterbalanced between participants. Data points exceeding +/- 2*SD* from the average RTs of plausible and not plausible answers were excluded from the analyses, i.e., 2.8% of datapoints (*n* = 135).

Data Analysis Data analysis and visualization for all the experiments were carried out with R (RCore Team, 2019) and

RStudio (v4.0.3). Plausibility judgments for each experiment (1 = yes; 0 = no) were modeled with a mixed effects binomial logistic regression using glmer() function from "lme4" R's package (Bates et al., 2015). The model featured Category (abstract vs. concrete), Type of Follow-up (Experiment 1: uncertainty vs. certainty; Experiment 2: curiosity vs. end; Experiment 3: causal vs. spatio-temporal), and their interaction as fixed effects, with participants and sentences as random intercepts. Significance of the main effects and interactions was assessed with Wald Chi-Squared tests using the Anova() function from "car" R's package (Fox & Weisberg, 2019). RTs for each experiment were first log-transformed and then analyzed separately for "plausible" and "not plausible" responses with linear mixed models performed with the lmer() function from "lme4" R's package (Bates et al., 2015). The structure of the models was identical to the one used for plausibility judgments. Post-hoc contrasts were carried out with the "emmeans" R's package (Lenth, 2021) using Tukey's adjustment for multiple comparisons. Sample size for each experiment was computed using the wp.logistic() function from the WebPower R's package (Zhang et al., 2023). To achieve an 80% power with an alpha of .05, and an expected probability of .8 for the outcome 1 (plausible) when the level of the predictor is 0 (e.g., abstract–uncertain) and .2 for the outcome 0 (not plausible) when the level of the predictor is 1 (e.g., abstract–certain) the power analysis suggested a sample size of N = 26 participants for each experiment.

Results and Discussion

Plausibility Judgments We found a significant main effect of Category, χ^2 (1) = 4.26, p =.038, Type of Follow-up, χ^2 (1) = 972.91, p <.001, and a significant interaction between Category and Type of Follow-up, χ^2 (1) = 102.07, p <.001. Post-hoc contrasts showed that for both abstract, z =19.687, p <.0001, and concrete sentences, z =26.763, p <.0001, participants judged more plausible certainty follow-ups compared to uncertainty follow-ups. However, we also found that uncertainty follow-ups were considered more plausible for abstract than for concrete sentences, z = 6.066, p <.0001, whereas there was no difference within certainty follow-ups, p =.66 (see Figure 1).

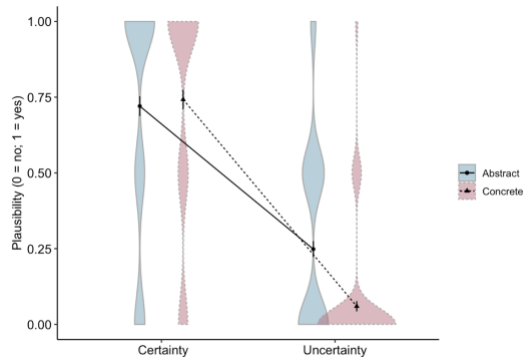

Figure 1: Plausibility judgements for abstract and concrete sentences as a function of follow-ups.

Response Times: Plausible We found a significant main effect of Type of Follow-up, F(1, 1059.93) = 155.04, p < .001, showing that participants were overall slower with follow-ups signaling uncertainty (M uncertainty = 1264.09 *ms*; SD = 424.38; M certainty = 958.91 *ms*; SD = 388.75). No other main effect or interaction reached significance, all p_s > .062.

Response Times: Implausible We found a significant main effect of Type of Follow-up, F(1, 1061.05) = 27.40, p < .0001, and a significant two-way interaction between Category and Type of Follow-up, F(1, 1105.80) = 5.77, p = .016. No other main effect reached significance (Category p = .11). Post-hoc contrasts showed that participants are faster to judge as not plausible abstract sentences when they are paired with uncertainty follow-ups compared to when they are paired with certainty follow-ups, t(984) = -2.020, p = .043. Likewise, they are faster to judge as not plausible concrete sentences when they are paired with uncertainty follow-ups compared to when they are paired with certainty follow-ups, t(1508) = -5.390, p < .001. We also found that participants responded slower to abstract–uncertainty pairings than concrete–uncertainty pairings, t(67.1) = 3.748, p < .001. There was instead no difference within certainty follow-ups, t(187.5) = -0.5364, p = .716. Overall then, follow-ups signaling certainty were judged as more plausible than those indicating uncertainty, regardless of the concreteness ~ abstractness of sentences. This was also reflected in longer RTs for plausible answers featuring uncertainty follow-ups, compared to those featuring certainty follow-ups. However, pairings with uncertainty follow-ups preceded by abstract sentences were judged as more plausible than those preceded by concrete sentences. Indeed, participants were also slower in judging abstract–uncertainty pairings as not plausible compared to concrete–uncertainty pairings (see Figure 2).

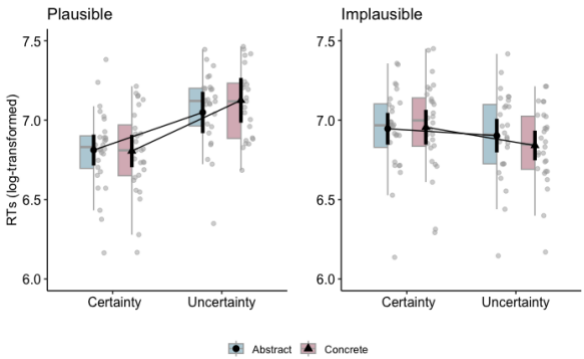

Figure 2: RTs for abstract and concrete sentences as a function of follow-ups. Raw data in the background are aggregated over participants.

Experiment 2: Abstractness and Curiosity

Experiment 2 asks whether abstractness elicits more curiosity than concreteness. We hypothesized that with abstract sentences participants rely more on others to understand better their meaning or tend to ask for more information (i.e., they are more curious) compared to concrete sentences. So, we expected participants to judge more plausible for abstract sentences expressions signaling curiosity, or the need for more specifications compared to expressions signaling the willingness to end the conversation, and the opposite for concrete sentences. We also expected a difference in RTs to abstract and concrete sentences as a function of the type of follow-up.

Method

Twenty-eight Italian speakers were recruited at the University of Bologna (M_{age} = 20.96; *SD* = 2.52, age range = 18 – 29). The procedure and data analysis were identical to those of Experiment 1, as well as linguistic stimuli except for the follow-up expressions that were selected from the pilot study (b) based on their scores on ratings of curiosity vs. willingness to end the conversation, i.e., "Describe; "Tell me more" vs. "Ok"; "Thanks". We removed 2.84% of datapoints (*n* = 134) for further analyses.

Results and Discussion

Plausibility Judgments We found a significant main effect of Category $\chi^2(1)$ = 19.20, *p* <.0001, and a significant interaction between Category and Type of Follow-up, χ^2 (1) = 718.88, *p* <.0001. No other main effect reached significance, Category *p* = .15. In keeping with our predictions, post-hoc contrasts showed that participants judged as more plausible abstract sentences when they were paired with curiosity follow-ups compared to when they were paired with end follow-ups, *z* = 19.949, *p* <.0001. Conversely, concrete sentences were judged as more plausible when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, *z* = 18.268, *p* <.0001. We also found that abstract–curiosity pairings were judged as more plausible than concrete–curiosity pairings, *z* = 17.459, *p* <.0001, and abstract–end pairings were judged as less plausible than concrete–end pairings, *z* = -5.569, *p* <.0001 (see Figure 3).

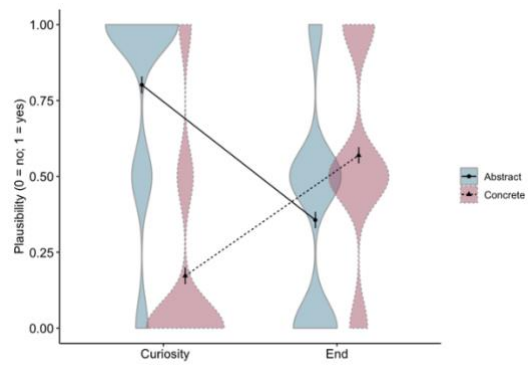

Figure 3: Plausibility judgements for abstract and concrete sentences as a function of follow-ups

Response Times: Plausible We found a significant main effect of Type of Follow-up, *F*(1, 1472.40) = 5.09, *p* = .024, and a significant interaction between Category and Type of Follow-up, *F*(1, 1446.66) = 20.50, *p* <. 0001. No other main effect reached significance, Category *p* = .11. Post-hoc contrasts showed that participants were faster to judge plausible concrete sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, *t*(1122) = -4.153, *p* <.0001, whereas there was no difference within abstract sentences across types of follow-ups, *t*(2036) = 1.858, *p* =. 063. We also found that participants responded faster to abstract–curiosity pairings than concrete–curiosity pairings, *t*(161) = -3.791, *p* =.0002, whereas there was no difference between abstract–end pairings and concrete–end pairings, *t*(113) = 1.922, *p* =.057.

Response Times: Implausible We found a significant main effect of Category, *F*(1, 45.16) = 17.84, *p* <.0001, Type of Follow-up, *F*(1, 2161.16) = 41.40, *p* <.0001, and a significant interaction between Category and Type of Follow-up, *F*(1, 2180.16) = 4.54, *p* = .033. Post-hoc contrasts showed that participants were faster to judge as not plausible abstract sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, *t*(2070) = 5.279, *p* <.0001. Likewise, they were faster to judge as not plausible concrete sentences when they were paired with end follow-ups compared to when they were paired with curiosity follow-ups, *t*(2332) = 3.743, *p* =.0002. We also found that abstract–curiosity pairings were responded to slower than concrete–curiosity pairings, *t*(126.1) = 4.344, *p* <.0001. There was also a difference within end follow-ups, where abstract–end pairings were also responded to slower than concrete–end pairings, *t*(74.3) = 2.504, *p* = .014. So, in line with our predictions, participants judged abstract–curiosity pairings more plausible than both concrete–curiosity pairings and abstract–end pairings. This was partly confirmed by RTs, where abstract–curiosity pairings were judged as plausible faster than concrete–

curiosity pairings, and as not plausible slower than concrete–curiosity pairings (see Figure 4).

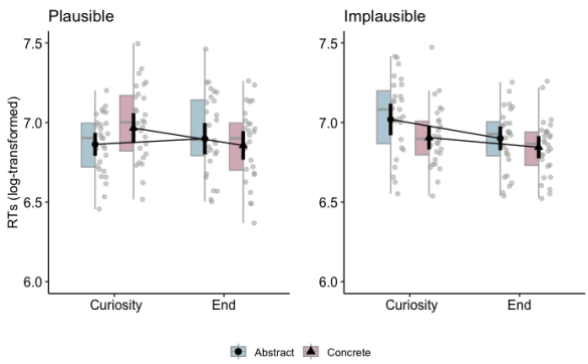

Figure 4: RTs for abstract and concrete sentences as a function of follow-ups.

Experiment 3: Abstractness and Types of Questions

Experiment 3 investigates whether abstract sentences evoke more interest in internal processes and causal mechanisms rather than spatial and temporal specifications compared to concrete sentences. We hypothesized that participants judge more plausible follow-ups for abstract sentences "why" and "who" questions (i.e., causal questions) compared to "where" and "when" questions (i.e., spatio-temporal questions), and the opposite for concrete sentences. We also expected a difference in RTs to abstract and concrete sentences as a function of the type of follow-up.

Method

Twenty-seven Italian speakers were recruited at the University of Bologna (M_{age} = 21.81; *SD* = 2.77, age range = 18 – 28). The procedure and stimuli were identical to those of Experiment 1 and 2, except for the follow-up expressions for which we selected the most frequently produced questions related to spatio-temporal and causal–agent issues from Villani et al. (2022), i.e., "where?/when?" and "why/for whom?", respectively. We removed 2.75% of datapoints (*n* = 125) for further analyses.

Results and Discussion

Plausibility Judgments We found a significant main effect of Category, $\chi^2(1) = 14.79$, $p < .0001$, Type of Follow-up, $\chi^2(1) = 123.05$, $p < .0001$, and a significant interaction between Category and Type of Follow-up, $\chi^2(1) = 60.78$, $p < .0001$. Contrary to our predictions, post-hoc contrasts showed that participants judged as more plausible abstract sentences when they were paired with spatio-temporal follow-ups compared to when they were paired with causal follow-ups, $z = 3.180$, $p = .001$. This, however, was true also for concrete sentences, $z = 13.189$, $p < .0001$, in line with our expectations. We also found that abstract–spatiotemporal pairings were judged as less plausible than concrete–spatiotemporal pairings, $z = -6.686$, $p < .0001$, in keeping with what we expected. There was instead no difference within causal follow-ups across types of sentences, $z = -1.088$, $p = .27$ (see Figure 5).

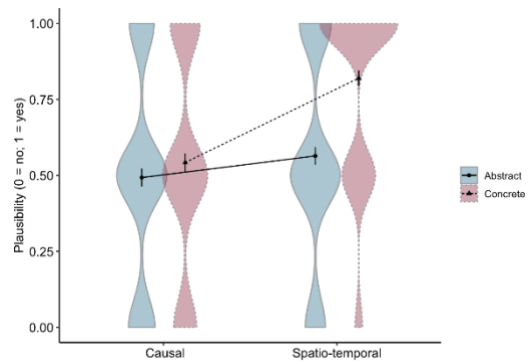

Figure 5: Plausibility judgements for abstract and concrete sentences as a function of follow-ups.

Response Times: Plausible We found a significant main effect of Category, $F(1, 39.28) = 13.72$, $p = .006$, showing that concrete sentences were responded to faster than abstract sentences (*M* concrete = 956.25 *ms*; *SD* = 340.13; *M* abstract = 994.64 *ms*; *SD* = 376.94), and a significant main effect of Type of Follow-up, $F(1, 2579.06) = 32.55$, $p < .0001$, showing that overall participants were faster with spatio-temporal follow-ups (*M* = 949.24 *ms*; *SD* = 341.86) compared to causal follow-ups (*M* = 1004.37 *ms*; *SD* = 374.23). The two-way interaction did not reach significance, $p = .095$.

Response Times: Implausible We found a significant main effect of Type of Follow-up, $F(1, 1675.18) = 4.27$, $p = .038$, and a significant interaction between Category and Type of Follow-up, $F(1, 1669.74) = 9.01$, $p = .002$. No other main effect reached significance, Category $p = .173$. Post-hoc contrasts showed that participants were faster to judge as not plausible concrete sentences when they were paired with causal follow-ups compared to when they were paired with spatio-temporal follow-ups, $t(1591) = 3.176$, $p = .001$. There was instead no difference within abstract sentences, $t(1730) = 0.738$, $p = .460$. We also found that spatio-temporal follow-ups were responded to faster when they were paired with abstract sentences compared to when they were paired with concrete sentences, $t(148.6) = -2.660$, $p = .008$. There was instead no difference within causal follow-ups across types of sentences, $t(69.5) = 0.825$, $p = .412$. So, overall causal follow-ups were always judged as being less plausible, regardless of the abstractness of sentences. This was also reflected by slower RTs for plausible answers to causal follow-ups. Spatio-temporal follow-ups instead were judged as more plausible for concrete sentences than for abstract sentences. Consistently, participants were slower to judge concrete sentences as being not plausible when they were followed by spatio-temporal follow-ups compared to abstract sentences (see Figure 6).

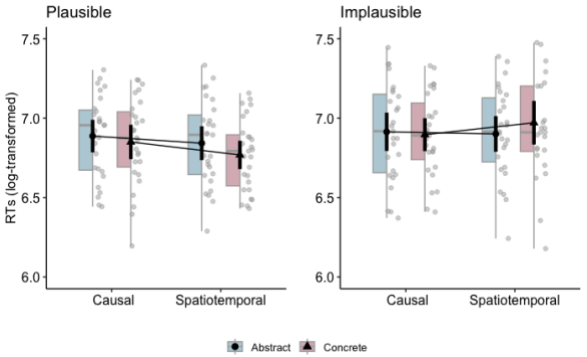

Figure 6: RTs for abstract and concrete sentences as a function of follow-ups.

General Discussion

In three preregistered experiments we asked participants to evaluate the plausibility of linguistic exchanges referring to concrete and abstract concepts. Specifically, we predicted that abstract sentences would be evaluated as more plausible with follow-up expressions indicating uncertainty, curiosity, and why and who questions, while concrete sentences with follow-ups indicating certainty, end of the conversation, and spatiotemporal questions.

First, we found that abstract concepts are more associated with uncertainty than concrete ones. Participants considered uncertainty expressions more plausible when matched with abstract than with concrete sentences, and data on RTs consistently showed that participants were slower to judge as implausible abstract–uncertainty pairings than concrete–uncertainty pairings. This is in line with single words ratings, showing that people's confidence in knowing word meanings is lower with abstract than concrete words (Mazzuca et al., 2022). Crucially, these findings conceptually replicate and extend with a novel task the results of a production task (Villani et al., 2022). Theoretically, our findings are in line with proposals linking abstract concepts processing with inner monitoring mechanisms (Binder et al., 2005; Fernyhough & Borghi, 2023; Shea, 2018).

Second, our results show that, with pairings involving abstract sentences, people seem to think of an interaction that continues. Indeed, with abstract sentences, pairings with curiosity follow-ups are evaluated as more plausible and processed faster than with concrete sentences, while the opposite is true for pairings with end follow-ups. The results thus indicate that, in conversations involving abstract concepts, people are willing to engage in longer and more in-depth interactions. Many could be the causes of this willingness, among which the higher variability across people and the higher uncertainty of the meaning of abstract sentences, which makes it necessary to understand better and interact longer to reach a common ground. Theoretically, this result is compatible with views suggesting that abstract concepts might lead to promoting social interaction. Evidence with single words has shown that abstract concepts elicit more social situations (Wiemer-Hastings & Barsalou, 2005) and are linked with socialness (Diveica et al., 2023; Pexman et al., 2023). In addition, recent research has highlighted the importance of social interaction for concepts and abstractness (Andrade-Lotero et al., 2023; Olsen & Tylén, 2023; Rączaszek-Leonardi & Zubek, 2023). More specifically, our results support recent views proposing that, due to their open character, abstract concepts might be more prone to be socially negotiated than concrete ones, which point to a specific single referent (Borghi, 2022).

Finally, Experiment 3 showed that spatio-temporal follow-ups are considered more plausible with concrete than abstract sentences and require more time to be judged as implausible when paired with concrete sentences. This indicates that people tend to locate concrete concepts spatiotemporally, consistent with the fact that their referents are spatially bounded objects or entities. Contrary to our expectations, we did not find that abstract sentences were judged more plausible with causal follow-ups, suggesting that causality is relevant for both kinds of sentences. The discrepancy with the results of Villani et al. (2022) might be due to the difference between a plausibility evaluation and a production task, the latter likely capturing more immediate associations.

Methodologically, our study has some particularities. Instead of using concrete and abstract words embedded in a single sentence, we used a simulated verbal exchange. In addition, sentences were derived from a previous production study, hence investigating concepts embedded in expressions taken from natural—although simulated—linguistic exchanges. Finally, this study represents both a conceptual replication and a consolidation with a more structured method of previous results. Clearly, this work is not exempt from limitations. For example, here we focus on differences between concrete and abstract concepts. However, many authors agree on the importance of focusing on different varieties of concrete and abstract concepts (Conca et al., 2021; Desai et al., 2018; Persichetti et al., 2023). Hence, in the future, we intend to address fine-grained differences in plausibility judgments employing different types of concrete and abstract concepts. Another potential limitation is that we did not consider concepts' hierarchical level (i.e., their specificity, Bolognesi & Caselli, 2023; Villani et al., 2024), although an ongoing study is devoted to this aspect.

To conclude, our findings have several implications for the literature on concepts in general—and abstractness in particular—but they are also relevant for studies on dialogue (Pickering & Garrod, 2021). In fact, overall, we found that when conversational topics are more abstract people judge them as requiring more effort to align with the interlocutor compared with concrete conversational topics. Operationally, this effort in reaching alignment is revealed in the higher uncertainty of follow-ups associated with abstract sentences and increased willingness to continue the conversation (curiosity) with pairings including abstract sentences. Hence, our results add layers to studies on dialogue showing that the concreteness ~ abstractness of the topic might modulate the conversational dynamics (Gandolfi et al., 2023).

Acknowledgments

The authors would like to thank all the Body, Action, and Language Lab (BalLab, https://sites.google.com/view/annaborghilab) for insightful remarks on early versions of this study. In particular, we thank Chiara Fini, Luca Tummolini, Angelo Mattia Gervasi, Chiara De Livio, and Valentina Rossi for their stimulating thoughts. The authors would also like to thank all members of ABSTRACTION research group (GRANT AGREEMENT: ERC-2021-STG-101039777) for comments and discussions. They are, in alphabetical order, Adele Loia, Andrea Amelio Ravelli, Claudia Collaccini, and Giulia Rambelli. Data presented in the present contribution have been collected at the Experimental Lab of the Department of Modern Languages, Literatures and Cultures of the University of Bologna, Italy (https://site.unibo.it/laboratorio-sperimentale/en). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the Experimental Lab.

Funding: Claudia Mazzuca was funded by Research Projects of National Relevance PRIN 2022 – PROJECT ASSO (Abstract conceptS and Social InteractiOn) and PRIN 2022 PNRR – PROJECT Democratizing Concepts (DeCo). Caterina Villani, Tommaso Lamarra, and Marianna Bolognesi were funded by ABSTRACTION European Union (GRANT AGREEMENT: ERC-2021-STG-101039777). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Anna Borghi was funded by PNRR funds, PE8 - AGE-IT - Spoke 4: Healthy aging. Project line led by AMB: "Older adults' relationship with the natural environment to enhance their cognitive performances and promote mental health".

References

Andrade-Lotero, E. J., Ortiz-Duque, J. M., Velasco-García, J. A., & Goldstone, R. L. (2023). The division of linguistic labour for offloading conceptual understanding. *Philosophical Transactions of the Royal Society B, 378*(1870), 20210360.

Banks, B., Borghi, A. M., Fargier, R., Fini, C., Jonauskaite, D., Mazzuca, C., ... & Woodin, G. (2023). Consensus paper: Current perspectives on abstract concepts and future research directions. *Journal of Cognition, 6*(1).

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of cognitive neuroscience, 17*(6), 905-917.

Bolognesi, M., & Steen, G. (2018). Editors' introduction: abstract concepts: structure, processing, and modeling. *Topics in cognitive science, 10*(3), 490-500.

Bolognesi, M. M., & Caselli, T. (2023). Specificity ratings for Italian data. *Behavior Research Methods, 55*(7), 3531-3548.

Borghi, A. M. (2023). *The Freedom of Words: Abstractness and the Power of Language*. Cambridge University Press.

Borghi, A. M. (2022). Concepts for which we need others more: The case of abstract concepts. *Current directions in psychological science, 31*(3), 238-246.

Borghi, A. M., Barca, L., Binkofski, F., & Tummolini, L. (2018). Abstract concepts, language and sociality: from acquisition to inner speech. *Philosophical Transactions of the Royal Society B, 373*(1752), 20170134.

Conca, F., Borsa, V. M., Cappa, S. F., & Catricalà, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neuroscience & Biobehavioral Reviews, 127*, 474-491.

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1752), 20170143.

Desai, R. H., Reilly, M., & van Dam, W. (2018). The multifaceted abstract brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1752), 20170122.

Diveica, V., Pexman, P. M., & Binney, R. J. (2023). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods, 55*(2), 461-473.

Dove, G. (2022). *Abstract concepts and the embodied mind: Rethinking grounded cognition*. Oxford University Press.

Fernyhough, C., & Borghi, A. M. (2023). Inner speech as language process and cognitive tool. *Trends in cognitive sciences*, 27(12):1180-1193.

Fini, C., Falcinelli, I., Cuomo, G., Era, V., Candidi, M., Tummolini, L., ... & Borghi, A. M. (2023). Breaking the ice in a conversation: abstract words prompt dialogs more easily than concrete ones. *Language and Cognition*, 1-22.

Fox, J. & Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks, CA.

Gandolfi, G., Pickering, M. J., & Garrod, S. (2023). Mechanisms of alignment: shared control, social cognition and metacognition. *Philosophical Transactions of the Royal Society B, 378*(1870), 20210362.

Henningsen-Schomers, M. R., & Pulvermüller, F. (2022). Modelling concrete and abstract concepts using brain-constrained deep neural networks. *Psychological research, 86*(8), 2533-2559.

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.1-1. https://CRAN.R-project.org/package=emmeans

Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more

iconic?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170137.

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*, 1271-1291.

Mazzuca, C., Fini, C., Michalland, A. H., Falcinelli, I., Da Rold, F., Tummolini, L., & Borghi, A. M. (2021). From affordances to abstract words: The flexibility of sensorimotor grounding. *Brain Sciences*, *11*(10), 1304.

Mazzuca, C., & Santarelli, M. (2022). Making it abstract, making it contestable: politicization at the intersection of political and cognitive science. *Review of Philosophy and Psychology*, 1-22.

Mazzuca, C., Falcinelli, I., Michalland, A. H., Tummolini, L., & Borghi, A. M. (2022). Bodily, emotional, and public sphere at the time of COVID-19. An investigation on concrete and abstract concepts. *Psychological research*, *86*(7), 2266-2277.

Olsen, K., & Tylén, K. (2023). On the social nature of abstraction: cognitive implications of interaction and diversity. *Philosophical Transactions of the Royal Society B*, *378*(1870), 20210361.

Persichetti, A.S., Shao, J., Denning, J.M., Gotts, S.J., & Martin, A. (2024) Taxonomic structure in a set of abstract concepts. *Frontiers in Psychology*, *14*, 1278744. doi: 10.3389/fpsyg.2023.1278744

Pexman, P. M., Diveica, V., & Binney, R. J. (2023). Social semantics: the organization and grounding of abstract concepts. *Philosophical Transactions of the Royal Society B*, *378*(1870), 20210363.

Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press.

Rączaszek-Leonardi, J., & Zubek, J. (2023). Is love an abstract concept? A view of concepts from an interaction-based perspective. *Philosophical Transactions of the Royal Society B*, *378*(1870), 20210356.

R Core Team (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Shea, N. (2018). Metacognition and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170133.

Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, *11*(3), 403-430.

Villani, C., Orsoni, M., Lugli, L., Benassi, M., & Borghi, A. M. (2022). Abstract and concrete concepts in conversation. *Scientific Reports*, *12*(1), 17572.

Villani, C., Loia, A. & Bolognesi, M.M. (2024). The semantic content of concrete, abstract, specific, and generic concepts. *Language and Cognition*. Published online 2024:1-28. doi:10.1017/langcog.2023.64.

Zhang, Z., Mai, Y., Yang, M., & Zhang, M. Z. (2018). Package 'WebPower'. Basic and Advanced Statistical Power Analysis Version, 72.