A new computational model of alignment and overspecification in reference

Albert Gatt[1,2], Martijn Goudbeek[1], Emiel Krahmer[1]

[1]*Tilburg University, The Netherlands*; [2]*University of Malta, Malta*

Overspecification

In visual domains, people often overspecify referring expressions. Rather than selecting only those properties of a target which single it out from its distractors, they often include properties which are not strictly required for identification. See Figure 1.

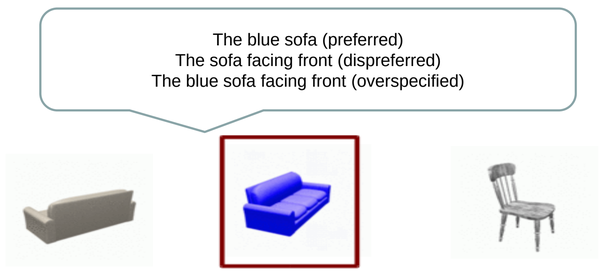

Figure 1. Referential options and overspecification.

Why do people overspecify?

Some properties are salient or preferred and are therefore included in a referring expression even if they have no discriminatory value (Pechmann 1989; Belke & Meyer 2002; Engelhardt *et al.* 2006). One reason for this is that reference production is incremental.

But what happens in interactive situations?

Our hypothesis is that in dialogue, priming and alignment can influence content selection:

- Even dispreferred properties may be more likely to be selected if primed.
- Speakers may be more likely to overspecify if primed to do so.

Alignment

Goudbeek and Krahmer (2010) used the experimental paradigm in Figure 2 and showed that:

- Speakers who are primed with a dispreferred attribute are more likely to use it for later descriptions of different targets.
- Although the primes used were not overspecified descriptions, people often overspecified anyway, using both a preferred and a dispreferred property.

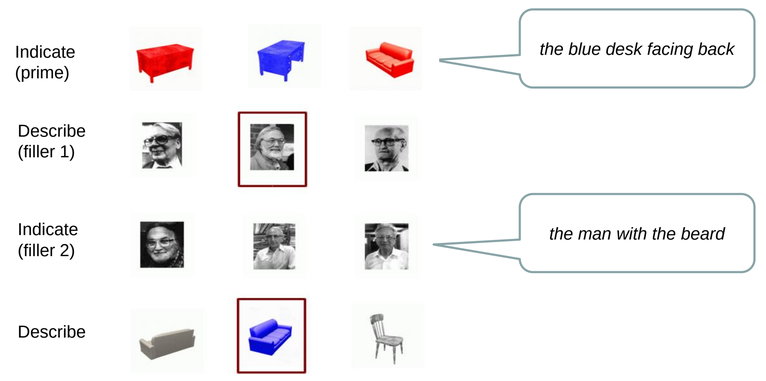

Figure 2. Structure of an experimental trial

Our experiment

- We primed participants with overspecified descriptions containing both a preferred and a dispreferred attribute (e.g. *the blue desk facing back*)
- We asked them to refer to a different object which could be identified using either attribute alone (*the blue sofa*, *the sofa facing front*).
- Participants overspecified over 50% of the time, using both the preferred and dispreferred attribute. See Figure 4.

Modeling

Existing models

Dale and Reiter's (1995) Incremental Algorithm models overspecification by selecting properties of a target in order of preference (e.g. colour before orientation), stopping when the description is distinguishing. The IA only overspecifies if it selects a preferred property before a dispreferred one, when the dispreferred one alone would suffice.

In our experimental scenarios, the IA would never overspecify (Figure 3), because a preferred property is always sufficient.

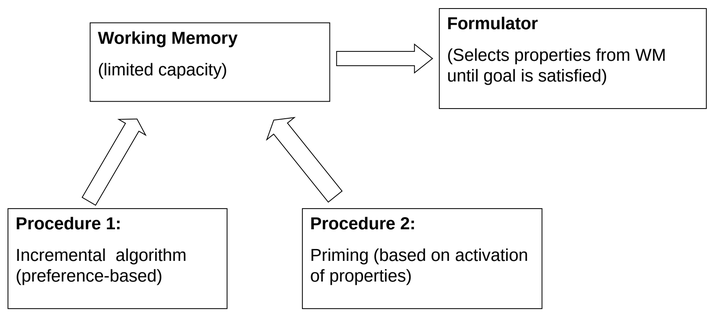

Figure 3. Parallel architecture for reference production

A new model

We developed a new computational model based on parallel, competing processes which select properties and populate a limited-capacity working memory buffer. This combines:

- A preference-based procedure à la Dale and Reiter (1995), which selects properties based on preference/salience.
- A priming-based procedure which selects properties based on their degree of activation in preceding discourse.

Results

Comparing the parallel model to human data

We ran our parallel model on the experimental trials, computing the proportion of overspecified descriptions produced and comparing it to the Incremental Algorithm as our baseline.

We found no statistically significant difference between our model and the human data. By contrast, the Incremental Algorithm has a 0% rate of overspecification.

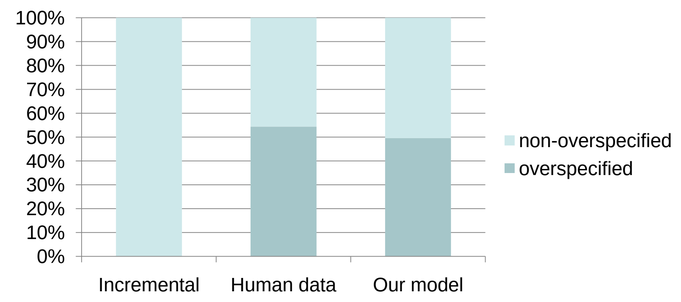

Figure 4. Experimental and modelling results

Conclusions

- Our experiment shows evidence that overspecification is influenced by priming/alignment.
- We have proposed a new computational model which combines preference-based heuristics and priming in a parallel framework.
- Our model matches the human data far better than the Incremental Algorithm.

Future work

- We are seeking to tune the model by learning parameters from the data.
- We are also testing non-deterministic versions of the Incremental Algorithm.

Universiteit van Tilburg

http://bridging.uvt.nl/

University of Malta
L-Università ta' Malta