Predicting Discourse Relations: The Processing Benefit of a Connective

Marian Marchal (Saarland University), Merel Scholman (Utrecht University), Ted Sanders (Utrecht University), Vera Demberg (Saarland University)
marchal@coli.uni-saarland.de

Rationale. Comprehenders make predictions at various linguistic levels [e.g. 1, 2]. To illustrate, when hearing *Lizzy was tired,* one might make predictions about whether the speaker will next discuss the cause or the consequence of Lizzy's tiredness (a relation prediction), what a specific consequence might be (a semantic prediction; e.g. *drink coffee, go to bed*), and how a specific consequence will be formulated (a lexical-syntactic prediction; e.g. *make a cappuccino, took a sip of her coffee).* It is unclear, however, to what extent predictions of discourse relations (DRs) [3] influence processing beyond semantic or lexical-syntactic predictability, since these factors are confounded in previous work. Here, we examine (1) whether DR predictability explains processing difficulty *beyond* other levels of predictability and (2) whether the processing benefit provided by a connective [e.g. 4, 5] can be explained by enhanced prediction or has an additional effect.

Method. We operationalize DR predictability as **relation surprisal** (RS), the negative log probability of the DR type given the context, and take **semantic information value** (SIV) [6] as a measure of semantic predictability. These were calculated based on continuations in a human (n = 160) cloze task. **GPT2 surprisal** (GS) without context served as an estimate of lexical-syntactic predictability. We conducted a region-by-region self-paced reading (SPR) study (n=121) as well as an eye-tracking-while-reading (ET) study (n=79), in which native English speakers read 24 target stories containing cause-consequence sentence pairs as in Table 1. In the explicit but not the implicit condition, these DRs were marked with the connective *therefore*. We analyzed log-transformed response times (RT) from SPR and first-pass (FP) and total fixation (TF) duration from ET. Using mixed-effects piecewise structural equation modeling (pSEM, Figure 1) [7], we estimated the direct and indirect effects of the predictors of interest, while controlling for trial and length.

Results. First, we examine how connective presence influences predictability (see Table 2). As expected, RS is higher in the implicit condition. There was no significant effect of connective on GS, but RS predicts SIV, and as such the connective indirectly facilitates semantic predictions. With respect to processing difficulty, SIV positively predicted all three reading measures, providing evidence for semantic prediction (see Table 3). GS only predicts ET reading measures. Contrary to expected, RS negatively predicted TF, suggesting that more expected relations are read *slower* when accounting for facilitation through semantic prediction. Crucially, there was a significant effect of connective beyond predictability for all measures except for TF.

Conclusion. We show that the connective increases the predictability of upcoming material, and that predictability influences reading times, though sometimes in unexpected ways. The effects of predictability should thus be taken into account when analyzing the facilitating effect of the connective. We find that the connective facilitates processing beyond making upcoming material more predictable.

conn	pred	item text	RS	SIV	GS
Context		Angela used to live in a small flat in Atlanta.			
exp	high	She didn't pay rent for months. She **was evicted**	0.23	0.57	13.16
imp	high	She didn't pay rent for months. Therefore, she **was evicted**	0	0.72	13.82
exp	low	She had over fifteen cats. She **was evicted**	1.15	1.19	13.16
imp	low	She had over fifteen cats. Therefore, she **was evicted**	0	1.21	13.82
Context		… by her landlord. Angela decided to move to a rural area.			

Table 1. Example of an item in each condition, along with relation RS, SIV and GS estimates. Note that the manipulation of predictability was binary, but a continuous measure was included in the analysis.

	predictor	path	type	β	95% CI	
GS	conn	c	direct	.15	[-.01,.30]	
	length	f	direct	.62	[.48,.72]	*
SIV	conn	b	direct	.02	[-.20,.24]	
	length	e	direct	.27	[-.11,-.41]	*
	RS	d	direct	.32	[.18,.47]	
	conn	ad	indirect	-.16	[-.25,-.09]	*
RS	conn	a	direct	-.50	[-.58,-.41]	*

Table 2. (In)direct effects of the presence of a connective, the predictability measures and length on the different predictability measures. These are independent of the reading time measures. * indicates significance at the .05 level. Paths refer to Figure 1. Connective was deviation-coded (imp: -1; exp: 1).

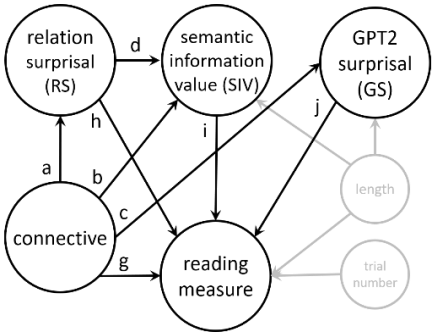

Figure 1. Structure of the pSEM. Note that in the model for the ET measures, there was an additional path between trial number and GS.

			RT			FP			TF		
predictor	path	type	β	95% CI		β	95% CI		β	95% CI	
conn	g	direct	-.06	[-.08,-.03]	*	-.10	[-.15,-.05]	*	-.06	[-.10,-.01]	*
GS	j	direct	.02	[-.01,.05]		.10	[.02,.15]	*	.07	[.03, .15]	*
SIV	i	direct	.03	[.01,.05]	*	.08	[.04,.13]	*	.14	[.10,.19]	*
RS	h	direct	.02	[-.01,.04]		-.07	[-.13,-.03]	*	-.02	[-.07,.03]	
conn	ah+adi+...	indirect	-.01	[-.02,.01]		.04	[.01,.08]	*	.00	[-.04,04]	
RS	di	indirect	.01	[.00,.02]	*	.03	[.01,.05]	*	.05	[.02,.08]	*
conn	g+ah+…	total	-.07	[-.09,-.05]	*	-.06	[-.10,-.01]	*	-.06	[-.10,-.01]	*
RS	h+di	total	.03	[.01,.05]	*	-.05	[-.10,-.00]	*	.03	[-.04,.08]	

Table 3. Direct, indirect and total effects of the predictors of interest on the various reading measures. The estimates for trial and length are not presented due to lack of space. Paths refer to Figure 1.

References

[1] Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31,* 32–59. **[2]** Heilbron, M., Armeni, K., Schoffelen, J. M., Hagoort, P., & De Lange, F. P. (2020). A hierarchy of linguistic predictions during natural language comprehension. *Proceedings of the National Academy of Sciences*, 119(32). **[3]** Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics,* 25(1), 1-44. **[4]** Cozijn, R., Noordman, L. G., & Vonk, W. (2011). Propositional integration and world-knowledge inference: Processes in understanding 'because' sentences. *Discourse Processes, 48*, 475–500. **[5]** Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes, 52,* 47–76. **[6]** Giulianelli, M., Wallbridge, S., & Fernández, R. (2023). Information value: Measuring utterance predictability as distance from plausible alternatives. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 5633–5653). Singapore: ACL. **[7]** Lefcheck, J. S. (2016). piecewisesem: Piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, 7, 573–579.