Recent developments in formal pragmatics

Part 1/3: Optimality Theory

The Arctic University of Norway, November 2018

Kjell Johan Sæbø

BiOT emerged at the turn of the millennium as a fusion of Radical Pragmatics and Optimality Theoretic Semantics (Blutner 2000). (Benz and Mattausch 2011: 1)

Radical Pragmatics (Levinson 2000, Atlas 2005, building on earlier work): Pragmatics loops back into semantics; implicatures influence content

Optimality theoretic semantics: Choices of interpretations are governed by a competition among alternative candidate interpretations

Blutner (1998, 2000) extended this original version by taking also alternative forms into account that the speaker could have used, but did not. (van Rooij and Franke 2015)

What is optimal is not just interpretations with respect to forms, but rather form-interpretation pairs.

Figure 1: The roots of pragmasemantics

1 Strong optimality and scalar implicatures

Consider the query (1) and the three possible responses in (2).

(1) How often are you satisfied with the quality of the sex?

(2) a. Sometimes.
    b. Often.
    c. Always.

These three responses form an entailment scale, a so-called Horn scale.

Under a definition of optimality like (3) and the four assumptions 1.–4., the pairing of the content sometimes, not often with the form sometimes and the pairing of the content often, not always with the form often are optimal.

(3) Optimality of \(<f,c>\)

A form-content pair \(<f,c>\) is optimal iff for any \(<f',c>\) or \(<f,c'>\),

\[ P(c/\lfloor f \rfloor) \geq P(c/\lfloor f' \rfloor) \text{ and } P(c/\lfloor f \rfloor) \geq P(c'/\lfloor f \rfloor) \]

\(P(c/\lfloor f \rfloor)\) is the probability of the truth of content \(c\) given \(f\)'s literal content.

<table>
<thead>
<tr>
<th>(P(x/\lfloor y \rfloor))</th>
<th>(\exists, &lt;n)</th>
<th>(&gt; n, \neg\forall)</th>
<th>(\forall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sometimes</td>
<td>(\Rightarrow \frac{1}{3})</td>
<td>(\frac{1}{3})</td>
<td>(\frac{1}{3})</td>
</tr>
<tr>
<td>often</td>
<td>0</td>
<td>(\Rightarrow \frac{1}{2})</td>
<td>(\frac{1}{2})</td>
</tr>
<tr>
<td>always</td>
<td>0</td>
<td>0</td>
<td>(\Rightarrow 1)</td>
</tr>
</tbody>
</table>

Table 1: Scalar implicatures as optimal interpretations

The non-zero non-optimal form-content pairs are blocked.

1. Stronger scalemates are relevant
2. Sender authority: common ground that S knows

---

3. No **bias**: prior probabilities are evenly distributed

4. Same **cost**: no significant differences in complexity among scalemates

The first two assumptions are preconditions for the definition (3) to apply; in a *wh* question context like (1), both are given.\(^2\)

The last two assumptions can be lifted to provide more complex measures, to which we will have occasion to return in due course.


## 2 Case study: simple versus complex reflexives

A theory based on Bergeton (2004) and Eckardt (2001) can explain the low acceptability of *self* in (4), but not the necessity of *self* in (5):

(4) Narcissus speiler seg (# selv).
Narcissus reflects SEG (# SELF)

(5) Narcissus beundrer seg # (selv).
Narcissus admires SEG # (SELF)

The keys to explaining both are, first, the focus structures in (6) and (7), second, OT pragmatics applied to the two alternatives in (7) (Sæbø 2009):

(6) a. Narcissus [speiler seg]\(F\).
Narcissus mirrors SEG
b. # Narcissus [speiler]\(F\) [seg selv]\(F\).
Narcissus mirrors SEG SELV

(7) a. # Narcissus [beundrer seg]\(F\).
Narcissus admires SEG
b. Narcissus [beundrer]\(F\) [seg selv]\(F\).
Narcissus mirrors SEG SELV

\(^2\)Here are two cases where relevance and sender authority are not (yet) given:

(i) “Is this tequila distilled twice, as required by the Mexican government?” “Yes.”

(ii) At this point, we can already say that half of the cats found their way home.

The infelicity of (6b) can be explained with its focus presupposition (8):

(8) **Focus presupposition of Narcissus speiler seg selv**
There are propositions \(\phi\) such that \(\exists P \sim \text{speiler}\) and \(\exists y \exists \text{Narcissus}\) such that \(\phi = P(y)(\text{Narcissus})\), and there are propositions \(\psi\) such that \(\exists Q \sim \text{speiler}\) and \(\exists z \sim \text{Narcissus}\) such that \(\phi = Q(z)(\text{Narcissus})\).

This would be satisfied if there were a prior probability that someone speiler someone else or if there were to be alternatives to seg selv in the discourse. That is difficult, but for beundrer the presupposition is easily satisfied.

But the infelicity of (7a) cannot be explained with its focus presupposition:

(9) **Focus presupposition of Narcissus beundrer seg**
There are propositions \(\phi\) such that \(\exists P \sim \text{beundrer seg}\) such that \(\phi = P(\text{Narcissus})\).

This is easily satisfied, regardless of alternatives to seg.

But BiOT can predict that the optimal interpretation of (7a) includes the anti-presupposition that the focus presupposition of (7b) is not justified –

(10) **Focus implicature of Narcissus beundrer seg** (loosely)
There is no prior probable or salient alternative to beundrer, or there is no prior probable or salient alternative to seg.

In other words, the verb should be sufficiently predictable from the reflexive, or vice versa, which is not the case when the verb is ‘admire’.

## 3 M-implicatures and weak optimality

The ‘Division of Pragmatic Labor’ (Horn 2004: 16): “as a result of general pragmatic interactions, unmarked expressions are generally used to convey unmarked messages and marked expressions are generally used to convey marked messages” (Davis and Potts 2010: 42).

### 3.1 The simple and the stereotypical

In the world of Dostoevsky’s *Besy* ‘Demons’, (11) is false while (12) is true.
Recent developments in formal pragmatics
Optimality Theory

Recent developments in formal pragmatics
Optimality Theory

3.2 The brief and the vague

Observation → Round number words tend to have round interpretations (Krifka 2002)

Thus (15a) can be true at the same time as (15b) is false but (15c) is true:

(15)  a. The distance between Tromsø and Vadsø is 400 km.
      b. The distance between Tromsø and Vadsø is 409 km.
      c. The distance between Tromsø and Vadsø is 418 km.

This can be explained in Directional Optimality Theory if it is assumed that
– 409 and 418 are more costly expressions than 400, and
– approximate interpretations are preferred over precise ones.

Again, the pairing of the worse form with the worse content comes out as
weakly optimal – it is deblocked because its two competitors are blocked:

<table>
<thead>
<tr>
<th></th>
<th>±.5</th>
<th>±20</th>
</tr>
</thead>
<tbody>
<tr>
<td>409</td>
<td>⇒.2</td>
<td>.4</td>
</tr>
<tr>
<td>400</td>
<td>.4</td>
<td>⇒.6</td>
</tr>
</tbody>
</table>

Table 3: Brief and vague or else elaborate and precise

Note, though, that Krifka (2007) substitutes a Game theoretic account.

4 Outlook

For two reasons, BiOT has faded into the background of pragmatic theory.

4.1 Assimilation to Game Theory

Over the first decade of the new millennium, BiOT was gradually superseded
by Game Theory as the dominant framework of formal pragmatics (indeed,
Dekker and van Rooij (2000) called BiOT “an application of game theory”).

→ Topic for the third installment (GT and the Rational Speech Act model)!
4.2 A challenge: embedded implicatures

Embeddings (Chierchia 2004) → A local approach is better than a global one

(16) Usually you may only take an apple. So, if you may take an apple or take a pear, you should bloody well be pleased. (Kamp 1973: 279)

Blutner’s response (2006: 11):

I will argue that both approaches can coexist in optimality theoretic pragmatics: a global theory describes the principal forces that direct communication – it has a diachronic dimension . . . ; a local theory describes the actual synchronic dimension – it explains how online, incremental interpretation of complex sentences is possible. The connection . . . results from assuming that the results of global optimization fossilize into a local mechanism . . . (my emphasis)

→ Topic for the second, next installment (The grammatical theory)!

References


