

De re interpretation in belief reports--An experimental investigation

Introduction The subject of the sentence *In 1978, the president built a huge skyscraper in NYC* can, under a *de re* reading refer to the current president, or under a *de dicto* reading can refer to the holder of the presidential role within the temporal operator (*in 1978*). Despite the variety of theoretical approaches to formalize this distinction in natural language semantics (Cresswell & von Stechow, 1982; Keshet, 2008; Pearson, Forthcoming; Percus, 2000; von Fintel & Heim, 2011; see a summary in Keshet & Schwarz, 2019), it is generally accepted that both interpretations should be generated by the grammar. Yet in some contexts one reading may be argued to be strongly inaccessible, as, e.g. argued by Nelson (2019) for the lack of *de re* meaning in (1).

(1) Sally *believes* that *her brother* is happy. (Nelson, 2019: supplement (3))

[supposed *de re* context: Sally observes someone laughing, not knowing he is her brother.]

As far as we know, there has yet to be any quantitative investigation of both *de re* and *de dicto* readings under intensional operators and the factors affecting their accessibility. Following the literature, we expect the grammar to generate both readings with additional heuristics resolving this ambiguity in conversation. The heuristics in our case refers to contexts that can be devised in which one reading is, in fact, *unavailable* or *inaccessible*. Here, we test the availability and accessibility of *de re* readings under two different intensional operators (both attitude embedding verbs: *believe*, *say*) and two different QUDs (*real world facts*, *subject's mental state*) to lay the experimental groundwork for understanding factors that go into the ambiguity, and occasional lack thereof.

Experiment Participants (128 native speakers of English) were recruited through Amazon Mechanical Turk to participate in an agreement rating task with three mistaken identity stories (Table 1). Regarding the belief reporter, the belief holder, and a target entity that is the subject of the belief, our contexts vary in **whether the belief holder can identify the entity in the real world** under a particular description (e.g. "the poem that Nina wrote" in Table 1): if so, our sentences make both *de re* and *de dicto* readings true; if not, the sentence is only true with a *de re* interpretation. Second, since *de re* interpretation is evaluated in the real world and *de dicto* in a belief world, we created an explicit question before the target sentence such that **the question asks about either real world facts or the status of the belief world**. With a between-subjects design, we hypothesize that an interaction effect exists--where the sentence only permits *de re* reading but the QUD targets the belief world, the agreement rating would be lower. Third, we compared between subjects **the use of *believe* and *say* verbs** in terms of this *de re/de dicto* accessibility, predicting that since *say* prefers literal report and the description of the target entity is changed in the report, *say* would have overall lower agreement rating. In the end our 2x2x2 entirely between-subjects design has its dependent variable as participants' agreement with a belief report statement. The agreement rate was collected by shifting a slider scale from numerical value 50 to somewhere between 'highly disagree' (0) and 'highly agree' (100). 377 trials remained for analysis after data pruning.

Data Analysis & Result We analyzed two features from the data: (1) the direction to which the slider moved indicates the basic category of agreement or disagreement (Fig.1); (2) the distance from the center indicates how much participants agreed or disagreed (Fig. 2). Both figures indicate that all conditions elicit more agreement than disagreement, supporting the availability of *de re* interpretation. More specifically, in the *believe* verb condition with only *de re* interpretation and a belief-world emphasis, people were significantly less likely to agree than when both *de re* and *de dicto* are set to be true (Fisher's exact test: odds ratio = 0.122, $p^{***} < .001$; odds ratio = 0.319, $p^* < .05$, Fig.1), presenting a *de re* reading with decreased

inaccessibility. Furthermore, *believe* elicited more agreement than *say* in the *de dicto* + *de re* condition with real-world emphasis (odds ratio = 4.583, $p^* < .05$). Yet through a mixed-effects logistic regression model which predicts the binary ‘direction’ data with predictors including belief holder’s knowledge, QUD, verb type, and their interactions, we only found a main effect of belief holder’s knowledge ($\beta = -0.97$, $z = -2.79$, $p = .005$) with no interaction effect as predicted.

Conclusion Clearly, the accessibility of *de re* reading can be manipulated by contextual factors such as the belief holder’s knowledge of the entity. Our next step is to increase the number of tested stories for a clearer interpretation of our predicted interaction and to test the accessibility of *de dicto* and *de re* interpretation mediated by other intensional operators such as temporal phrases to better understand the principles that affect these readings in the intensional domain.

Table 1: Schema of test conditions within the example of *poem*

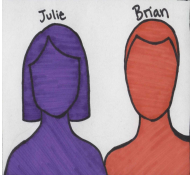

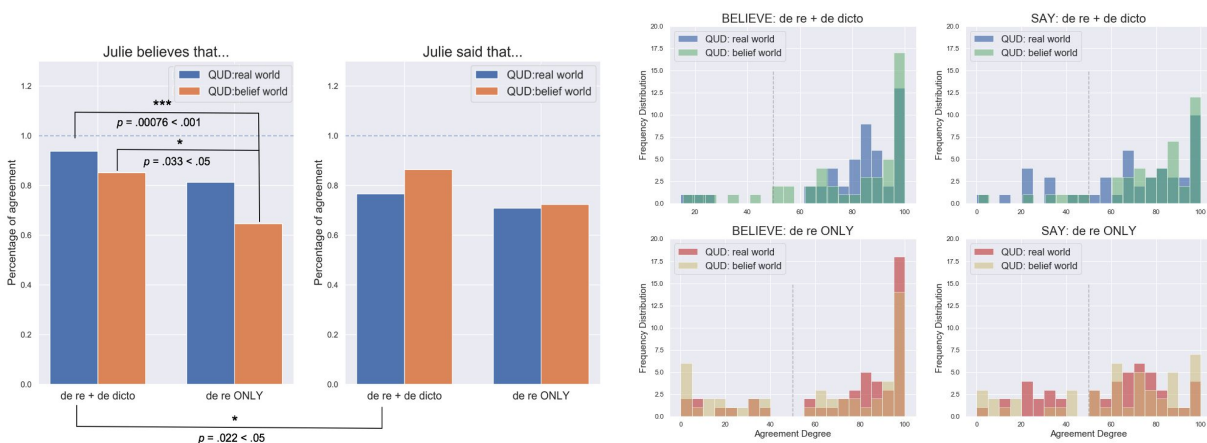
 <p>[CONTEXT] Julie is one of several judges of a poetry competition and Brian is the editor of the poetry collection. Julie encounters a fabulous poem (without knowing who wrote it) and tells Brian: “This poem is awesome. I believe it is going to win the prize.” Brian looks into the submission pile and discovers that Nina was the author.</p> 	
<p>[CONDITION: <i>de re</i> ONLY] However, Brian doesn’t share his discovery with Julie about the authorship of this poem. So Julie doesn’t know who wrote this promising poem.</p>	<p>[CONDITION: <i>de re</i> & <i>de dicto</i>] Brian shares his discovery with Julie about the authorship of this poem. So Julie knows that it was Nina who wrote this promising poem.</p>
<p>[CONTEXT DIALOGUE] Jim: “Do you know which poem will win the prize?” Brian: “I think they are still finalizing the result, but Julie had some idea.”</p>	
<p>[CONDITION: QUD-real world emphasis] Jim: “So which one will win?”</p>	<p>[CONDITION: QUD-belief world emphasis] Jim: “So what does Julie think?”</p>
<p>[CRITICAL SENTENCE expected to have different readings under different conditions] Brian: “Well, she believes/said that the poem that Nina wrote will win the prize.”</p>	

Figure 1 (left): Percentage of agreement under 2x2x2 design (with Fisher’s exact test score)

Figure 2 (right): Distribution of distance from center under 2x2x2 design



References: Cresswell & von Stechow. (1982). Keshet. (2008). Keshet & Schwarz. (2019). Marti, L. (2006). Nelson, M. (2019). Pearson, H. (Forthcoming). Percus, O. (2000). von Fintel, K., & Heim, I. (2011).