Naturalistic Reading Time Data Support Information Locality

Introduction Characterizing the processing difficulty of a word (or any linguistic unit) has been an essential component of psycholinguistic research [1, 2]. According to *expectation-based theories* [3, 4], the processing difficulty of a word depends on its predictability given the preceding context. Words with higher contextual probability are easier to process. By contrast, *locality-based theories* [5, 6] hold that as the linear distance between two co-dependents increases, the memory representation of the early co-dependent becomes weaker due to decay or interference from other material in memory, making it more difficult to be retrieved and integrated at the later co-dependent. Although both theories have received abundant support [e.g., 7-12], it remains an open question how they can be theoretically and empirically reconciled. Research on the *interactions of expectations and locality* is therefore crucial for building a complete theory of sentence processing, but so far prior work has yielded mixed results [13-15]. To shed light on this issue, we used data from naturalistic reading time (RT) corpora in English [7, 16-18] (**Table 1**) to provide broad-coverage evaluations of two hypotheses (*Information Locality* vs. *Prediction Maintenance*) that make *divergent predictions regarding how expectation and locality interact.*

The ***Information Locality*** hypothesis [19, 20] states that words that highly predict each other are constrained to be close to each other. Thus, locality should be stronger when expectation is high (**Fig. 1**). But the ***Prediction Maintenance*** hypothesis [13] predicts that strong expectation can cancel locality effects (**Fig 2**). When two co-dependents highly predict each other, the cost of retrieval at the later co-dependent will be lower, considering that it might already be preactivated.

Methods We first parsed the texts from the corpora using the Stanford Neural Dependency Parser [21], if parses were not provided by the corpora, and extracted all dependencies (see ex. 1). Following [19], we formalized *expectation* as Head-Dependent Mutual Information (HDMI; see ex. 2) and *locality* as Dependency Length (DL; number of intervening words) [22, 23]. We fitted linear mixed effects models on the log transformed RTs of the later co-dependents (model in ex. 3), with DL, HDMI, and their interaction, and two word-level factors as fixed effects, all scaled. For eye-tracking, first-path duration and total reading times were analyzed. We first ran analyses for each dataset separately. We also ran a meta-analysis collapsing all datasets (for eye-tracking, only total viewing times were included). With the full datasets, we additionally ran exploratory analyses based on head directions (according to UD [24] standards) and whether the dependency involves only core arguments (i.e., verbs and nouns).

Results For all analyses, we report the sign of the coefficient of the model output. A positive sign means that RTs increase as the factor (e.g., DL, HDMI) increases. Results of by-dataset analysis are summarized in **Table 2**. We found that all datasets show locality effects, whereby RTs increase as DL increases. Two datasets also show expectation effects, whereby RTs decrease as HDMI increases. More importantly, two datasets show *information locality* effects, whereby the effect of DL becomes more positive (i.e., leading to more increase in RTs) as HDMI increases. No datasets provide support for the *prediction maintenance* hypothesis. Results of the meta-analysis and exploratory analyses are reported in **Table 3**. The meta-analysis provides evidence for locality and expectation effects, as well as *information locality* effects. In the exploratory analyses, all show support for locality effects, and the majority show expectation and *information locality* effects. Again, no analyses support the *prediction maintenance* hypothesis.

Conclusion Using data from naturalistic RT corpora, we provided broad-coverage evidence for the *information locality* hypothesis: Locality effects are enhanced with high expectations. We showed that naturalistic RT corpora can provide a good source of evidence that corroborates controlled experimentation and be used to test multiple theoretical predictions against each other [10-12]. This is work-in-progress, and we plan to include more English datasets in our analysis, and use Bayesian hierarchical models for the meta-analysis. We also note that current study is limited to English: Effects of locality and expectation and their interaction profiles may differ from language to language. Future work should investigate cross-linguistic datasets.

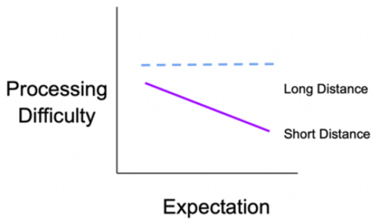


Fig 1: Conceptual predictions of *Information Locality*

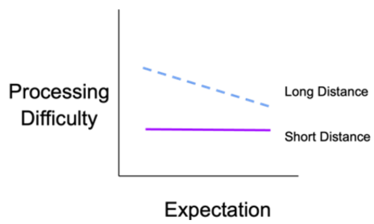


Fig 2: Conceptual predictions of *Prediction Maintenance*

Datasets	Dependencies	Participants
Natural Stories **SPR** (Futrell et al., 2020)	~9.5k	178
Natural Stories **A-Maze** (Boyce & Levy, 2022)	~9.5k	95
Provo **Eye-Tracking** (Luke & Christianson, 2018)	~2.6k	84
Brown **SPR** (Smith & Levy, 2013)	~6.7k	35

Table 1: Datasets included in analysis

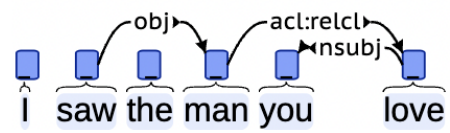
Datasets	DL	HDMI	DL*HDMI
Natural Stories SPR RT	**Positive (p<0.001)**	**Negative (p<0.001)**	**Positive (p<0.01)**
Natural Stories A-Maze RT	**Positive (p<0.001)**	Negative (p=0.109)	Positive (p=0.85)
Brown SPR RT	**Positive (p<0.01)**	Positive (p=0.62)	**Positive (p<0.001)**
Provo ET first path duration	**Positive (p<0.01)**	Positive (p<0.09)	Positive (p=0.59)
Provo ET total reading time	**Positive (p<0.01)**	**Negative (p<0.001)**	Positive (p=0.20)

Table 2: Results of by-dataset analyses.

Analysis	DL	HDMI	DL*HDMI
Full Analysis	**Positive (p<0.001)**	**Negative (p<0.001)**	**Positive (p<0.01)**
Head-Final Dependencies	**Positive (p<0.001)**	Negative (p=0.39)	**Positive (p<0.001)**
Head-Initial Dependencies	**Positive (p<0.01)**	**Negative (p<0.001)**	Negative (p<0.51)
Core Dependencies	**Positive (p<0.001)**	**Negative (p<0.001)**	**Positive (p<0.01)**
Non-Core Dependencies	**Positive (p<0.001)**	**Negative (p<0.001)**	**Positive (p<0.01)**

Table 3: Results of meta-analysis and exploratory analyses

1. Example of a UD parse [24] containing a head-final *nsubj* relation (i.e., subject-verb dependency), a head-initial *obj* relation (i.e., verb-object dependency), and a head-final acl:relcl (i.e., RC head-RC verb dependency)



2. HDMI = $\log \frac{p(h,d)}{p(h)p(d)}$
   - p(h,d) stands for how many times that pair occurs together in a dependency
   - p(h)p(d) stands for how many times in total the head and the dependent occur in the corpus
   - We use these estimates in the HDMI formula to compute HDMI for any given pair of word categories. For word categories, we use more fine-grained part-of-speech tags from UD [24]

3. <u>*Example formula*</u> (for Natural Stories SPR): log(RT) ~ DL * HDMI + Word_Length + Frequency + (1|Subject)

References [1] Miller & Chomsky (1963) Handbook of Mathematical Psychology. **[2]** Frazier & Fodor (1978) Cognition **[3]** Hale (2001) NAACL **[4]** Levy (2008) Cognition **[5]** Gibson (1998) Cognition. **[6]** Lewis & Vasishth (2005) Cogn. Sci. **[7]** Smith & Levy (2013) Cognition **[8]** Grodner & Gibson (2005) Cogn. Sci. **[9]** Wu, Kaiser, & Vasishth (2018) Cogn. Sci. **[10]** Demberg & Keller (2008) Cognition. **[11]** Shain et al. (2016) CL4LC. **[12]** Shain et al. (2023) HSP. **[13]** Husain, Vasishth, & Srinivasan (2014) Plos One. **[14]** Schwab, Ming, & Liu (2022) JEP:LMC **[15]** Safavi, Husain, & Vasishth (2016) Front. Psychol. **[16]** Futrell et al. (2020) Lang. Resour. Eval. **[17]** Boyce & Levy (2022) Glossa Psycholinguistics. **[18]** Luke & Christiansen (2018) Behav. Res. Methods. **[19]** Futrell et al. (2019) Depling. **[20]** Futrell, Gibson, & Levy (2020) Cogn. Sci. **[21]** Chen & Manning (2014) EMNLP. **[22]** Liu (2008) J. Cogn. Sci. **[23]** Futrell Mahowald & Gibson (2015) PNAS. **[24]** Nivre et al. (2016) LREC'16.