Discourse marker predictions in reading are not lexical
Merel Scholman, Hannah Rohde & Vera Demberg

1. Introduction

A large body of research has provided evidence that readers and listeners generate expectations about upcoming content, and in some cases even predict specific words, during language comprehension (for a review, see Kuperberg & Jaeger, 2016). The predicted content tends to be processed faster once encountered, presumably due to the pre-activation of concepts or specific words. However, research on predictive processing has focused mainly on processing at the level of syntax and semantics. In the current contribution, we focus on discourse relation predictions, which require comprehenders to establish semantic-pragmatic dependencies (such as `cause-consequence`, `contrast`) within and between sentences, as well as across passages.

Previous work suggests that comprehenders can indeed make predictions about upcoming discourse relations (e.g., Asr & Demberg, 2020; Barthel, Tomasello & Liu, 2022; Blumenthal-Dramé, 2021; Crible, 2021; Dery & Koenig, 2015; Hoek, Rohde, Evers-Vermeul & Sanders, 2020; Rohde & Horton, 2014; Scholman, Rohde & Demberg, 2017; Scholman, Demberg & Sanders, 2020; Schwab & Liu, 2020; Xiang & Kuperberg, 2015; Köhne-Fuetterer, Drenhaus, Delogu & Demberg, 2021; Yi & Koenig, 2021). Such relations can be explicitly signalled by discourse connectives and cue phrases such as *although* or *as a result*, but they need not be. Most studies have focused on whether readers can use a discourse connective (e.g., *because*, *however*) to predict upcoming dependencies and upcoming content, but not whether the prediction pertains to a specific upcoming discourse marker. This is what is investigated in the current contribution.

Whether readers can generate expectations that take the form of specific lexical predictions is still an unresolved question. Results from a range of studies are taken to indicate that comprehenders can indeed predict specific lexical items in certain contexts (DeLong et al., 2005; Köhne-Fuetterer et al., 2021; Van Berkum et al., 2005; Wicha et al., 2004). These studies focus mainly on predictions of gender noun class or phonological form. However, recent studies have challenged this view and suggested that comprehenders may rely more on general expectations and statistical regularities rather than precise lexical predictions. For example, Nieuwland et al. (2018) failed to replicate the effect found in DeLong et al. (2005) in a large-scale replication study.

The current study investigates whether comprehenders generate lexical predictions of specific upcoming contrast markers or content-based expectations of upcoming relations (irrespective of the type of marker that is used). We test this in Dutch, which has multiple pairs of two-part discourse markers:

- "aan de ene kant" (similar to *on the one hand* and referred to from now on as OT1H) and "aan de andere kant" (*on the other hand*, OTOH);

- "enerzijds" (also equivalent to *on the one hand*, referred to as EZ from now on) and "anderzijds" (also meaning *on the other hand*, referred to as AZ).

The two pairs are highly similar in meaning and usage. These markers typically co-occur with their parallel partner to signal a contrastive dependency: 73% of OT1H occurrences are followed specifically by OTOH, and 88% of EZ occurrences are followed by AZ in the Dutch Sonar corpus (Oostdijk et al., 2013). However, mixing the pairs (e.g., EZ~OTOH) can happen in natural language as well (1% of occurrences).

We test what form predictions take in a self-paced reading study with three conditions: OT1H~OTOH, EZ~OTOH, and no marker for the first contrastive element followed by OTOH. Under a lexical prediction account, readers pre-activate specific discourse markers based on the form of the marker used to express the first contrastive element, in which case we should see faster reading times on OTOH following OT1H compared to EZ. However, under the relation prediction account, people should expect a contrast relation rather than a particular surface form, in which case any contrastive marker should suffice.

2. Method

Participants. 138 native speakers of Dutch (age range 18-52 years; mean age 27 years; 59 female) participated in this experiment. Data from two additional participants were excluded because they didn't pass the comprehension question accuracy check (they scored <70%). Participants were recruited via Prolific and received 3.00 GBP for their participation. They were unaware of the purpose of the experiment.

Materials. The experimental stimuli consisted of twelve passages.[1] Each item consisted of an introductory sentence, followed by contrast1 and contrast2. Contrast2 was always marked with OTOH; contrast1 was marked with OT1H, marked with EZ, or unmarked. The three conditions are illustrated in Passage 1.

(1) De backpackster was haar trip naar Hawaii aan het plannen.
    a. **Aan de ene kant** leek het haar leuk om daar te surfen.   [OT1H]
    b. **Enerzijds** leek het haar leuk om daar te surfen.   [EZ]
    c. Het leek haar leuk om daar te surfen.   [∅]
    Aan de andere kant leek het haar ook leuk om te snorkelen in de oceaan.
    (The backpacker was planning her trip to Hawaii. [Aan de ene kant / Enerzijds / ∅] / she thought it would be fun to surf there. On the other hand, she thought it also seemed fun to snorkel in the ocean.)

The items were interspersed with 60 filler items for an unrelated study. The stimuli were counterbalanced across three lists, with each story appearing in a different condition in each list. The participants were randomly assigned to one of the lists.

2.1. Procedure

Participants were recruited via Prolific, after which they were directed to a website hosted by PC Ibex (Schwarz & Zehr, 2021), where they completed the moving window self-paced reading experiment. Participants were presented with a verification statement after 25% of the items. The entire study lasted on average 15 minutes.

Reading times were measured on two regions: "Aan de andere kant" (the target OTOH region) and the three words following the target region (the spillover region). We analyzed the reading times on these regions separately in linear mixed effects models with log-transformed reading times. The variable for condition was contrast-coded using simple coding, with the EZ condition being the reference level. A centered covariate of trial order was included in the models to account for any variance due to participants' reading times speeding up over the course of experiment.[2]

3. Key findings

Figure 1 shows the reading times on the target and spillover region per condition. The results of the linear mixed effect regression models are summarized in Table 1. On the target region, the effect of condition was significant: OTOH was read faster when it followed EZ compared to when no cue was used to express contrast1. Crucially, there was no significant difference between reading times of OTOH following EZ or OT1H. This suggests that comprehenders readily interpret any type of contrastive marker to express contrast2, rather than have a preference for the parallel partner matching the contrast1 marker. The results on the spillover region showed no significant differences between any of the conditions.

---

[1]To reduce the possibility of rapid expectation adaption effects (i.e. learning effects) throughout the experiment (Fine et al., 2013), we presented a larger number of participants with a smaller number of items.
[2]Model specification target region: `log(rt) ∼ trial+cond + (1+cond|subject) + (1+cond|item)`
Model specification spillover region: `log(rt) ∼ trial+cond + (1+cond|subject) + (1+cond||item)`
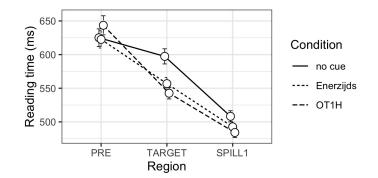
Figure 1: Mean reading times and error bars (SE) per region and condition.

| | Region | | | | | | | |
| | Target | | | | Spillover | | | |
	β	SE	t	p	β	SE	t	p
(Intercept)	6.27	0.02	282.16		6.15	0.02	271.744	
Trial order	-0.00	0.00	-13.28	<**.001**	-0.00	0.00	-14.80	<**.001**
Condition.No cue	0.06	0.01	3.95	<**.001**	0.02	0.01	1.84	.09
Condition.OT1H	-0.03	0.01	-1.90	.07	-0.02	0.01	-1.77	.10

Table 1: Regression coefficients and test statistics from linear mixed-effects models for the effect of condition for each region, Exp. 1. α significant at .025.

4. Discussion

This study was designed to test whether comprehenders predict a specific lexical element to mark contrast2, or whether their predictions are rather at the discourse level, in which case any (appropriate) lexical element marking contrast2 should suffice. The results provide evidence for the latter: OTOH was read faster when contrast1 was marked, but no difference was found between conditions in which EZ or OT1H was used to mark contrast1. The current study therefore does not provide evidence for the all-or-nothing activation of a specific lexical item before it is encountered in the linguistic input. Rather, our results support a graded prediction account, which assumes a formation of partial predictions that do not necessarily include word identity (Luke & Christianson, 2016). As part of the presentation, we will also report on a follow-up eyetracking-while-reading study that investigates the time-course of the facilitative effect of a connective.

In sum, this contribution establishes expectation-driven effects at the discourse level, showing comprehenders' awareness of the discourse dependency established by a discourse marker along with their flexibility in identifying and integrating discourse relations with different markers.

References

Asr, F. T., & Demberg, V. (2020). Interpretation of discourse connectives is probabilistic: evidence from the study of but and although. *Discourse Processes*, *57*, 376–399. doi:`10.1080/0163853X.2019.1700760`.

Barthel, M., Tomasello, R., & Liu, M. (2022). Online comprehension of conditionals in context: A self-paced reading study on wenn ('if') versus nur wenn ('only if') in German. *Linguistics Vanguard*, *8*. doi:`10.1515/lingvan-2021-0083`.

Blumenthal-Dramé, A. (2021). The online processing of causal and concessive relations: comparing native speakers of English and German. *Discourse Processes*, *58*, 642–661. doi:`10.1080/0163853X.2020.1855693`.

Crible, L. (2021). Negation cancels discourse-level processing differences: Evidence from reading times in concession and result relations. *Journal of Psycholinguistic Research*, *50*, 1–26. doi:`10.1007/s10936-021-09802-2`.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117–1121. doi:`10.1038/nn1504`.

Dery, J. E., & Koenig, J.-P. (2015). A narrative-expectation-based approach to temporal update in discourse comprehension. *Discourse Processes*, *52*, 559–584. doi:`10.1080/0163853X.2014.966293`.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PloS one*, *8*, 1–18. doi:`10.1371/journal.pone.0077661`.

Hoek, J., Rohde, H., Evers-Vermeul, J., & Sanders, T. J. (2020). Scolding the child who threw the scissors: Shaping discourse expectations by restricting referents. *Language, Cognition and Neuroscience*, (pp. 1–18). doi:`10.1080/23273798.2020.1852292`.

Köhne-Fuetterer, J., Drenhaus, H., Delogu, F., & Demberg, V. (2021). The online processing of causal and concessive discourse connectives. *Linguistics*, *59*, 417–448. doi:`10.1515/ling-2021-0011`.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*, 32–59. doi:`10.1080/23273798.2015.1102299`.

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive psychology*, *88*, 22–60. doi:`10.1016/j.cogpsych.2016.06.002`.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A. et al. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, e33468. doi:`10.7554/eLife.33468`.

Oostdijk, N., Reynaert, M., Hoste, V., & Schuurman, I. (2013). The construction of a 500-million-word reference corpus of contemporary written Dutch. In *Essential Speech and Language Technology for Dutch: Results by the STEVIN-programme*. Springer Verlag. doi:`10.1007/978-3-642-30910-6\_13`.

Rohde, H., & Horton, W. S. (2014). Anticipatory looks reveal expectations about discourse relations. *Cognition*, *133*, 667–691. doi:`10.1016/j.cognition.2014.08.012`.

Scholman, M. C., Rohde, H., & Demberg, V. (2017). "On the one hand" as a cue to anticipate upcoming discourse structure. *Journal of Memory and Language*, *97*, 47–60. doi:`10.1016/j.jml.2017.07.010`.

Scholman, M. C. J., Demberg, V., & Sanders, T. J. (2020). Individual differences in expecting coherence relations: Exploring the variability in sensitivity to contextual signals in discourse. *Discourse Processes*, *57*, 844–861. doi:`10.1080/0163853X.2020.1813492`.

Schwab, J., & Liu, M. (2020). Lexical and contextual cue effects in discourse expectations: Experimenting with German 'zwar... aber'and English 'true/sure... but'. *Dialogue & Discourse*, *11*. doi:`10.5087/dad.2020.203`.

Schwarz, F., & Zehr, J. (2021). Tutorial: Introduction to pcibex–an open-science platform for online experiments: Design, data-collection and code-sharing. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 15–16). virtual.

Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from erps and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443. doi:`10.1037/0278-7393.31.3.443`.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*, 1272–1288. doi:`10.1162/0898929041920487`.

Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience*, *30*, 648–672. doi:`10.1080/23273798.2014.995679`.

Yi, E., & Koenig, J.-P. (2021). Grammar modulates discourse expectations: Evidence from causal relations in English and Korean. *Language and Cognition*, *13*, 99–127. doi:`10.1017/langcog.2020.29`.