

# Once Bitten, Twice Shy:

## The Impact of Predictive Validity on Anticipatory Processing During Sentence Comprehension

Sofia Pelica<sup>1</sup>, and Sofia Frade<sup>2</sup>

<sup>1</sup>Iscte - University Institute of Lisbon; <sup>2</sup>Iscte - University Institute of Lisbon, CIS-Iscte

The father told him to go brush his **eggs**.

### BACKGROUND

Readers can use contextual constraints to predict upcoming words.

**Word Expectancy** is the likelihood of a specific word being predicted based on the sentential context.

(e.g., in the given illustration, teeth is more likely to be predicted than eggs; also known as cloze probability; Taylor, 1953)

**Prediction Validity** refers to the overall likelihood of encountering a correct prediction in the experiment as a whole.

(e.g., the validity would be higher if the participant had read a considerable number of congruent sentences; Brothers et al., 2017)

Congruent Sentential Context

Higher Word Expectancy & Facilitates Processing

(Brothers et al., 2015, 2017; Huettig, 2015)

### Is the activation of lexical predictions *automatic* or *strategic*?

**Strategic** because generating specific predictions at the form level may require additional processing, which...

- may be metabolically costly
- and readers may suppress it when unfavorable.

(e.g., in a low validity environment; Cevoli et al., 2022; Heyman et al., 2015; Ito et al., 2016; Kuperberg & Jaeger, 2016)

**Automatic** because readers unconsciously activate associated concepts while processing a sentence, thus...

- with a constraining context, pre-activation occurs even at the phonological/orthographic level.

(e.g., Huettig, 2015; Nieuwland, 2019; Pickering & Gambi, 2018)

Support from **Brothers et al. (2017)**'s findings.

Need for Replication

But other findings support the automatic-activation.

(e.g., DeLong et al., 2018)

### Replication of Word Expectancy and Prediction Validity effects.

#### REPLICATION OF BROTHERS ET AL. (2017)

99 participants (90 females)

$M_{age} = 22.64$ ,  $SD_{age} = 3.72$

Participants were randomly allocated to **high** or **low** validity conditions.

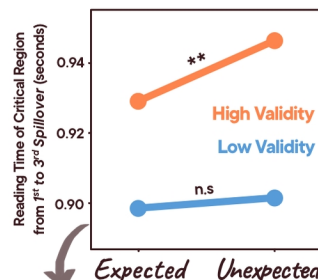
**1st Block**  
120 expected  
120 unexpected

**2nd Block**  
30 expected +  
30 unexpected

A self-paced reading task on Pavlovias was used.

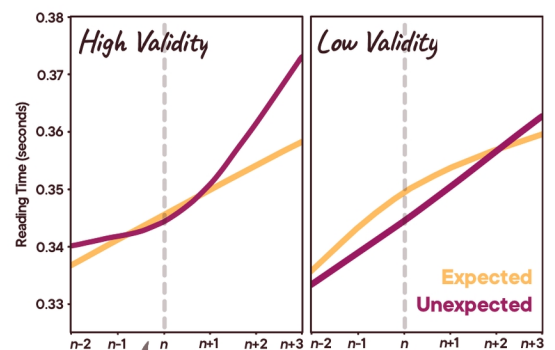
Only critical and spillover words from the 2nd Block were analyzed.

Expected: Antes de cozinhar a Antonia vestia sempre um avental **que foi bordado** à mão.  
Unexpected: Como ficou em segundo lugar também recebeu um avental **que foi bordado** à mão.



Interaction Effect was replicated,  $b = 0.02$ ,  $t(5802) = 2.06$ ,  $p = .039$ .

**Unexpected Words** led to significantly longer reading times than **Expected Words**,  $b = -0.02$ ,  $p < .001$ , with a stronger effect in the **High Validity** condition (but validity alone did not have a significant effect).



The Non-Linear Predictive Relationship of Reading Time throughout the Critical and Spillover Words was significant for all terms ( $p < .001$ ).

- (1) Readers reduce their anticipatory processing in low validity environments,
- (2) anticipatory processing can become costly in high validity environments, and
- (3) predictive validity alone may not fully explain these effects. Other factors may play a role!

**Limitations:** We did not consider other factors such as sentence complexity and reading proficiency.

APPE

Put me in your pocket



sofiapelica.github.io