Theory: Grice’s maxims of quantity [1] requires speakers to choose the one among possible forms that is as informative as necessary for communicating the intended meaning and not to exceed that. The Uniform Information Density [2] further considers incremental perception of language stimuli by the interlocutor as a basis of how information should be distributed across utterances. We apply this theory to the choice of writers on when to use a discourse connective between two sentences in construction of discourse relations given that linguistic features of the first sentence can be predictive of the relation sense.

Hypothesis: highly informative discourse connectives should appear in cases where relational surprisal would be high if the connective is not used.

\[ \text{Surprisal encountering Arg2} = -\log p(\text{Arg2} | \text{Arg1}) = -\log p(\text{Arg2} | \text{R}, \text{Arg1}) - \log p(\text{R} | \text{Arg1}) \]

\[ \text{Given R is the relation between Arg1 and Arg2, i.e., } p(\text{R} | \text{Arg1}, \text{Arg2}) = 1 \]

Corpus study: Chosen alternative relations often have some type of negation in their Arg1 [3]. By analysis of the Penn Discourse Treebank [4] we investigate whether:

1. Negation in arg1 is a strong marker of the relation sense (normalized point-wise mutual information analysis [5])
2. Relational surprisal wrt. this feature is higher in explicit than implicit cases

Chosen Alternative relations with Arg1Neg from PDTB:

“They didn’t panic during the first round of selling this morning. Instead, they bought on weakness and sold into the strength, which kept the market orderly.” --- Explicit

“I would say this is not bad news; [instead] this is a blip” --- Implicit

The UID mechanism applied to incremental inference of discourse relations suggests that the predictability of a relation can be considered as a factor for dropping explicit discourse connective.

<table>
<thead>
<tr>
<th>Arg1Neg</th>
<th>Arg2Neg</th>
<th>Explicit</th>
<th>Implicit</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE</td>
<td>FALSE</td>
<td>14857</td>
<td>12155</td>
<td>27012</td>
</tr>
<tr>
<td>FALSE</td>
<td>TRUE</td>
<td>1975</td>
<td>2153</td>
<td>4128</td>
</tr>
<tr>
<td>TRUE</td>
<td>FALSE</td>
<td>2126</td>
<td>1758</td>
<td>3884</td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>500</td>
<td>518</td>
<td>1018</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>19458</td>
<td>16584</td>
<td>36042</td>
</tr>
</tbody>
</table>

• Negation in the first argument of a discourse relation changes the distribution of upcoming discourse relations.

• In particular, the Chosen Alternative relation is more expected following a negation. Therefore, the explicit connector can be omitted.

\[ P(\text{Chosen alternative} | \text{Arg1Neg}) \]

- Explicit: 1.5% (Significant diff. at p < 0.001)
- Implicit: 5.3%

• Relational surprisal affects writers’ choice of inserting / omitting discourse connectors.

Reference: