The role of discriminatory power in reference production

Roger P.G. van Gompel, University of Dundee
r.p.g.vangompel@dundee.ac.uk

In almost all computational models of concept selection during reference production, the discriminatory power of a property plays an important role. Discriminatory power (or informativeness or utility) refers to the number of distractors that a property that is true of the referent rules out, for example, the number of alternative referents that a property such as "red" rules out when an object is referred to as "the red basket".

Models differ in the exact way in which discriminatory power plays a role. In the greedy algorithm (Dale, 1989), properties are selected in order of discriminatory power until a fully discriminating expression has been found. In the incremental algorithm (Dale & Reiter, 1995), discriminatory power plays a less dominant role, because the order of property selection is based on preference (e.g., colour is more preferred than size, independent of discriminatory power). However, discriminatory power plays a role because only properties that rule out at least one distractor are considered. A more recent model, PRO (Van Gompel, Van Deemter, Gatt, & Krahmer, 2019) also assumes that only discriminatory properties are selected, but in addition, it claims that if there are any properties that rule out all distractors, then speakers first select one such property before adding others. Finally, the rational speech act theory (Degen, Hawkins, Graf, Kreiss, & Goodman, 2020) claims that the likelihood with which a property is selected is proportional to the proportion of distractors it rules out.

We conducted two reference production experiments to investigate discriminatory power. In Experiment 1, participants described pictures to a confederate, who had to select the correct object. We manipulated the discriminatory power of the colour of the target object and the borders around it. Each property was either fully discriminatory (ruling out all distractors), partly discriminatory (ruling out only one distractor) or non-discriminatory, see Fig. 1. In all conditions, merely mentioning the object (e.g., basket) was sufficient to identify the target, so the use of either colour or border always resulted in an overspecified expression.

Experiment 1 showed a strong preference for expressions that included colour (e.g., "red basket") rather than border ("basket in triangle"), see Fig. 1. Comparison between the C part – B part and C full – B part conditions showed that speakers produced the preferred property colour more often (using either colour+object or colour+border+object, $p < .05$, logit mixed effect models) when it was fully discriminatory than when it was only partly discriminatory, and a comparison between C non – B non and C full – B non showed that colour was also used more when it was fully discriminatory than when it was non-discriminatory. In contrast, there was no evidence that speakers preferred partly discriminatory colour over non-discriminatory colour: There was no difference between C part – B part and C non – B non or between C part – B full and C non – B full. Finally, discriminatory power of the less preferred border property had no effect on how often speakers used border in their descriptions, regardless of whether it was fully, partly or non-discriminatory.

Experiment 2 was conducted online and participants were told that they had to describe a target object to identify it from the other objects. Despite the absence of a listener, speakers produced more overspecified colour+object and colour+border+object descriptions than in Experiment 2. As in Experiment 1, colour was preferred to border. However, there was no evidence that speakers took into account discriminatory power, that is, property selection was unaffected by whether colour or border were fully, partly or non-discriminatory.

We conclude that speakers only take into account discriminatory power in the presence of a listener. Even then, discriminatory power plays a limited role: Speakers take into account whether a salient/preferred property (i.e., colour) is fully discriminatory, but perhaps surprisingly, it does not matter whether a property is partly discriminatory or not discriminatory at all. No concept selection model fits completely with these findings, and they will need to be further specified to account for them.

Figure 1: Conditions (arrow indicates target)

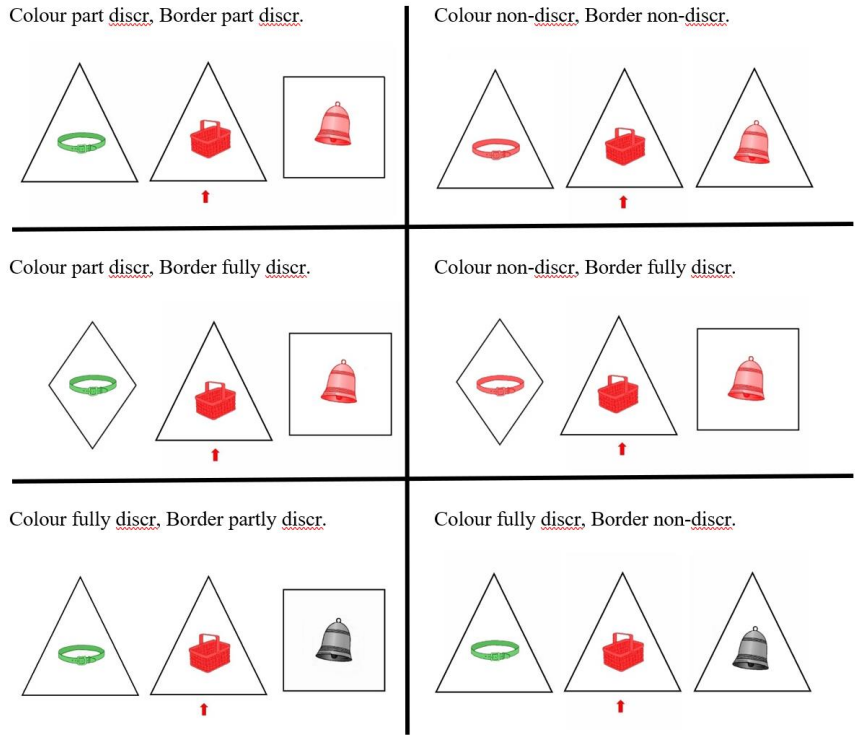
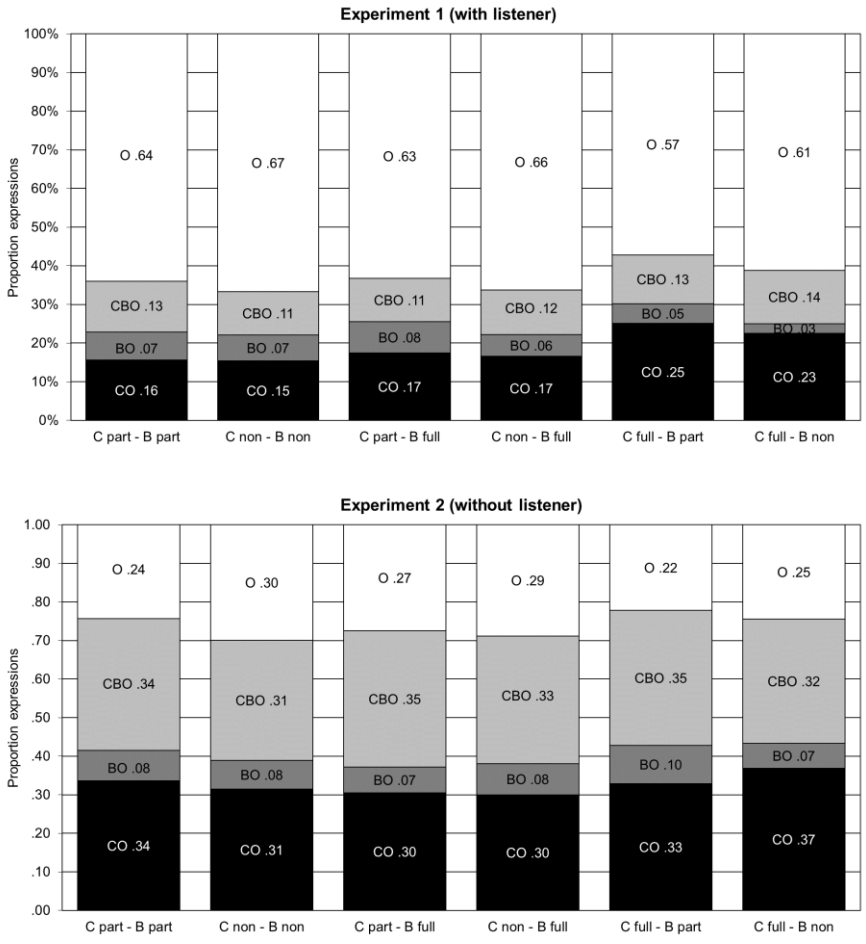


Figure 2: Results



C = colour, B = border, CO = colour+object expression, BO = border+object expression, CBO = colour+border+object expression, O = object only expression