**Background**. This study directly compares two proposed measures of morphological processing, incremental morpheme surprisal and static relative entropy, while introducing new, thus far untested, variables into these models through a lexical decision task in Spanish. Since the dual-route debate [1-2], morphological processing theories have shifted towards an information theoretic explanation of the processing phenomena. Two main models, implementing processing in terms of incremental predictive steps or static information gain, have emerged but have not yet been directly compared. **Surprisal** models [3-4] are defined as the incremental expectation of upcoming morphemes. In the morphological processing literature, these models have largely been explored in Magnetoencephalographic (MEG) neurolinguistic studies [5-14]. The measure has been used as an index of morphological decomposition and has been correlated with M170 amplitude (visual word processing component). These studies primarily compute surprisal as the conditional probability of encountering a morpheme given a preceding morpheme. **Static Relative Entropy (RE)** models measure the amount of information, in bits, that a stem carries, quantified by the deviation between the stem-relative probability distribution of the stem's inflected forms and the class level distribution of identically inflected words. This measure has been correlated with response times across languages and parts of speech [15-23]. Despite being a priori compatible with an incremental processing model, this formulation of RE is explicitly static. It is defined as information associated with each stem, regardless of the stem's morphological ending—thus making no predictions about the local relationships between morphemes.
**Methods.** Participants conducted a visual lexical decision task either online (n=54) or in-person with simultaneous MEG recording (n=23) in Spanish. The task consisted of 108 verbal and pseudo-stems (distributed across 3 inflection classes), conjugated with 13 different endings (fig. 1). The trials were split into 2 and 3 suffixed data depending on the number of suffixes associated with each ending. The MEG task consisted of an 84 stem subset of the task.
**Models.** (fig. 2) The simple RE model included a static RE measure per stem (in line with models from [15-23]). The full RE model extended previously proposed RE models by including a suffix frequency variable, which provides an additional axis of information available to the processing system per word. The local Surprisal model included transition probabilities of all adjacent morphemes (in line with models from [5-14]). The nonlocal Surprisal model included all variables from the local Surprisal model along with a nonlocal transition probability between the stem and final (person/number) suffix—a predictive relationship that is predicted from statistical parsing of the data [24] but not predicted from linguistic theory [25-26].
**Behavioral Results.** (fig. 3-5) Linear mixed effect models were fit to response times. Models were compared using AIC [27]. Full RE and local Surprisal models outperform control & simple RE models, but do not differ amongst themselves. Nonlocal Surprisal model outperforms the models for 2 suffixed data but doesn't differ from the local Surprisal model for 3 suffixed data.
**MEG Results**. Data collection is finished. However, due to issues with the localizer task [28], localization of the region implicated in morphological decomposition is ongoing.
**Discussion.** The model comparison shows that the nonlocal Surprisal model outperforms most other models for the behavioral data, suggesting that morphological processing is incremental and predictive—making use of both local and nonlocal predictions. The similarity in variance explained by the full RE and local Surprisal models stems from the addition of the suffix frequency in the RE model. Suffix frequency, is, however, likewise compatible with an incremental processing model. In a post-hoc comparison, the local Surprisal model with the additional suffix frequency outperformed the full RE model. Future analyses will focus on the MEG data, to test whether only local surprisal, as predicted by localist theoretical models of morphology [25-26], or both local and nonlocal surprisal, as predicted by purely statistical parsing [24], correlates with the M170 component. Despite this ongoing analysis, the behavioral data points to incremental models consistently outperforming static information theoretical models in single word processing in the visual modality.

**Fig. 1. Stem Endings of Experimental Stimuli**

| Data | Tense | -AR Class | | -ER Class | | -IR Class | |
|---|---|---|---|---|---|---|---|
| 2 Suffixed Data | Imperfect | aba | aba\|mos | ía | ía\|mos | ía | ía\|mos |
| | | aba\|s | aba\|n | ía\|s | ía\|n | ía\|s | ía\|n |
| | | aba | aba\|n | ía | ía\|n | ía | ía\|n |
| | Future | ar\|é | ar\|emos | er\|é | er\|emos | ir\|é | ir\|emos |
| | | ar\|ás | ar\|án | er\|ás | er\|án | ir\|ás | ir\|án |
| | | ar\|á | ar\|án | er\|á | er\|án | ir\|á | ir\|án |
| 3 Suffixed Data | Conditional | ar\|ía | ar\|ía\|mos | er\|ía | er\|ía\|mos | ir\|ía | ir\|ía\|mos |
| | | ar\|ía\|s | ar\|ía\|n | er\|ía\|s | er\|ía\|n | ir\|ía\|s | ir\|ía\|n |
| | | ar\|ía | ar\|ía\|n | er\|ía | er\|ía\|n | ir\|ía | ir\|ía\|n |

The above chart represents person (1st, 2nd, 3rd) and number (sg., pl.) suffixes per tense (y-axis) and verbal inflection class (x axis). Proposed morphological boundaries are marked here with | .

**Fig. 2. Statistical Models with Relevant Variables**

| Model | Example | Variables of Interest |
|---|---|---|
| Control Model | bail aba s<br>dance past 2.sg | Avg. bi-letter frequency<br>Surface log frequency |
| Simple RE | bail aba s<br>dance past 2.sg | Avg. bi-letter frequency<br>Surface log frequency<br>Static Relative Entropy |
| Full RE | bail aba s<br>dance past 2.sg | Avg. bi-letter frequency<br>Stem log frequency<br>Static Relative Entropy<br>Suffix log frequency |
| Local Surprisal | bail aba s<br>dance past 2.sg | Avg. bi-letter frequency<br>Stem log frequency<br>Transition Probability—Stem to Suffix 1<br>Transition Probability—Suffix 1 to 2 |
| Nonlocal Surprisal | bail aba s<br>dance past 2.sg | Avg. bi-letter frequency<br>Stem log frequency<br>Transition Probability—Stem to Suffix 1<br>Transition Probability—Suffix 1 to 2<br>Transition Probability—Stem to Suffix 2 |

**Fig. 3. Two Suffixed Data Model Comparison**

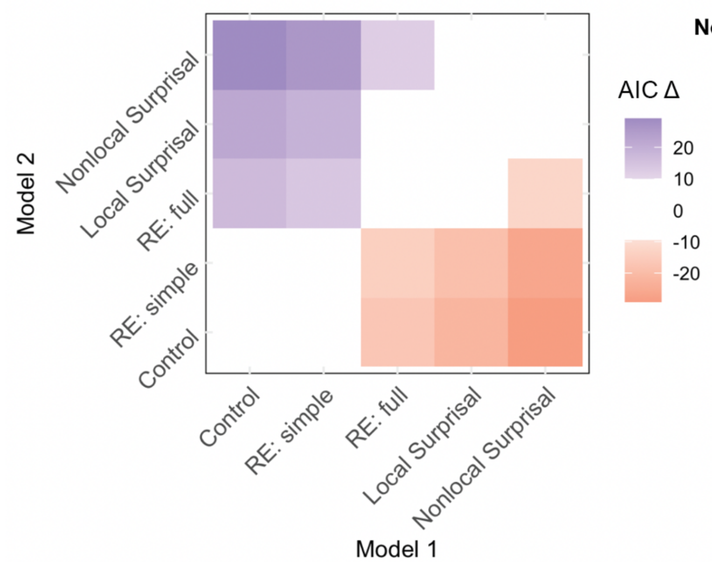


**Fig. 4. Three Suffixed Data Model Comparison**

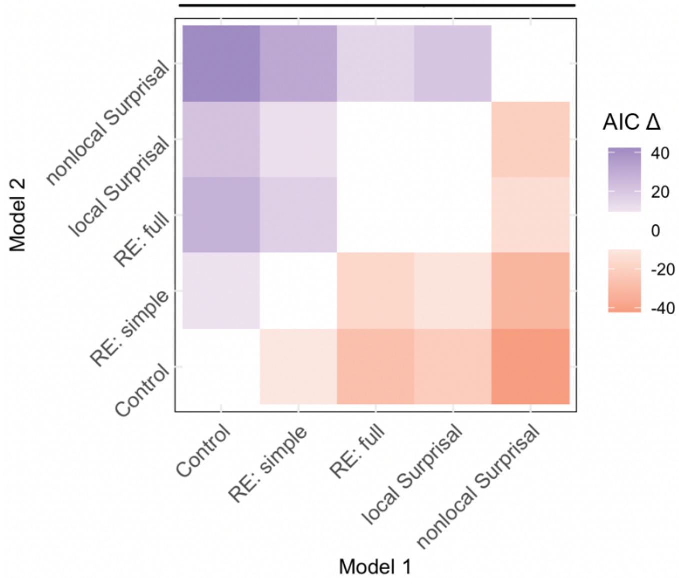


**Fig. 5. Selected Model Output**

| Model | Variable | 2 Suffixed Data | | 3 Suffixed Data | |
|---|---|---|---|---|---|
| | | Estimate | P-val | Estimate | P-val |
| **Simple RE** | RE | 1.994 | 0.63 | -19.698 | <0.05 |
| **Full RE** | RE | 0.774 | 0.85 | -17.942 | <0.05 |
| | Suffix Freq. | -15.377 | <0.01 | -28.832 | <0.01 |
| **Local TP** | TP 1 | -3.727 | 0.46 | -3.254 | 0.686 |
| | TP 2 | -19.356 | <0.01 | -21.082 | <0.05 |
| | TP 3 | | | -19.628 | <0.01 |
| **nonlocal TP** | TP 1 | -2.913 | 0.574 | -3.708 | 0.636 |
| | TP 2 | -20.451 | <0.01 | -23.076 | <0.01 |
| | TP 3 | | | 2.649 | 0.785 |
| | nonlocal TP | 23.486 | 0.442 | -299.594 | <0.01 |

[1] Pinker & Ullman 2002; [2] McClelland & Patterson 2002; [3] Hale 2001; [4] Levy 2008; [5-14] Zweig & Pylkkänen 2009; Solomyak & Marantz 2009; Solomyak & Marantz 2010; Lewis et al. 2011; Fruchter & Marantz 2015; Neophytou et al. 2018; Gwilliams & Marantz 2018; Stockall et al. 2019; Wray et al. 2022; Cayado et al. 2024; [15-23] Kostić et al. 2003; Baayen & Moscoso del Prado 2005; Milin et al. 2009a; Milin et al. 2009b; Nenadić et al. 2016; Hendirx et al. 2017; Filipović-Đurđević & Gatarić 2018; Filipović-Đurđević & Milin 2019; Nenadić et al. 2021; [24] Aristia 2024 (Dissertation) [25] Halle & Marantz 1993; [26] Embick 2010: [27] Akaike 1974; [28] Gwilliams & Marantz 2016;