The behavioral cost of building composite referents in reading comprehension

Introduction. To comprehend real-life texts, the reading mind needs to keep track of key information across an unfolding discourse. On one view, as natural language unfolds, the mind continuously builds and updates a discourse-level situation model. [1] Consider this example: "*The cook was washing some pots. They were so greasy that she had to use a sponge of steel wool. It made her job much easier.*" Here, 'cook', 'pots', and 'sponge' introduce *discourse referents*—elements that can be referred to ('*They* [=pots] *were so greasy.*'). Prior work shows that introducing new referents and accessing old ones are behaviorally [2] and neurally [3] distinct operators. But languages can also refer to more complex entities. For instance, the pronoun 'it' refers not to the simple referents 'sponge', 'steel', or 'wool', but rather to 'a sponge of steel wool'—a composite referent. Indeed, prior work shows that humans do interpret some pronouns as referring to composite referents, not simple ones. [4] However, it remains unclear whether the mind builds composite referents from simple referents in the situation model, or what the behavioral cost associated with such an operator is.

Methods. 43 participants (ages: 28.5±4.3; 23 female) read 72 short, 5-sentence stories sentence by sentence in a self-paced reading paradigm. Across 3 conditions, all sentences in a set were identical, except for the 4^{th}, critical sentence (Table 1), which ended with 3 simple referents (***simple₃***: 'wool, sponges and steel'), 2 simple referents (***simple₂***: 'steel wool and sponges') or 2 simple referents that form a composite referent (***composite***: 'sponges of steel wool'; Fig. 1). Crucially, lexical items and number of words were identical across conditions. Conditions in one set were divided and counterbalanced across 3 lists; each participant saw one list. After each story, a True/False comprehension task appeared, targeting any sentence with equal probability. We hypothesized that if the reading mind indeed builds composite discourse referents, then ***composite*** reading times (RTs) should be longer than ***simple₂*** RTs; and based on prior work, ***simple₃*** RTs should also be longer than ***simple₂*** RTs, due to one additional simple referent. Finally, we hypothesized that if introducing simple referents and building composite ones are distinct operators, this *could* lead to unequal RTs (***composite***≠***simple₃***).

Results. We regressed critical sentence RTs against nuisance regressors (Table 2): list, age, gender, number of words, average word frequency, syntactic complexity (3 regressors from a Principal Component Analysis over 11 metrics), daily reading hours, and random intercepts per subject and stimulus set. Compared to this reduced model, adding our experimental condition significantly improves model fit (p=.00213; Fig. 2; Table 2). Pairwise comparisons (Fig. 3; Table 2) revealed that building a composite referent increases RTs (***composite***>***simple₂***; p=.0012); this holds when equalizing—across conditions—whether the first simple referent has one word (sponge) or two words (steel wool). Adding an extra simple referent also increases RTs (***simple₃***>***simple₂***; p=.0351). Follow-up two one-sided tests (TOST [5]) for equivalence failed to reject the hypothesis that the difference between ***simple₃*** and ***composite*** is smaller than the smallest effect size of interest, which we set to 50ms ($p_{diff<50ms}$=.343; $p_{diff>-50ms}$=.028). Similar analyses on the outro sentence (#5) revealed no significant effect of Condition (Fig. 2).

Discussion. Our results suggest that, in addition to tracking simple referents, the reading mind also strings them into composite referents. We also replicate prior findings showing that adding an extra simple referent increases RTs. However, it remains unclear whether building a composite referent is the same mental operator as adding a simple referent; future studies may tackle this question using a different paradigm (e.g., neurally). Our results suggest that models of discourse-level comprehension should account for composite referents. They also prompt the question: Does the reading mind build composite referents recursively? Consider, for example, "*John's student likes pets. Her dog is very cute. Its paws are tiny.*" Here, '*its [=John's student's dog's] paws*' theoretically refers to a hierarchically deeper composite referent than '*her [=John's student's] dog*'. Does this referent hierarchy matter behaviorally and/or neurally?

Table 1. Example stimulus set. All sentences were identical except for the critical region in the 4th sentence.

	Sentence		Condition
1	John visited the new abstract art exhibition yesterday afternoon.		
2	On display were many innovative and original art pieces.		
3	He saw a painting made of vivid colors and swirling shapes.		
4	He also saw a sculpture made from	wool, sponges and steel.	***simple***$_3$
		steel wool and sponges.	***simple***$_2$
		sponges of steel wool.	***composite***
5	It was a very popular exhibit with quality pieces.		
	The colors in the painting were lively. (True / False)		TASK

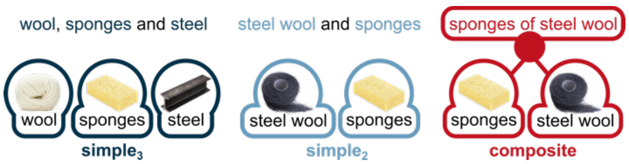

Fig. 1. Graphic representation of the referent structure in the three experimental conditions. **simple**$_2$ and **simple**$_3$ involve introducing two and three simple referents (icons), respectively. **composite** involves introducing two simple referents that make up a composite referent (red circle).

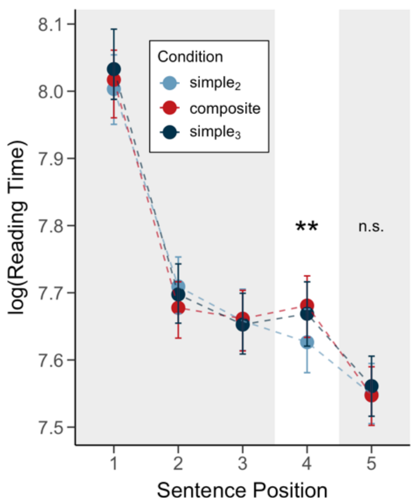

Fig. 2. Mean and standard deviation of log-transformed RTs per sentence, per condition. Significance symbols refer to the effect of adding Condition to the reduced model (Table 2).

Table 2. Statistical model and regression results. Log-transformed reading times (RT) in sentence 4 are modeled using a reduced (gray) and a full model (gray+**Condition**). Continuous variables (italics) are scaled and centered; categorical variables appear in bold. For the experimental manipulation (Condition), ***simple***$_2$ is the base level. Regression estimates and corresponding t values are shown for intercept and for fixed effects with $|t| > 2$. Model comparison with likelihood ratio test compared the full and reduced models. Post-hoc pairwise comparisons are over estimated marginal means (corrected for multiple comparisons with the Tukey method).

Model	log(RT$_4$) ~ Intercept + **Condition** + **List** + *Num_of_words* + *Syntax_complexity* + *Average_frequency* + *Age* + **Gender** + *Reading_hours_per_day* + (1 | Participant) + (1 | Set)	
Fixed Effect (full model)	**Estimate**	***t* value**
Intercept	-0.2432	-1.249
Condition (***composite***–***simple***$_2$)	0.1158	3.353
Condition (***simple***$_3$–***simple***$_2$)	0.0827	2.500
Num_of_words	0.1576	4.424
Model comparison (full–reduced): $\chi^2(2) = 12.305$, $\boldsymbol{p = 0.00213}$ **(**)**		
Pairwise comparisons:	***composite***–***simple***$_2$, $\boldsymbol{p = 0.0012}$ **(**)**	
	simple$_3$–***simple***$_2$, $\boldsymbol{p = 0.0351}$ **(*)**	
	composite–***simple***$_3$, $p = 0.4661$	

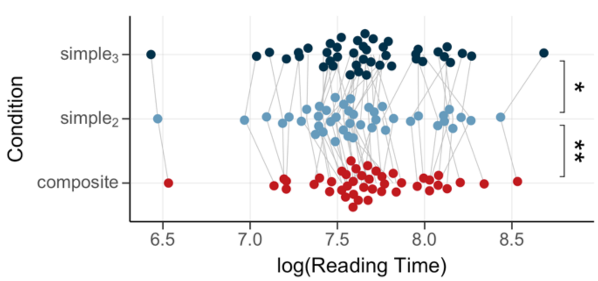

Fig. 3 (bottom). Log-transformed RTs on critical sentence by condition. Each dot represents average RT of one participant per condition. Gray lines connect each participant's data. Significance symbols refer to pairwise comparisons over estimated marginal means (Table 2).

References: [1] O'Brien, E. J., & Cook, A. E. (2015). Models of discourse comprehension. *Handbook on reading*, 217-231. **[2]** Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Mem & Cogn*, *12*, 489-497. **[3]** Coopmans, C. W., & Nieuwland, M. S. (2020). Dissociating activation and integration of discourse referents: Evidence from ERPs and oscillations. *Cortex*, 126, 83-106. **[4]** Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *J Mem Lang*, 53(2), 292-313. **[5]** Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Adv Meth Prac Psych Sci*, 1(2), 259-269.